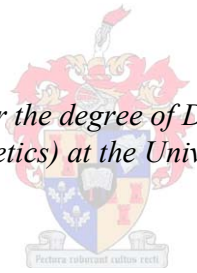
Investigation of the genetic aetiology of Parkinson's disease in South Africa

by
Rowena J. Keyser



Dissertation presented for the degree of Doctor of Philosophy (Human Genetics) at the University of Stellenbosch

Promoter: Dr. Soraya Bardien
Faculty of Health Sciences
Department of Biomedical Sciences

Co-promoter: Prof. Jonathan Carr
Faculty of Health Sciences
Division of Neurology

March 2011

Declaration

By submitting this dissertation electronically, 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.

Date: March 2011

Abstract

Parkinson's disease (PD), a neurodegenerative movement disorder characterized by resting tremors, bradykinesia, postural instability and rigidity, is due to a selective loss of dopaminergic neurons in the substantia nigra. Non-motoric symptoms include autonomic, cognitive and psychiatric problems. PD has been suggested to result from environmental factors, genetic factors or a combination of the two. Evidence has mounted over the last 13 years supporting the involvement of a significant genetic component. Mutations in the *parkin, PINK1, DJ-1, ATP13A2, SNCA*, and *LRRK2* genes have been conclusively associated with PD.

The aim of the present study was to establish the first study on the genetic etiology of PD in South African patients. Patients from the various South African ethnic groups with predominantly early-onset PD and/or a positive family history were recruited. Varying numbers of study participants (ranging from 88-205) were used for the different sections of this study depending on their availability at the time of the experiments and the specific clinical criteria applied. Mutation screening was conducted using High-resolution melt (HRM) analysis, DNA sequencing and multiplex ligation-dependent probe amplification (MLPA).

HRM analysis and sequencing of the known PD genes identified the following mutations: *parkin* (T113fsX163), *PINK1* (Y258X), and *LRRK2* (G2019S and R1441C). Using haplotype analyses, the five South African *LRRK2* G2019S-positive patients were found to share a common ancestor with other G2019S haplotype 1-associated families reported worldwide.

Two commercially available MLPA kits, SALSA P051 and P052, were used to assess the study participants for exon dosage mutations. Exonic deletions and insertions in *parkin* were identified in five patients. In addition, a family with a whole-gene triplication mutation of *SNCA* was identified. This is the 4th family worldwide to have this specific mutation which leads to a severe phenotype with autonomic dysfunction and early-onset dementia.

The CAESAR (CAndidatE Search And Rank) bioinformatic program was used to select novel candidate genes for PD. CAESAR produced a ranked list containing known PD causing genes as well as novel candidates. The *MAPT* and *SNCAIP* genes were selected from the list of ten highest scoring genes. HRM analysis identified novel sequence variants in both genes with unknown functional significance that warrants further study.

A novel 16bp deletion (g.-6_+10del) in the promoter region of *DJ-1* was identified in one PD patient. The functional significance of this variant was investigated using a Dual-Luciferase Reporter assay. The variant was found to significantly reduce luciferase activity in two separate cell lines, HEK293 and BE(2)-M17 neuroblastoma cells, both with and without oxidative stress ($p<0.0001$), and we proposed that the 16bp sequence might be important in transcriptional regulation of *DJ-1*. In addition, the activity of three transcription factors (AhR, ARNT and HIF-1) with binding sites within the deletion sequence may be influenced by the variant.

In conclusion, mutation screening resulted in the identification of mutations in six patients in *parkin*, six patients in *LRRK2*, one patient in *PINK1* and one patient in *SNCA*. In addition, a number of novel sequence variants were identified with unknown functional significance. Investigating the genetic basis of PD in the unique South African ethnic groups has shown that the known PD associated genes play minor roles in causing the disease in this population which indicates the possible involvement of other as yet unidentified PD genes. Innovative bioinformatic and wet bench experimental strategies are therefore urgently needed to identify new candidate genes for PD.

Opsomming

Parkinson se siekte (PS), ‘n neurodegeneratiewe bewegings-siekte, gekarakteriseer deur rustende spiersametrekkings, bradykinesia, posturale onstabiliteit en rigiditeit, onstaan as gevolg van geselekteerde verlies van dopaminergiese neurone in die substantia nigra. Nie-motoriese simptome sluit in outonome, kognitiewe en psigiatriese afwykings. Dit is voorgestel dat PS ontwikkel as gevolg van omgewings- en genetiese faktore of ‘n kombinasie van die twee. Daar was ‘n toename in bewyse vir die verantwoordelikheid van die genetiese komponent oor die afgelope 13 jaar. Mutasies in die *parkin, PINK1, DJ-1, ATP13A2, SNCA*, en *LRRK2* gene word met PS geassosieer.

Die doel van hierdie studie was om vir die eerste keer die genetiese etiologie van PS in Suid-Afrikaanse pasiënte te ondersoek. Pasiënte van die verskillende Suid-Afrikaanse etniese groepe, met hoofsaaklik vroeë-aanvang PS en/of ‘n positiewe familie-geskiedenis, was gebruik. Wisselende getalle van studie-deelnemers (van 88-205) was gebruik vir die verskillende dele van die studie, afhangende van hul beskikbaarheid op die tyd van die eksperimente en die spesifieke kliniese kriteria wat van toepassing was. Mutasie-analiese was uitgevoer deur middel van Hoë-resolusie smelting (HRS)-analiese, DNS volgorde-bepaling en multipleks ligasie-afhanklike ‘probe’ amplifikasie (MLPA).

HRS-analiese en DNS volgorde-bepaling van die bekende PS gene het die volgende mutasies deïdentifiseer: *parkin* (T113fsX163), *PINK1* (Y258X), en *LRRK2* (G2019S en R1441C). Haplotiepe-analiese het gevind dat vyf Suid-Afrikaanse *LRRK2* G2019S patiente ‘n gemeenskaplike voorvader deel met ander wêreldwyd gerapporteerde *LRRK2* haplotiepe 1-geassosieerde families.

Twee kommersieel beskikbare MLPA ‘kits’, SALSA P051 en P052, was gebruik om die deelnemers te toets vir exon-dosis mutaties. Exon-delesies en invoegings in *parkin* was gevind in vyf patiente. ‘n Familie met ‘n volle geen triplikasie van *SNCA* was gevind. Dit is die 4de familie wêreldwyd wat die spesifieke mutasie het en dit lei tot ‘n erge fenotiepe met outonomiese afwykings en vroeë-aanvang dementia.

Die ‘CAESAR (CAndidatE Search And Rank)’ bioinformatiese program was gebruik om nuwe kandidaat PS gene te selekteer. Die program het ‘n lys kandidaat gene, wat beide bekende

geassosieerde en nuwe bevat, opgelewer. Die *MAPT* en *SNCAIP* gene was gekies uit tien gene met die hoogste tellings. HRS analiese het nuwe DNS volgorde variante in beide gene gevind. Die funksies van die variante is tans onbekend en moet verder ondersoek word.

'n Onbekende 16bp delesie (g.-6_+10del) in die promotor area van *DJ-1* was gevind in een PS patient. 'n Dubbel-lusiferase rapporteerder eksperiment was uitgevoer om die funksie van die variant te ondersoek. Die variant het die lusiferase-aktiwiteit aansienlik verlaag in twee afsonderlike sel lyne, HEK293 en BE(2)-M17 neuroblastoma selle, met en sonder oksidatiewe spanning ($p<0.0001$). Dit was voorgestel dat die 16bp volgorde dalk belangrik kan wees vir transkripsionele regulasie van *DJ-1*. Die variant mag dalk ook die aktiwiteit van drie transkripsie faktore (AhR, ARNT and HIF-1) met bindings plekke in die delesie- volgorde, beïnvloed.

Ter afsluiting, mutasie analiese het gelei tot die identifikasie van mutasies in ses patiente in *parkin*, ses patiente in *LRRK2*, een patient in *PINK1* en een patient in *SNCA*. 'n Aantal nuwe variante was gevind met ombekende funksies. Ondersoek van die genetiese basis van PS in die uniek Suid-Afrikaanse etniese groepe het gevind dat die bekende PS gene nie 'n groot rol speel in die ontwikkeling van die siekte in die populasie nie. Dit is moontlik dat ander onbekende PS gene hier verantwoordelik is vir die siekte. Dit is dus belangrik om innoverende bioinformatiese en eksperimentele strategieë toe te pas om nuwe kandidaat-gene, vir PS, te identifiseer.

Table of Contents

Acknowledgments

I would like to express my sincere gratitude to the following people and organizations that contributed to the completion of my PhD degree:

- Dr. Soraya Bardien and Prof. Jonathan Carr, for your guidance and support during this project. Thank you for everything that I have learned from you and for giving me the opportunity to conduct this study with your guidance;

- My parents, for all your support, encouragement and prayers throughout my university studies. Thank you for being there for me. It would not have been possible without you;

- My family and friends, for your encouragement throughout this project;

- Ekow Oppon, for your love and support through everything. You inspired me and your encouragement helped me to keep on going. Thank you. Medaase, Mido Wu;

- Coworkers in the laboratory, for all your support and practical help throughout this project. Thank you for everything my friends;

- Dr Suzanne Lesage and Prof Alexis Brice for accommodating me in your laboratory in Paris, France. Celine Dupuits, for teaching me and making me feel welcome. Merci Beaucoup;

- The Oppenheimer Memorial Trust, Movement Disorder Society and Dean's Travel Fund (Stellenbosch Univeristy) for travel grants to attend the Movement Disorder congress and to spend two months on a research visit to Paris, France;

- Prof Paul van Helden, for providing financial support;

"Ask the Lord to bless your plans and you will be successful in carrying them out"

Prov. 16:3

List of Abbreviations

AAO:	Age at onset
AD:	Alzheimer's disease
ad:	Autosomal dominant
Amp^r:	Ampicillin resistance
ALP:	Autophagy-lysosome pathway
ALS:	Amyotrophic lateral sclerosis
ANK:	Ankyrin repeat
ar:	Autosomal recessive
ARM:	Armadillo
ATP13A2:	ATPase type 13A2
BCL2L1:	Bcl-2-like protein 1
C106:	Cysteine residue 106
CAESAR:	Candidate search and rank program
CBS:	Corticobasal syndrome
CDCrel-1:	Cell division control-related protein-1
COMT:	Catechol-O-methyl transferase
COR:	C-terminal of Roc
DAT:	Dopamine transporter
DBS:	Deep brain stimulation
DJ-1:	Oncogene DJ-1
dsDNA:	Double stranded DNA
EO:	Early-onset
ETC:	Electron transport chain
FBXO7:	F-box only protein 7
FTDP:	Frontotemporal dementia with parkinsonism
GBA:	β glucocerebrosidase
GD:	Gaucher's disease
GIGYF2:	GRB10 interacting GYF protein 2
GWAS:	Genome-wide association studies
HD:	Huntington's disease
HEK293:	Human embryonic kidney
HRM:	High-resolution melt
HTRA2:	HtrA serine peptidase 2
IBR:	In-between-RING
KRS:	Kufor-Rakeb syndrome
LBs:	Lewy bodies
L-DOPA:	Levodopa
LO:	Late-onset
LRR:	Leucine rich repeat
LRRK2:	Leucine-rich repeat kinase 2
MAO-B:	Monoamine-oxidase-B
MAPKKK:	Mitogen-activated protein kinase kinase kinase
MAPT:	Microtubule-associated protein tau
MCS:	Multiple cloning site
MLPA:	Multiplex ligation-dependent probe amplification
$MPDP^+$:	1-methyl-4-phenyl-2,3-dihydropyridinium
MPP^+:	1-methyl-4-phenyl-pyridinium
MPTP:	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MTS:	Mitochondrial-targeting domain

NAC:	Non-amyloid-β component
O:	Ocular signs
OMIM:	Online mendelian inheritance in man
PARK2:	Parkin
PCR:	Polymerase chain reaction
PD:	Parkinson's disease
PET:	Positron Emission Tomography
PGC-1α:	Peroxisome proliferator-activated receptor-gamma coactivator -1, alpha
PINK1:	PTEN-induced kinase 1
PLA2G6:	Phospholipase A2, group VI
PolyPhen:	Polymorphism Phenotyping
PSP:	Progressive supranuclear palsy
RanBP2:	RAN binding protein 2
REM:	Rapid eye movement
RING:	Really Interesting New Gene
RNAi:	RNA interference
Roc:	Ras of complex proteins
ROS:	Reactive oxygen species
RPH:	Relative peak height
SANBI:	South African National Bioinformatics Institute
SCA2:	Spinocerebellar ataxia type2
SIFT:	Sorting Intolerant From Tolerant
SNCA:	α-synuclein
SNCAIP:	Synphilin-1
SNCB:	β-synuclein
SNpc:	Substantia nigra pars compacta
SNPeffect:	Single Nucleotide Polymorphism effect
SNPs3D:	Single Nucleotide Polymorphisms 3D
SSCP:	Single strand conformation polymorphism
ssDNA:	Single stranded DNA
TH:	Tyrosine hydroxylase
TM:	Transmembrane
Ubl:	Ubiquitin-like
UCHL1:	Ubiquitin carboxyl-terminal esterase L1
UPS:	Ubiquitin proteasome system
UTR:	Untranslated region
XRE:	Xenobiotic Response Elements

List of Figures

Chapter 6

Chapter 7

Chapter 8

Chapter 9

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 5

Chapter 6

Chapter 7

Chapter 9

Chapter 1

Introduction

1.1 Introduction

Parkinson's disease (PD) (OMIM#168600) is a debilitating neurodegenerative disorder, which is currently without cure, and significantly impairs the sufferer's motor skills. Neuropsychiatric disturbances, such as cognition, mood and behavior problems, are also associated with PD in addition to the abnormality of movement. PD is the second most common neurodegenerative disorder after Alzheimer's disease (AD), and is characterized by the progressive and selective loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc).

PD was first described in 1817 by the English physician Dr. James Parkinson (figure 1.1) as *Paralysis Agitans* in his essay entitled "An essay on the shaking palsy". About 60 years after the publication of the essay, the French neurologist Dr. Jean Martin Charcot was the first to truly recognize the importance of Parkinson's work and decided to name the disease after the English physician.

Figure 1.1 Photograph of Dr. James Parkinson (1755-1825)

The molecular pathways leading to onset of PD are unclear, but it is generally accepted that it may result from environmental factors (including exposure to neurotoxins), genetic factors (mutations in specific PD linked genes) or a combination of the two. The pathways involved in PD pathogenesis are proposed to include mitochondrial dysfunction, oxidative damage, abnormal protein accumulation and protein phosphorylation, all of which potentially affect dopaminergic neuronal function and survival (Thomas & Beal 2007; Cookson & Bandmann 2010). The discovery of genes involved in familial forms of PD has strengthened the role of

genetic factors in development of the disease. Discovery of these genes has made it possible to investigate the pathological mechanisms that lead to disease development. In addition, they have identified possible targets for the development of neuroprotective therapies.

1.2 The epidemiology of Parkinson's disease

1.2.1 Age at onset of PD

It has been found that increasing age is one of the strongest risk factors for developing PD (Marion 2001; Siderowf 2001). Since PD affects mainly individuals of older age, it has been shown to be more common in developed countries where life expectancy is greater. PD has been found to affect 1-2% of the population over the age of 65 years and increases to 5% in individuals 85 years and older (de Rijk *et al.* 2000; Fahn 2003). However, the age at onset of the disorder is widely variable and ranges from juvenile to very late in life. Individuals with an age at onset (AAO) before 20 years are considered to have juvenile-onset PD. Early-onset PD has been variably defined across studies as age at onset <40-50 years (Lucking & Brice 2000; Periquet *et al.* 2003; Mata *et al.* 2004) and individuals developing PD after 50 years of age are referred to as late-onset PD (Pankratz & Foroud 2007).

1.2.2 Gender variance

PD is more prevalent among men (19.0 per 100 000) than among women (9.9 per 100 000) (Van Den Eeden *et al.* 2003) and epidemiological and clinical data suggest that estrogen has neuroprotective properties against PD in women (Dluzen & McDermott 2000; Shulman 2007; Gillies & McArthur 2010). The potential mechanisms by which estrogen may act as a neuroprotectant include antioxidative functions, inhibition of the monoamine oxidase enzyme, activation of neurotrophins, and by increasing the blood flow to facilitate the clearance of potential neurotoxins from the brain (Dluzen & McDermott 2000). Furthermore, women are exposed to different environmental and occupational risk factors, which might influence their susceptibility to develop PD. There might also be gender specific genetic influences leading to fewer women being diagnosed with PD.

1.2.3 Geographical and ethnic variability in PD prevalence

PD occurs worldwide, but fewer cases are reported in Africa than in Europe or North America (Okubadejo *et al.* 2006). This could be due to genetic and environmental diversity, and different population strata. The prevalence of PD in Africa is lowest in individuals from western (Ghana, Liberia, Nigeria) and eastern (Ethiopia, Kenya, Somalia, Tanzania, Uganda) African countries (Okubadejo *et al.* 2006). In these countries less than 4% of the population are 60 years of age or older and the life expectancy is usually less than 57 years (Okubadejo *et al.* 2006). However, Africa is experiencing a demographic transition which will result in the population becoming older by the year 2015. Diseases mostly affecting the elderly, such as PD, could therefore become more common in these countries (Heligman *et al*., 2000).

The estimated prevalence rate of PD varies globally from 7 per 100 000 to 657 per 100 000 individuals per year (Tekle-Haimanot *et al.* 1990; Zhang & Roman 1993; Melcon *et al.* 1997). The prevalence rate for Africa ranges from 7 to 43 per 100 000 individuals per year (Ashok *et al.* 1986; Tekle-Haimanot *et al.* 1990) (Attia *et al*., 1993). The prevalence rate for Europe has been found to be between 100 and 200 per 100 000 individuals (von Campenhausen S. *et al.* 2005), which is similar to prevalence rates reported for North Africa (Ashok *et al.* 1986) (Attia *et al*., 1993). The prevalence rate for China has been reported to be about 1700 per 1000 000 individuals, although this figure is for AAO $\geq$ 65 years, so is not really comparable (Zhang *et al.* 2005b).

Given the above, it is not surprising that the incidence of PD varies between different ethnic groups. The highest incidence has been reported to be among Hispanics (16.6 per 100 000), followed by non-Hispanic Whites (13.6 per 100 000), Asians (11.3 per 100 000), and Blacks (10.2 per 100 000) (Van Den Eeden *et al.* 2003). PD might be less common in Black and Asian people than in those of European origin, however, incidence reports have been conflicting and may be due to differences in case-ascertainment methods between studies (Alves *et al.* 2008).

1.3 Pathological characteristics

PD is characterized pathologically by progressive and profound loss of neuromelanin containing dopaminergic neurons in the SNpc situated in the midbrain (figure 1.2). The dopaminergic neurons have high levels of melanin and therefore give the substantia nigra (Latin for 'black substance') a unique appearance (figure 1.3 a and b). Widespread neurodegeneration in the central nervous system also occurs, with the SNpc being involved after involvement of more caudal regions of the brainstem and the olfactory bulb (Lang & Lozano 1998). Symptoms of PD appear when about 70-80% of dopaminergic neurons have been lost. The loss of these neurons leads to decreases in the levels of the neurotransmitter dopamine at the nerve terminals in the striatum and causes dysregulation of the motor circuits (nigrostriatal system) that project throughout the basal gangli (figure 1.4) (Gibb & Lees 1991; Lang & Lozano 1998). The consequences of the cell loss are impaired coordination of movement as well as autonomic, cognitive and psychiatric problems.

Another pathological characteristic is intra-cytoplasmic proteinaceous inclusions known as Lewy bodies (LBs) that occur in the surviving neurons of the SNpc and other brain regions (figure 1.3 c). The LBs are enriched in filamentous α-synuclein as well as other proteins that are in most cases highly ubiquitinated. The presence of LBs is a requirement for the definitive diagnosis of PD. However, LBs have been reported to be absent in a few cases with parkin or LRRK2-associated disease that have undergone autopsy (Hayashi *et al.* 2000; Zimprich *et al.* 2004)

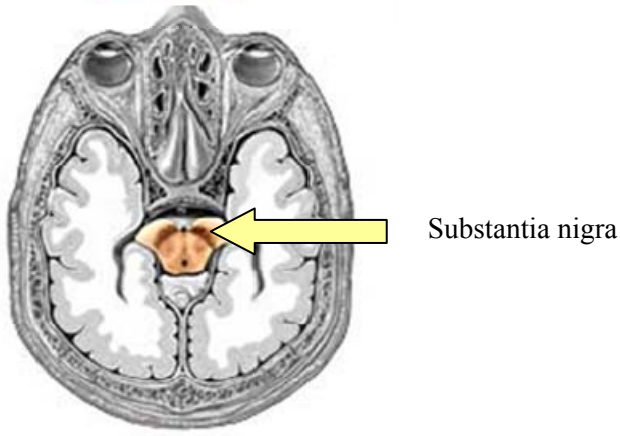


Figure 1.2 Illustration of the location of the substantia nigra in the brain. A transverse/horizontal section through the midbrain shows the position of the substantia nigra (taken from http://health.allrefer.com/health/parkinsons-disease-substantia-nigra-and-parkinsons-disease.html)

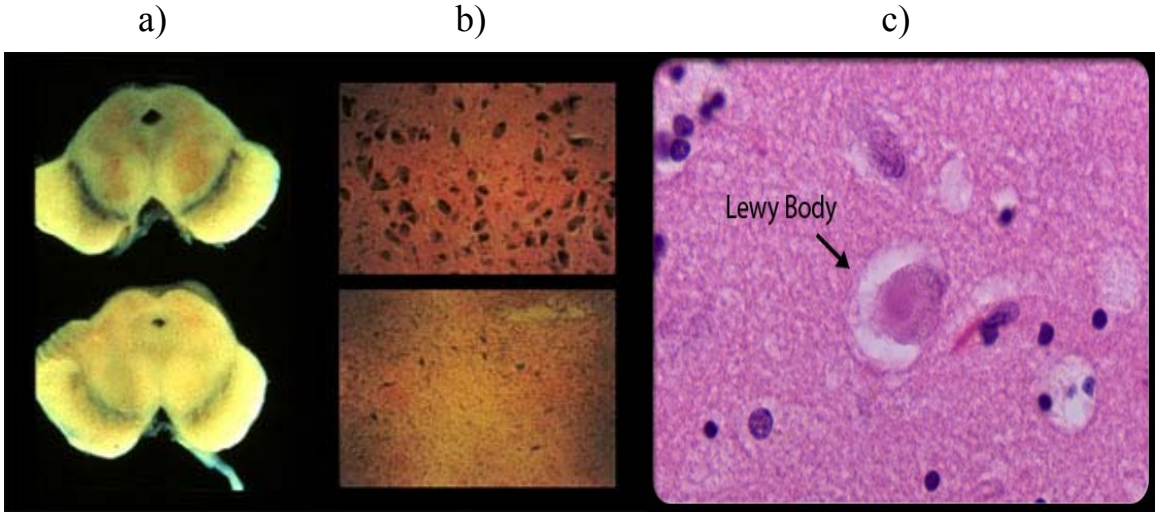


Figure 1.3 Microscopic images of the pathology of the substantia nigra. a) Normal (above) and abnormal (below) pigmentation of the neurons of the substantia nigra. Absence of pigmentation is indicative of dopaminergic cell death. b) Microscopic section of the substantia nigra showing normal (above) and abnormal (below) distribution of neuromelanin containing neurons. c) Microscopic image of the intra-cytoplasmic proteinaceous inclusions known as Lewy bodies (indicated by an arrow) (adapted from (Shulman 2007).

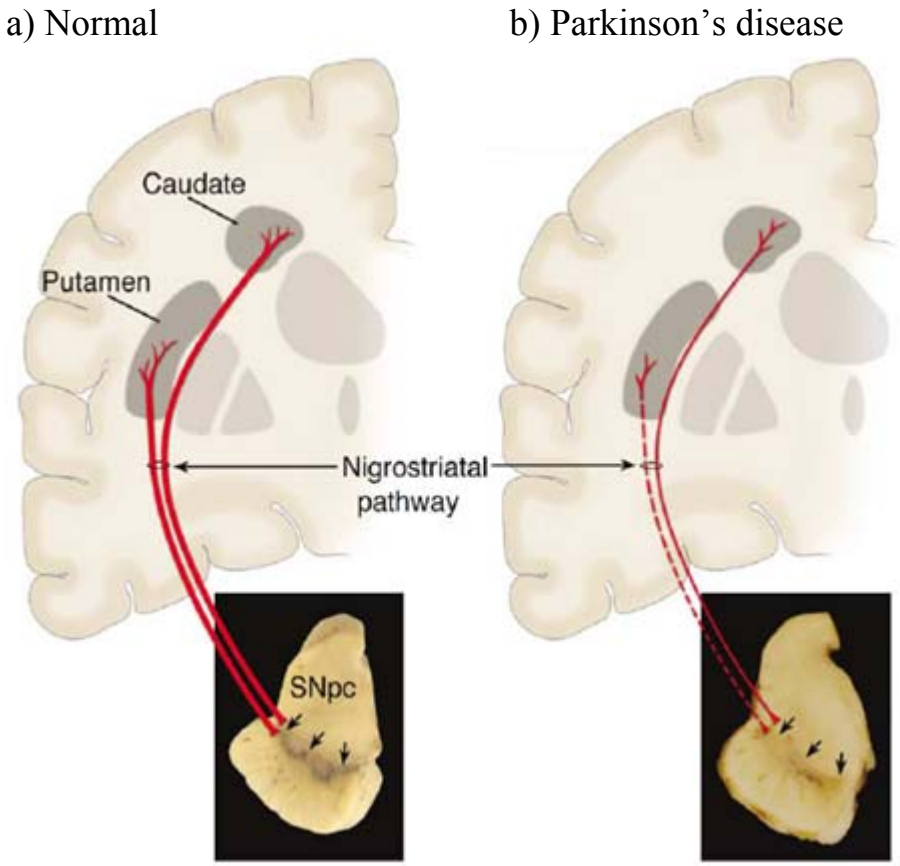


Figure 1.4 The nigrostriatal system in Parkinson's disease. a) Normal pathway showing dopaminergic nerve fibers (thick red line) extending from the substantia nigra (SNpc) to the striatum (i.e. putamen and caudate nucleus). Arrows indicate the nerve cell bodies in the SNpc. b) Abnormal pathway showing fewer nerve fibers extending from the SNpc to the putamen (dashed line) than to the caudate nucleus (thin red line). Arrows indicate the loss of cell bodies and depigmentation in the SNpc (taken from (Dauer & Przedborski 2003).

1.4 Clinical characteristics

1.4.1 Motor symptoms

PD patients develop severe motor disabilities about 4-6 years after being diagnosed with the disease. The four most common motor symptoms include resting tremor, rigidity, bradykinesia and postural instability.

Resting tremor

Resting tremor is in most cases the first neurological sign of PD, and is the one that motivates patients to visit a physician for a diagnosis. Tremors in PD patients do not impair daily activities of living to a large extent, since they predominantly occur when the individual is at rest and decreases during voluntary movement. Tremor normally occurs in the hand or foot on one side of the body and will involve the other side as the disease progresses (Siderowf 2001).

Rigidity

Rigidity (i.e. stiffness or inflexibility) is described as increased resistance to passive movements of a patient's limbs. Rigidity may cause pain and cramping and decreases the range of motion of the patient (Siderowf 2001).

Bradykinesia

Bradykinesia (*brady-*, 'slow', *kinisi*, 'motion') is characterized by slowness of movement. Bradykinesia, hypokinesia (reduction in movement amplitude) and akinesia (absence of normal unconscious movement) can result in paucity of normal facial expression, decreased volume of the voice, drooling, decreased size and speed of handwriting, and decreased stride length during walking (Dauer & Przedborski 2003).

Postural instability

Some patients lose normal postural reflexes that lead to falls which in some cases might be sufficiently severe to cause them to be confined to a wheelchair. A common symptom of PD is freezing, which is the inability to initiate a voluntary movement such as walking.

The quality of life of the PD patient is significantly impaired by these symptoms due to the fact that it takes longer to perform everyday tasks such as eating and getting dressed. Furthermore, several non-motor symptoms also occur and include constipation, olfactory dysfunction,

depression, anxiety, and Rapid Eye Movement (REM) sleep behavior disorder (Parkinson's Disease Foundation; http://www.pdf.org/en/symptoms).

1.4.2 Non-motor symptoms

PD patients develop cognitive abnormalities in which they become passive or withdrawn. They often develop depression with dementia occurring more frequently in older patients. It has been reported that 20-30% of PD patients develop dementia with 65% having this abnormality by the age of 85 years (Siderowf 2001). Psychiatric disturbances have been observed in 30% of PD cases. Patients may also experience hallucinations which are usually visual with delusions, and may also suffer from agitation or aggression. Some individuals become paranoid towards partners or other family members (Naimark *et al.* 1996). Sensory symptoms such as numbness, aching, tingling and muscle soreness have been observed (Snider *et al.* 1976). Constipation often occurs and may worsen with medication. Urinary incontinence, sexual dysfunction, excessive sweating, and sleep disturbances have also been observed with daytime drowsiness and insomnia (Partinen 1997; Davie 2008).

1.5 Genetic involvement in PD pathogenesis

Environmental factors were initially thought to be the predominant cause of PD. Some of these factors include exposure to pesticides and herbicides (e.g. dieldrin, paraquat and rotenone), neurotoxins (e.g. 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine, MPTP), rural living and well-water drinking, heavy metals (e.g. iron and manganese), and head trauma (Elbaz & Tranchant 2007). In the last 13 years, there has been a substantial increase in the evidence for the involvement of a genetic component and mutations have been found in six different genes in cases of autosomal recessive (ar) and dominant (ad) PD (Lesage & Brice 2009). Recessive forms of PD arise from mutations in the *parkin* (*PARK2*), PTEN-induced kinase 1 (*PINK1*), oncogene DJ-1 (*DJ-1*) and ATPase type 13A2 (*ATP13A2*) genes. The dominant forms of PD arise from mutations in the α-synuclein (*SNCA*) and leucine-rich repeat kinase 2 (*LRRK2*) genes. A summary of the various PD genes is shown in Table 1.1. To date, most PD-causing mutations (>150) have been found in *parkin* (Parkinson's disease Mutation Database, http://grenada.lumc.nl/LOVD2/TPI/home.php).

Table 1.1 Genes involved in Parkinson's disease

Locus	Gene	Chromosome	Form of PD	Inheritance	AAO (yrs)	Mutations
PD associated genes with conclusive evidence						
PARK1/ PARK 4	*SNCA*	4q21	EOPD	ad	20-85	A30P, E46K, A53T, genomic duplications/triplications
PARK2	*Parkin*	6q25.2-27	Juvenile and EOPD	ar	16-72	Point mutations, exonic rearrangements
PARK6	*PINK1*	1p35-36	EOPD	ar	20-40	Point mutations, rare large deletions
PARK7	*DJ-1*	1p36	EOPD	ar	20-40	Point mutations and large deletions
PARK8	*LRRK2*	12p11.2	LOPD	ad	32-79	7 pathogenic mutations, including the common G2019S
PARK9	*ATP13A2*	1p36	Juvenile ar Kufor-Rakeb syndrome and parkinsonism	ar	11-16	Point mutations
PD associated genes of unknown relevance/ inconclusive evidence						
PARK3	unknown	2p13	LOPD	ad	60s	Not identified
PARK5	*UCHL1*	4p14	LOPD	ad	55-58	One mutation in a single PD sibling pair
PARK10	unknown	1p32	unclear	ad	50-60	Not identified
PARK11	*GIGYF2*	2q36-q37	LOPD	ad	33-68	Seven missense variants
PARK12	unknown	Xq	unclear	unclear	-	Not identified
PARK13	*Omi/HTRA2*	2p13	unclear	unclear	49-77	Two missense variants
PARK14	*PLA2G6*	22q13.1	Juvenile ar levodopa-responsive dystonia-parkinsonism	ar	18-26	Two missense variants
PARK15	*FBXO7*	22q12-q13	EO ar parkinsonism-pyramidal syndrome	ar	10-19	Three point mutations
PARK16	unknown	1q32	unclear	unclear	-	Not identified
Not assigned	*SNCAIP*	5q23.1-q23.3	LOPD	unclear	63-69	R621C, various SNPs with association
Not assigned	*MAPT*	17q21.1	FTDP, O, CBS	ad	25-76	Haplotype H1, various SNPs with association
Not assigned	*SCA2*	12q24.1	EOPD	ad	45-59	Low-range interrupted CAG expansions in SCA2
Not assigned	*GBA*	1q21	EOPD	ar	40-50	Point mutations

AAO: age at onset; ad: autosomal dominant; ar: autosomal recessive; ATP13A2: ATPase type 13A2; CBS: corticobasal syndrome; EOPD: early-onset PD; FBXO7: F-box only protein 7; FTDP: frontotemporal dementia with parkinsonism; GBA:β glucocerebrosidase; GD: Gaucher's disease; GIGYF2: GRB10 interacting GYF protein 2; LOPD: late-onset PD; LRRK2: Leucine-rich repeat kinase 2; MAPT: microtubule-associated protein tau; O: ocular signs; HTRA2: HtrA serine peptidase 2; PD: Parkinson's disease; PINK1: PTEN-induced kinase 1; PLA2G6: phospholipase A2, group VI; SCA2: Spinocerebellar ataxia type2; SNCA: α-synuclein; SNCAIP: synphilin-1; UCHL1: ubiquitin carboxyl-terminal esterase L1 (Adapted from (Lesage & Brice 2009) and (Wider *et al.* 2010a).

1.5.1 Genes involved in autosomal recessive PD

1.5.1.1 *Parkin*

Parkin (OMIM 600116; chromosome 6q25) comprises 12 exons and encodes a 465 amino acid protein. The protein contains an N-terminal ubiquitin-like (Ubl) domain, a central RING (Really Interesting New Gene) domain (R0) and a C-terminal RING domain consisting of two RING finger motifs (R1 and R2) separated by an In-between-RING (IBR) domain (figure 1.5). The *Parkin* gene was first associated with PD in 1998 when exonic deletion mutations were identified in Japanese families with autosomal recessive juvenile Parkinsonism (Kitada *et al.* 1998). Since then, more than 150 different mutations have been identified including numerous point mutations and exonic rearrangements such as duplications and deletions (Periquet *et al.* 2003; Sun *et al.* 2006; Shadrina *et al.* 2007). Mutations in this gene have been found to be a common cause of early-onset Parkinsonism. However, mutations have also been found in cases of late-onset PD (>50-60 years of age) (Foroud *et al.* 2003; Sun *et al.* 2006). The mutations are not restricted to any of the functional domains of parkin and occur throughout the entire gene. Approximately half of the reported mutations fall in the category of exonic deletions or duplications (Hedrich *et al.* 2004). Individuals with *Parkin* mutations are reported to have slower disease progression, symmetrical onset, and in some cases, early-onset dystonia and levodopa responsiveness (Lohmann *et al.* 2003).

Single heterozygous *Parkin* mutations have been identified at higher frequencies than homozygous or compound heterozygous mutations. It has been proposed that heterozygous *Parkin* mutations could contribute to the development of PD by functioning as susceptibility factors. Heterozygous *Parkin* mutations have been shown to lead to later onset of PD (Foroud *et al.* 2003). Furthermore, fluorine-18-labelled dopa Positron Emission Tomography (PET) studies have found that dysfunction of dopaminergic neurons also occurred in individuals heterozygous for *Parkin* mutations, however the dysfunction was less severe than in homozygous individuals (Hilker *et al.* 2001). A possible mechanism by which heterozygous mutations might cause PD is haploinsufficiency where expression of a reduced amount of wild-type protein is insufficient to maintain normal function. It has also been suggested that some heterozygous mutations are more pathogenic than others because of a more severe consequence on structure and function of the protein (Klein & Lohmann-Hedrich 2007).

Parkin functions as an E3-type, E2 enzyme-dependent ubiquitin ligase and plays a role in the ubiquitin proteasome system (UPS) by ubiquitination of target proteins for degradation (Shimura *et al.* 2000). The UPS is the predominant proteolytic system for degradation of cytosolic, secretory, and membrane proteins (Hershko & Ciechanover 1998). Abnormalities in the UPS have been linked to neurodegenerative disorders such as PD, AD, Huntington's disease (HD), Prion diseases, familial Amyotrophic lateral sclerosis (ALS), and polyglutamine expansion disorders (Ciechanover & Brundin 2003). It has been observed that inactivation of parkin leads to reduction in UPS mediated degradation of target proteins, and that the accumulated proteins result in selective toxicity to dopaminergic neurons (Shimura *et al.* 2000; Yang *et al.* 2003; Sriram *et al.* 2005). Mutations in *Parkin* have been reported to cause alterations in the cellular localization, solubility, binding and ubiquitination properties of the parkin protein (Cookson *et al.* 2003; Gu *et al.* 2003; Sriram *et al.* 2005). Some mutations lead to compartmentalisation of parkin away from its normal cytoplasmic distribution as well as the site of enzymatic activity (Sriram *et al.* 2005).

Furthermore, several putative substrates of parkin have been identified including Sept5/CDCrel-1 (cell division control-related protein-1) (Zhang *et al.* 2000), synaptotagmin XI (Huynh *et al.* 2003) and RanBP2 (RAN binding protein 2) (Um *et al.* 2006). The accumulation of one or several of these substrates is suggested to contribute to the selective death of dopaminergic neurons.

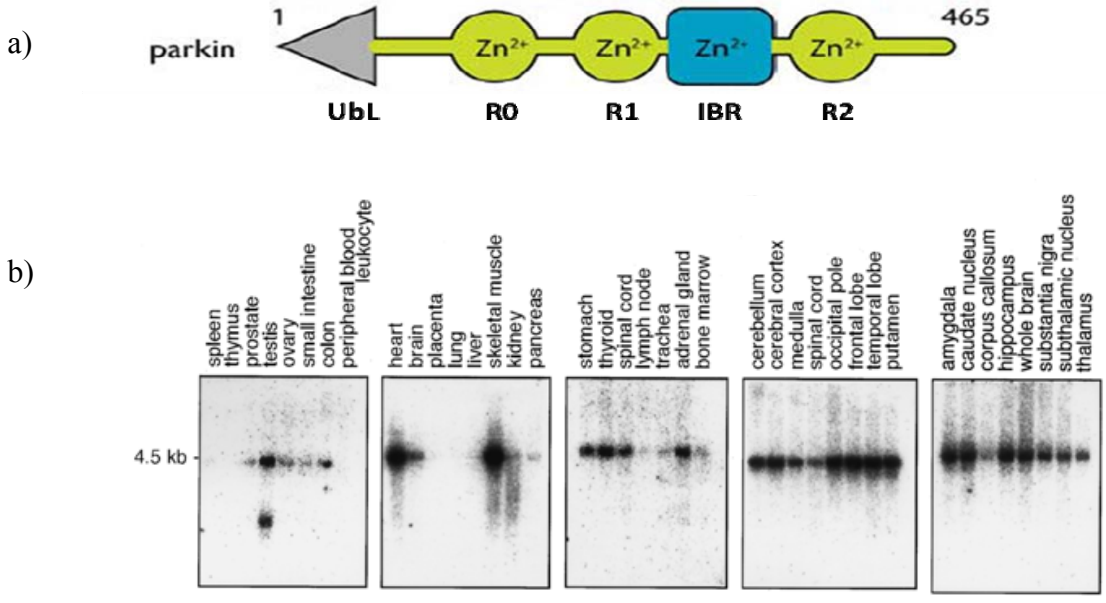


Figure 1.5 Schematic representation of the parkin protein and expression profile. a) Functionally important sites of parkin which contains the following domains: ubiquitin-like domain (Ubl), RING finger domains (R0-R2) and an In-between-RING domain (IBR) (taken from (Cookson & Bandmann 2010). b) Northern blot analysis of parkin in various human tissues showing ubiquitous expression (taken from (Kitada *et al.* 1998).

1.5.1.2 *PINK1*

PINK1 (OMIM 605909; chromosome 1p35-37) has 8 exons and encodes a highly conserved 581 amino acid protein that contains a mitochondrial-targeting domain (MTS), a putative transmembrane region (TM) and a conserved serine/threonine kinase domain (figure 1.6). The PINK1 protein, which is ubiquitously expressed in the human brain, has been shown to localize to the mitochondrion (Valente *et al.* 2004a; Gandhi *et al.* 2006). Mutations were first identified in 2004 in families of Italian and Spanish origin (Valente *et al.* 2004a) and have been found to be the second most common cause after the *parkin* gene of autosomal recessive early-onset PD. The frequency of *PINK1* mutations ranges from 1 to 7% in patients of different ethnicities (Rogaeva *et al.* 2004; Bonifati *et al.* 2005; Klein *et al.* 2005). Most *PINK1* mutations are located in the kinase domain and functional studies have shown that these mutations result in reduction in the enzymatic activity of PINK1 (Beilina *et al.* 2005; Sim *et al.* 2006). Little is known about the function of PINK1, but it is suggested to play a neuroprotective role against mitochondrial dysfunction and proteasomal-induced apoptosis (Valente *et al.* 2004a; Petit *et al.* 2005). PD patients with *PINK1* mutations have been reported to have atypical clinical features, such as psychiatric disturbances, dystonia at onset and sleep benefit (Valente *et al.* 2004b; Ephraty *et al.* 2007).

The involvement of *PINK1* in PD points towards two important factors. Firstly, this finding produced the first evidence that a kinase signaling pathway might be important in the pathogenesis of dopaminergic nigral cell death. Secondly, *PINK1* established a molecular link between the mitochondria and neurodegeneration in PD (Valente *et al.* 2004a; Abou-Sleiman *et al.* 2006). It has been shown that mitochondrial dysfunction plays an important role in the pathogenesis of PD (Onyango 2008) which will be discussed in Section 1.6.1. Interestingly, an interaction has been identified between parkin and PINK1, and it has been reported that they function together in the same pathway in maintaining mitochondrial integrity and function, with PINK1 functioning upstream of parkin (Park *et al.* 2006; Um *et al.* 2009).

a)

b)

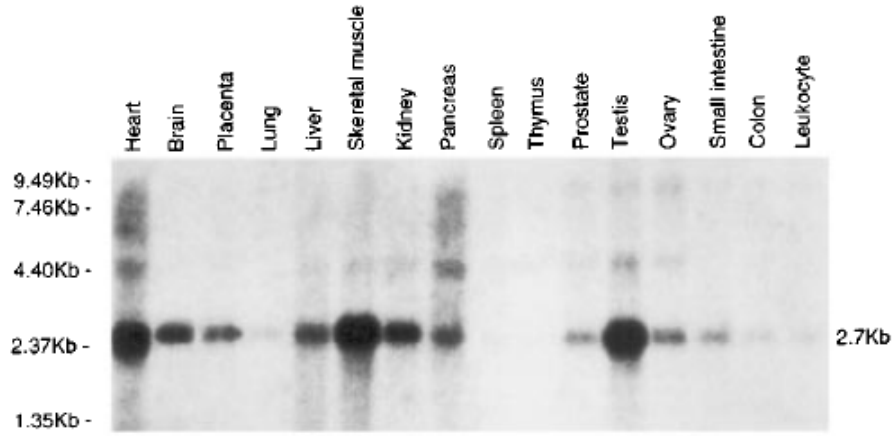


Figure 1.6 Schematic representation of the PINK1 protein and expression profile. a) The protein contains a mitochondrial-targeting domain (MTS), putative transmembrane region (TM) and a serine/threonine kinase domain (taken from (Cookson & Bandmann 2010). b) Northern blot analysis of PINK1 in various human tissues (taken from (Unoki & Nakamura 2001).

1.5.1.3 *DJ-1*

DJ-1 (OMIM 602533; chromosome 1p38) consists of 7 exons and encodes a highly conserved protein of 189 amino acids that belongs to the DJ-1/Thi/PfpI protein super family (figure 1.7a) (Nagakubo *et al.* 1997; Huai *et al.* 2003; Kahle *et al.* 2009). It is ubiquitously expressed in a variety of mammalian tissues including the brain (figure 1.7b; (Nagakubo *et al.* 1997). This gene was initially reported to be involved in oncogenesis and male rat infertility; however, it was later shown that *DJ-1* is associated with autosomal recessive early-onset PD, although such mutations are extremely rare (Nagakubo *et al.* 1997; Wagenfeld *et al.* 1998; Bonifati *et al.* 2003). Mutations include exon deletions, truncations, and homozygous and heterozygous point mutations, and have been shown to result predominantly in loss of function of the protein (Bonifati *et al.* 2003; Bonifati *et al.* 2004). Studies have shown that DJ-1 translocates to the mitochondria, where it is is proposed to play a role in protecting neurons from oxidative stress and protecting against mitochondrial damage (Canet-Aviles *et al.* 2004; Zhang *et al.* 2005a; Lev *et al.* 2008). Mitochondrial dysfunction leads to generation of reactive oxygen species (ROS), which causes damage to various cellular components.

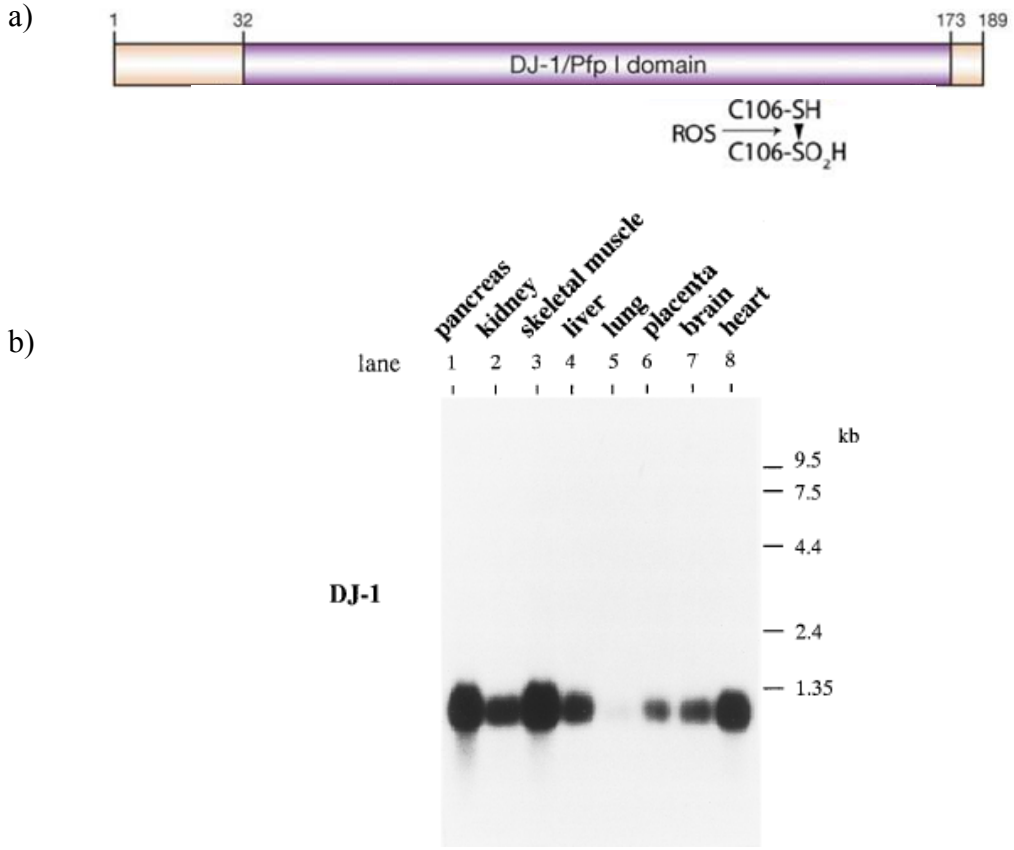


Figure 1.7 Schematic representation of the DJ-1 protein and expression profile. a) DJ-1 has only one domain called the DJ-1/Pfp I domain. The protein contains a critical Cysteine residue (C106) that can be modified in the presence of reactive oxygen species (ROS) to form a sulfinic acid (taken from (Moore *et al.* 2005). b) Northern blot analysis of DJ-1 in various human tissues (taken from (Nagakubo *et al.* 1997).

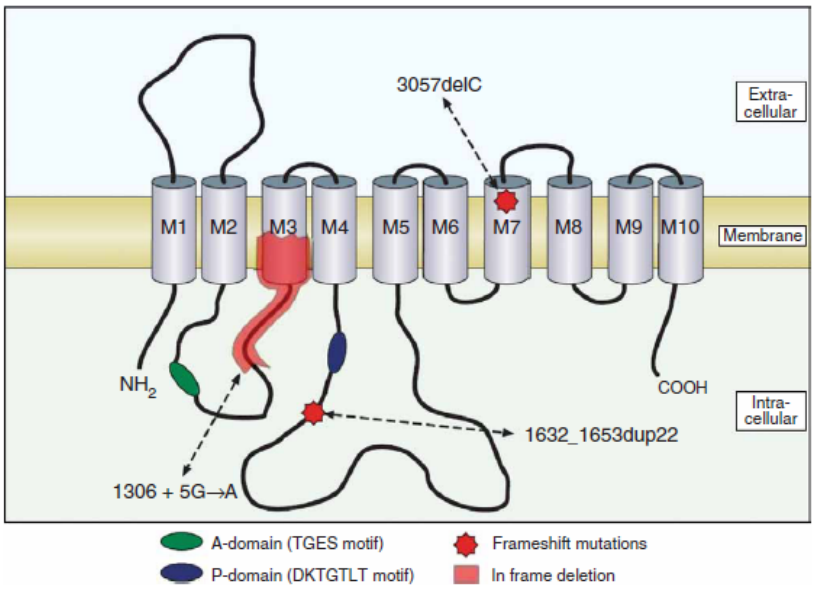
1.5.1.4 *ATP13A2*

ATP13A2 (OMIM 610513; chromosome 1p36) comprises 29 exons and encodes a large protein of 1,180 amino acids that contains 10 transmembrane domains which are located in lysosomal membranes (figure 1.8a) (Ramirez *et al.* 2006). Two motifs, TGES and DKTGTLT, are also found within the protein. ATP13A2 belongs to the P-type superfamily of ATPases that is involved in the transport of inorganic cations and other substrates across cell membranes (Schultheis *et al.* 2004). This protein is ubiquitously expressed and is also expressed throughout the brain (figure 1.8b), with highest levels in the SNpc, and is reported to be up-regulated in patients with late-onset PD (Ramirez *et al.* 2006). The involvement of this gene in PD was first discovered in 2006 with the identification of mutations in Jordanian and Chilean families with Kufor-Rakeb syndrome (KRS), which is a form of recessively inherited atypical parkinsonism (Ramirez *et al.* 2006). KRS has more widespread neurodegeneration, including dementia, and is clinically distinct from PD. Affected members of the Jordanian family harbored a homozygous 22 base pair duplication mutation (1632_1653dup22), which resulted in a frameshift and a

premature stop codon. Affected members of the Chilean family had a one base pair deletion in exon 26 (1019GfsX1021) that was inherited from the mother, and a splice site mutation (c.1305+5G>A) that was inherited from the father. The splice site mutation caused in-frame skipping of exon 13 which led to the removal of 111 nucleotides. These mutations were suggested to result in retention and proteasomal degradation of ATP13A2 in the endoplasmic reticulum as opposed to being inserted into the lysosomal membranes (Ramirez *et al.* 2006). Two additional mutations, F182L (Ning *et al.* 2008) and G504R (Di Fonzo *et al.* 2007), have also been reported. Interestingly, a genetic interaction has recently been reported in a study by Gitler and colleagues where the yeast homologue of human *ATP13A2* was found to interact with α-synuclein in yeast and subsequently suppress α-synuclein toxicity (Gitler *et al.* 2009).

The role played by *ATP13A2* in PD is however still questionable. A study conducted by Vilarino-Guell and colleagues identified 37 novel variants of which none segregated with the disease within kindreds (Vilarino-Guell *et al.* 2009). Case-control association studies gave negative results and ATP13A2 mRNA expression was not increased in PD brains compared to controls (Vilarino-Guell *et al.* 2009). They concluded that genetic variability in *ATP13A2* is unlikely to cause or influence the development of PD.

a)



b)

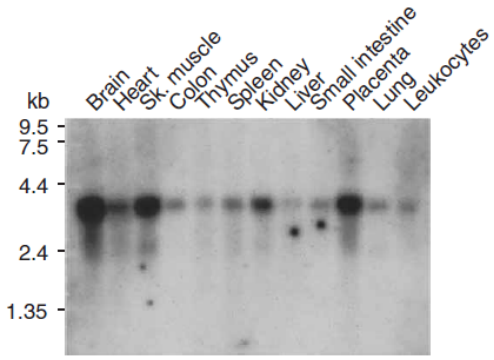


Figure 1.8 Illustration of the ATP13A2 transmembrane protein and expression profile. a) The protein contains 10 transmembrane domains (M1-10), a TGES motif and a DKTGTLT motif. The positions of frameshift mutations are indicated by red stars. The position of the in-frame deletion is shown by red shading. b) Northern blot analysis of ATP13A2 in various human tissues (taken from (Ramirez *et al.* 2006).

1.5.2 Genes involved in autosomal dominant PD

1.5.2.1 *SNCA*

α-Synuclein (*SNCA*) (OMIM 163890; chromosome 4q21) has 6 exons and encodes a 140 amino acid protein consisting of an N-terminal amphipathic region, a non-amyloid-B component (NAC) domain in the middle region, and an acidic C-terminal region (figure 1.9) (Jo *et al.* 2000). The protein is abundantly expressed throughout the brain with the highest levels reported in deeper layers of the cerebral neocortex, the hippocampus and the SNpc (Solano *et al.* 2000). α-Synuclein is a presynaptic protein and has been proposed to function in synaptic vesicle recycling, storage and compartmentalization of neurotransmitters and associates with vesicles and membrane structures (Abeliovich *et al.* 2000; Yavich *et al.* 2004). Of great importance, the protein has been found to be a major component of Lewy bodies and Lewy neurites in PD as well as other α-synucleinopathies (Spillantini *et al.* 1997).

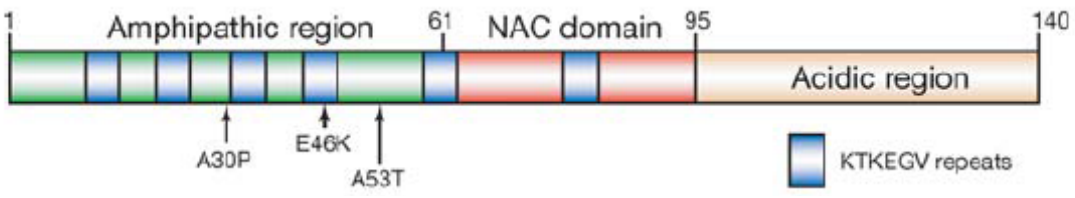
The involvement of the α-synuclein gene in PD was first reported in 1997 with the identification of an A53T substitution mutation in a large kindred with autosomal dominant PD (Polymeropoulos *et al.* 1997). Another two rare point mutations have also been found and include A30P (Kruger *et al.* 1998) and E46K (Zarranz *et al.* 2004). A30P, E46K and A53T mutant proteins have been shown to display an increased propensity for self-aggregation and oligomerization into protofibrils in comparison to wild-type proteins (Conway *et al.* 1998). Whole-gene multiplications (duplications or triplications) have also been identified in the α-synuclein gene and lead to over-expression of the protein. Duplication mutations of this gene give rise to a classical PD phenotype, whereas those affected with triplication mutations have

earlier onset, faster disease progression, marked dementia and frequent dysautonomia (Singleton *et al.* 2003; Chartier-Harlin *et al.* 2004; Ibanez *et al.* 2004; Ross *et al.* 2008a). Therefore, the higher the expression levels of α-synuclein, the more malignant the PD phenotype, which suggests that there is more widespread neurodegeneration in patients with higher levels of α-synuclein expression. This form of mutation is rare and less than five families worldwide have been reported to have an α-synuclein triplication, whereas duplication mutations have been reported in at least twelve families. Both recombination and duplication mechanisms have been shown to lead to α-synuclein multiplication (Ross *et al.* 2008a).

A susceptibility factor (REP1, D4S3481) for PD was identified in the promoter region of α-synuclein. REP1 is a dinucleotide repeat sequence that promotes normal gene expression (Touchman *et al.* 2001). It was found that REP1 allele length variability is associated with an increased risk for PD (Maraganore *et al.* 2006). A 263bp allele was associated with an increased risk for PD, while a 259bp allele was associated with a reduced risk for PD.

One of the mechanisms for the toxic effect of overexpression of wild-type or mutant α-synuclein has been proposed to be inhibition of the function of the proteasome in the UPS (Tanaka *et al.* 2001). In studies using Drosophila, it was shown that high levels of α-synuclein lead to abnormal protein aggregation and neurotoxicity in dopaminergic neurons (Feany & Bender 2000). Studies in mice have shown that overexpression of the protein led to the development of PD related features with findings of mislocalization and accumulation of insoluble α-synuclein in neurons in the neocortex, hippocampus and SNpc, as well as loss of dopaminergic terminals in the striatum, with associated motor disturbances (Masliah *et al.* 2000; Giasson *et al.* 2002).

a)



b)

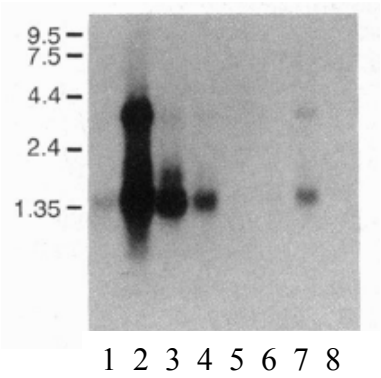


Figure 1.9 Schematic representation of the α-synuclein protein and expression profile. The protein contains an N-terminal amphipathic domain which contains six imperfect repeats (with a KTKEGV consensus motif), a non-amyloid-β component (NAC) domain, and an acidic C-terminal region. The positions of the three pathogenic missense mutations are indicated with arrows (taken from (Moore *et al.* 2005). b) Northern blot analysis of α-synuclein in various human tissues. 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 (Ueda *et al.* 1993).

1.5.2.2 *LRRK2*

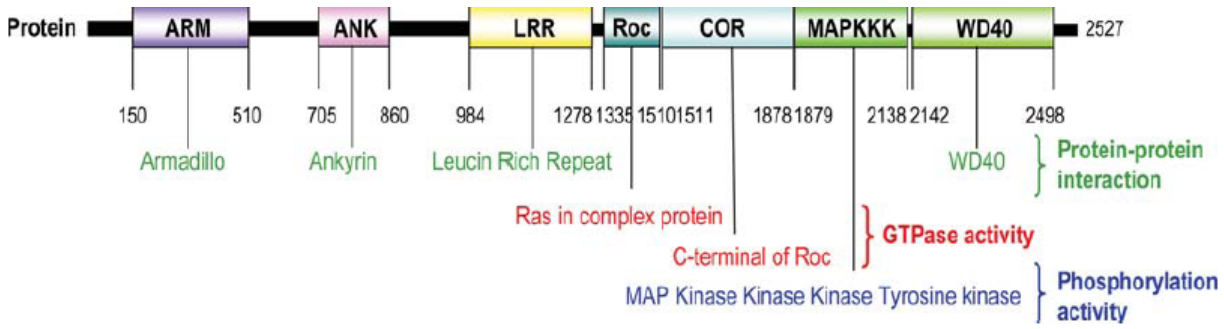
LRRK2 (OMIM 609007; chromosome 12q12) comprises 51 exons and encodes a 2,527 amino acid multi-domain protein (also known as dardarin) belonging to the ROCO protein family (Zimprich *et al.* 2004; Paisan-Ruiz *et al.* 2004). The physiological role of LRRK2 is unknown but the presence of multiple functional domains suggests involvement in a wide variety of functions (Thomas & Beal 2007). These functional domains include: Roc (Ras of complex proteins), COR (C-terminal of Roc), MAPKKK (mitogen-activated protein kinase kinase kinase) and WD40 (figure 1.10). The protein is suggested to be involved in the regulation of signal transduction cascades because of the presence of the Roc and MAPKKK domains (Guo *et al.* 2006).

LRRK2 is reported to be involved in approximately 10% of autosomal dominant familial PD and 3.6% of sporadic PD (Berg *et al.* 2005; Khan *et al.* 2005; Mata *et al.* 2005; Di Fonzo *et al.* 2006a; Johnson *et al.* 2007; Nichols *et al.* 2007; Xiromerisiou *et al.* 2007; Paisan-Ruiz *et al.* 2008). To date, more than 100 sequence variants have been identified in this gene of which seven (N1437H, R1441H, R1441C, R1441G, Y1699C, I2020T and G2019S) have been proven to be pathogenic. These mutations are located within the functional domains of the protein as well as in evolutionary conserved regions (Hedrich *et al.* 2006; Healy *et al.* 2008). The G2019S mutation, which occurs in the MAPKKK domain, is the most common mutation across diverse populations in both familial and sporadic PD and has been shown to increase the kinase activity of LRRK2 (West *et al.* 2005). The age-specific penetrance of the G2019S mutation has been

determined to be 28% at an age of 59 years, 51% at an age of 69 years, and 74% at an age of 79 years (Healy *et al.* 2008).

Three haplotypes (one major and two extremely rare haplotypes) have been identified that are present in carriers of G2019S. The first haplotype is most common and occurs in individuals of European, North and South African, and Ashkenazi Jewish origin (Lesage & Brice 2009). The second haplotype is rare and has been reported in five families of European origin (Zabetian *et al.* 2006a). The third haplotype which is also rare is most common in Japanese individuals but has also been reported in a Turkish family (Zabetian *et al.* 2006b; Pirkevi *et al.* 2009).

a)



b)

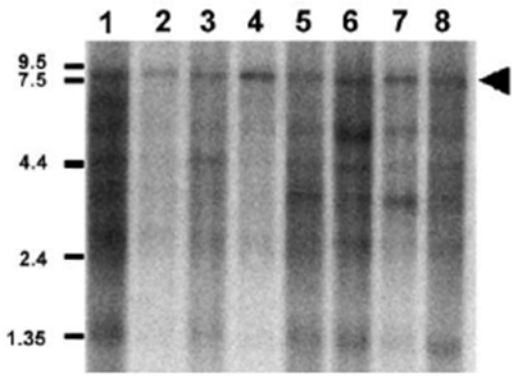


Figure 1.10 Schematic representation of the LRRK2 protein and expression profile. a) The protein contains the following domains: ARM, ANK, LRR, Roc, COR, MAPKKK, and WD40. Numbers indicate the amino acid positions (taken from (Lesage & Brice 2009). b) Northern blot analysis of LRRK2 in various human tissues. 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. The black arrowhead indicate the position of LRRK2 (taken from (Paisan-Ruiz *et al.* 2004).

1.5.3 Other genes associated with PD

In addition to the genes mentioned in Sections 1.5.1 and 1.5.2 a number of other genes have been reported to be linked to PD, but for many their involvement with this disorder is currently equivocal (Lesage & Brice 2009; Satake *et al.* 2009). Two of these genes, *SNCAIP* and *MAPT,* will be discussed in more detail as they have been selected as PD candidate genes in Chapter 6 of this study.

1.5.3.1 *SNCAIP*

Synuclein alpha interacting protein (*SNCAIP*) (OMIM 603779; chromosome 5q23.1-q23.3) has 11 exons and encodes a 919 amino acid protein (synphilin-1) that contains several protein-protein interaction domains including ankyrin repeat domains, a coiled-coil domain, and an ATP and GTP binding domain (figure 1.11a) (Engelender *et al.* 1999; Engelender *et al.* 2000). The protein is ubiquitously expressed throughout the human body including the SNpc, with highest levels in neurons (figure 1.11b; (Engelender *et al.* 1999). The importance of *SNCAIP* in PD was highlighted when it was found that two PD-linked gene products, α-synuclein and parkin, interact with synphilin-1 (Engelender *et al.* 1999; Chung *et al.* 2001). Furthermore, it was found that synphilin-1 is present in 80-90% of LBs in brain samples of PD patients (Wakabayashi *et al.* 2000).

Very few mutations in this gene have been found in PD patients and therefore it has not been conclusively linked to this disorder. Functional studies investigating the role played by an R621C variant that was identified in two sporadic German PD patients showed that the variant caused an increase in cell susceptibility to cellular stress (Marx *et al.* 2003). The R621C variant has been observed in control samples indicating that it is not likely to be pathogenic (Myhre *et al.* 2008). A study conducted by Li and colleagues found that synphilin-1 has protective effects *in vitro* (Li *et al.* 2010). They observed that overexpression of synphilin-1 protected cells against rotenone-induced cell death by means of reducing caspase-3 activation and poly (ADP-ribose) polymerase cleavage. Rotenone is an environmental toxin that inhibits mitochondrial complex I activity and it has been shown to induce dopaminergic neurodegeneration in rats and Drosophila (Betarbet *et al.* 2000; Coulom & Birman 2004). It is therefore suggested that *synphilin-1* might play a neuroprotective role in PD pathogenesis.

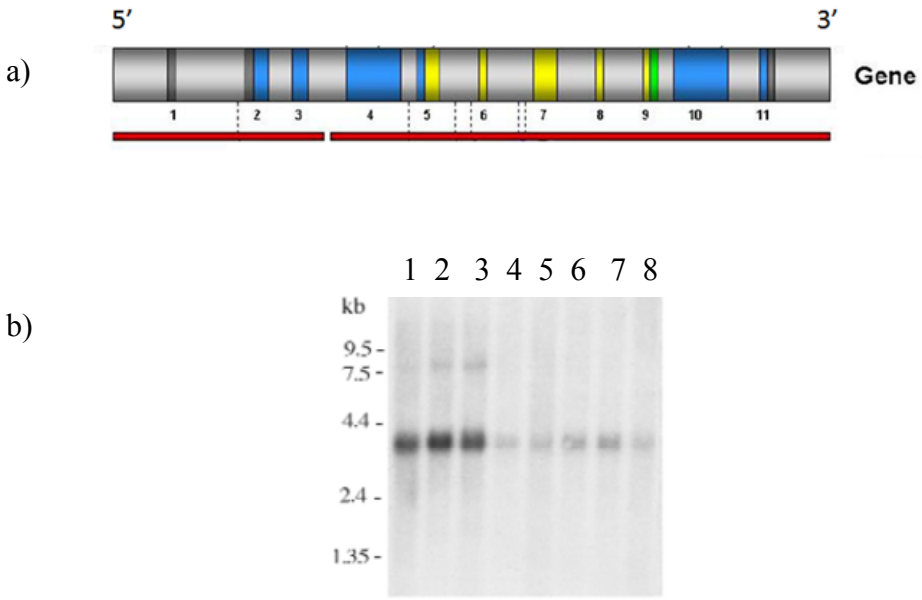


Figure 1.11 Illustration of the *SNCAIP* gene and expression profile. a) The positions of the ankyrin repeats (shaded in yellow), a coiled-coil domain (shaded in green) and translated regions (shaded in blue) are indicated. Numbers indicated the 11 exons of *SNCAIP* (taken from (Myhre *et al.* 2008). b) Northern blot analysis of synphilin-1 in various human tissues. 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 (Engelender *et al.* 1999).

1.5.3.2 *MAPT*

Microtubule-associated protein tau (*MAPT*) (OMIM 157140; chromosome 17q21.1) has 15 exons and encodes a 776 amino acid protein which is more frequently referred to as tau (figure 1.12). It has been found to be highly expressed in neurons and it is important for organizing and maintaining cell structure by modulating microtubules (Weingarten *et al.* 1975; Hirokawa 1994). Interactions between tau and microtubules occur via microtubule-binding repeat domains located in the carboxyl-terminus (Lee *et al.* 1989). Aggregation of the tau protein results in development of tauopathies and has been observed in several neurodegenerative disorders, such as Pick disease, AD, and disorders with parkinsonian features, such as progressive supranuclear palsy (PSP), corticobasal degeneration, and frontotemporal dementia with parkinsonism (FTDP-17) (Rademakers *et al.* 2004).

Genome-wide association studies (GWAS) recently revealed *MAPT* to be a risk factor for idiopathic PD (Simon-Sanchez *et al.* 2009; Edwards *et al.* 2010). Mutations in *MAPT* have been shown to be involved in autosomal-dominant FTDP-17 (Hutton *et al.* 1998). FTDP-17 patients with *MAPT* mutations present with personality, behavioral or cognitive changes that are associated with rapidly progressive Parkinsonism with poor response to levodopa treatment

(Wszolek *et al.* 1992). Mutations in *MAPT* account for ~5-10% of sporadic and 30% of familial frontotemporal dementia cases (Wider *et al.* 2010a). *MAPT* has a polymorphic inversion resulting in two main haplotypes: H1 and H2. H1 has been shown to be associated with an increased risk for PD (Farrer *et al.* 2002; Zabetian *et al.* 2007; Tobin *et al.* 2008; Wider *et al.* 2010b); however the functional variant within this haplotype still needs to be identified.

Overexpression of tau has negative effects on neurons that might play a role in the development of PD. In neuronal cell cultures increased tau inhibited intracellular transport along microtubules resulting in disruption of cell function and increased vulnerability of the cells to oxidative stress (Stamer *et al.* 2002). Accumulation of neurofilaments, microtubules and organelles were observed in transgenic mice that overexpressed tau, and were sufficient to cause damage to central nervous system neurons (Spittaels *et al.* 1999). In addition, it has been found that tau promotes the assembly of α-synuclein into fibrils that could further aggregate into LBs (Giasson *et al.* 2003).

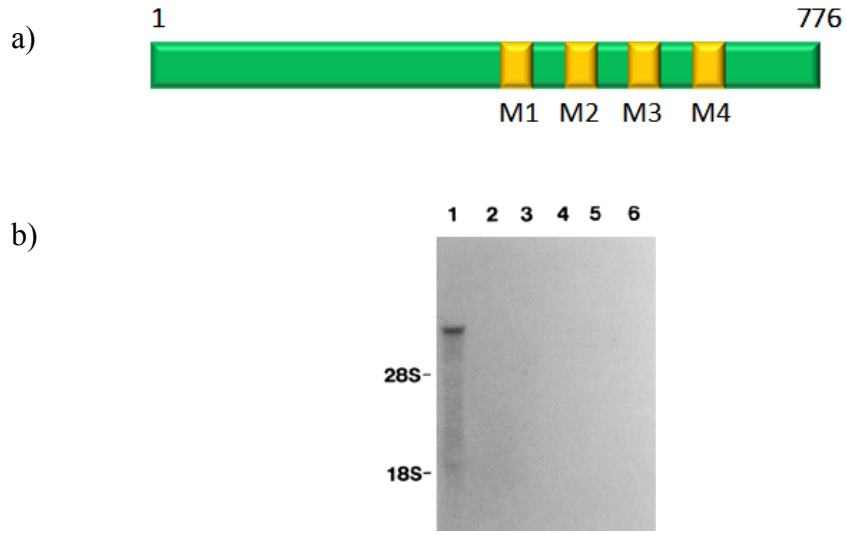


Figure 1.12 Schematic representation of the tau protein and expression profile. a) The four microtubule-binding repeat domains (M1-M4) are indicated. b) Northern blot analysis of tau showing expression in the brain only. Lane 1: brain, Lane 2: kidney, Lane 3: liver, Lane 4: spleen, Lane 5: stomach, Lane 6: thymus (taken from (Lewis *et al.* 1986).

1.5.4 PD susceptibility factors

Susceptibility factors are genetic variations that affect penetrance, age at onset, severity and progression of PD. A number of susceptibility factors have been shown to be associated with PD, and some of these will be briefly discussed. The REP1 variant which occurs in the promoter region of SNCA has been confirmed as a risk factor for PD and was discussed in section 1.5.2.1

(Maraganore *et al.* 2006). GWAS have identified association of variants in SNCA, *LRRK2* and *MAPT* with increased risk for PD (Simon-Sanchez *et al.* 2009). GWAS also showed association of the *MAPT* H1 haplotype and increased risk for PD (Farrer *et al.* 2002; Zabetian *et al.* 2007; Tobin *et al.* 2008; Wider *et al.* 2010b). Individuals who carried both the rs356219 SNCA SNP and the *MAPT* H1 haplotype have been shown to have double the risk of developing PD (Goris *et al.* 2007). The G2385R and R1628P variants in *LRRK2* have been found to confer susceptibility to PD but only in Asian populations (Di Fonzo *et al.* 2006b; Ross *et al.* 2008b). The β glucocerebrosidase (GBA) gene has also been found to be a susceptibility factor for PD in Ashkenazi Jews (Aharon-Peretz *et al.* 2004) and North Americans (Nichols *et al.* 2009) and is associated with an earlier age at onset. Also, the common A340T variant in *PINK1* has been found to contribute to the risk for late-onset PD in Chinese (Wang *et al.* 2006).

1.6 Molecular pathways implicated in PD pathogenesis

1.6.1 Mitochondrial dysfunction and oxidative stress

Evidence for the involvement of the mitochondria in neurodegeneration in PD emerged after a group of heroin addicts developed PD symptoms after accidental injection of a synthetic by-product of heroin production called MPTP (Langston *et al.* 1983). The active metabolite of MPTP is 1-methyl-4-phenyl-pyridinium ion (MPP^+) and has been found to selectively be taken up into dopaminergic neurons by means of the dopamine transporter. The method in which MPTP is metabolized in the brain is shown in figure 1.13 (Dauer & Przedborski 2003). This active metabolite inhibits mitochondrial complex-I catalytic activity in the electron transport chain (ETC) and results in increased oxidative stress as well as decreased energy production, leading to neuronal damage and death.

Oxidative stress is defined as an imbalance between ROS production and the antioxidant capacity of a cell and mitochondria have a central role in the generation of ROS. Mitochondrial dysfunction leading to increased ROS production can cause damage to various cellular components such as unsaturated lipids, proteins, and nucleic acids, and this has been implicated in various neurodegenerative disorders including PD, AD, and ALS (Lin & Beal 2006). Defects in the subunits and activity of mitochondrial complex-I have been observed in blood platelets and SNpc of PD affected patients (Keeney *et al.* 2006).

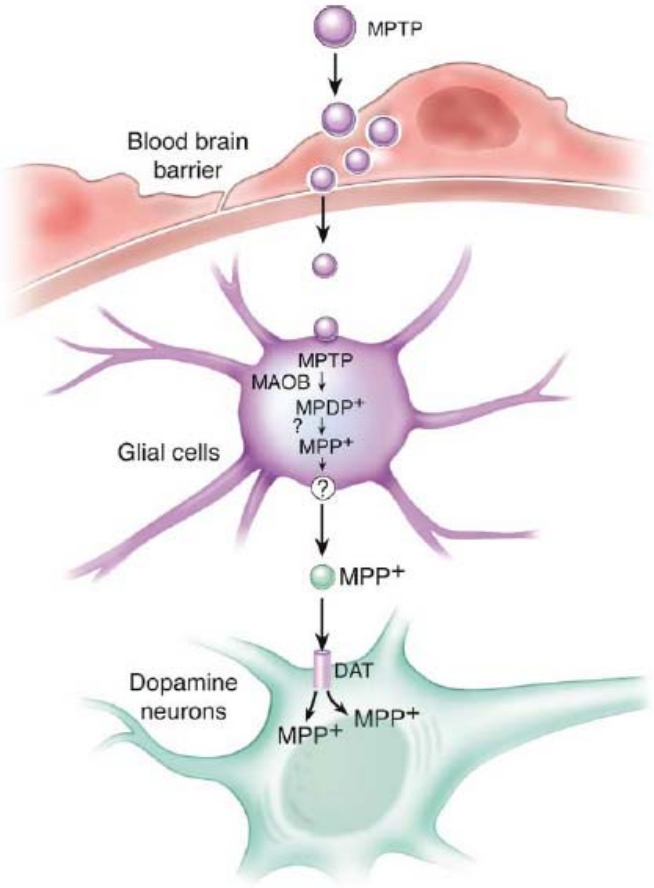


Figure 1.13 Illustration of MPTP metabolism. MPTP crosses the blood-brain barrier and is converted to $MPDP^+$ by MAO-B in non-dopaminergic cells. It is then converted to MPP^+ and released from the cell by an unknown mechanism. MPP^+ is then transported into dopaminergic cells via the dopamine transporter. $MPDP^+$: 1-methyl-4-phenyl-2,3-dihydropyridinium, MAO-B: Monoamine-oxidase-B, MPTP: 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine, MPP^+: 1-methyl-4-phenyl-pyridinium, DAT: dopamine transporter **(**taken from (Dauer & Przedborski 2003)**.**

As noted above, certain pesticides, such as rotenone and paraquat, have been found to produce similar pathological results by inhibiting mitochondrial complex-I activity (Lockwood 2000). Similarly, a complex-I inhibitor called annonacin, found in the tropical plant Annona muricata, has been suspected to be responsible for an atypical parkinsonsism in the French West Indies, and was found to promote dopaminergic neuronal death by impairing the process of energy production (Lannuzel *et al.* 2003).

The production of ROS may be a direct result of inhibition of the mitochondrial ETC or indirectly during the apoptotic process (programmed cell death) (Seaton *et al.* 1997). Mitochondria play an important part in controlling apoptosis, which involves the release of cytochrome c from the mitochondrial intermembrane space upon a specific trigger, in order to activate the cascade of caspases (cysteine proteases) responsible for degradation of the cell by

cleaving multiple cellular substrates (Gorman *et al.* 2000). In addition, changes in energy production by the mitochondria can induce apoptosis in neurons or increase their sensitivity to apoptosis (Gorman *et al.* 2000). Evidence has been found for the process of apoptosis in the SNpc of PD patients (Mochizuki *et al.* 1996).

Interestingly, many of the genes involved in PD (*parkin, PINK1, DJ-1*, α-synuclein and *LRRK2*) have been found to encode either mitochondrial proteins, or proteins associated with mitochondrial dependent cell death. The participation of the different PD genes in pathways encompassing the mitochondria as well as the UPS, which will be discussed in the following section, is illustrated in figure 1.14.

Studies with transgenic mice have shown that overexpression of α-synuclein impairs mitochondrial function, and can enhance the toxicity of MPTP with increased oxidative stress (Tanaka *et al.* 2001; Song *et al.* 2004). Parkin is known to associate with the outer mitochondrial membrane, where it protects against the release of cytochrome c and the activation of the caspases (Darios *et al.* 2003). Oxidized DJ-1 translocates to the mitochondria intermembrane space and matrix where it down-regulates the PTEN-tumor suppressor protein and protects cells form oxidative stress induced apoptosis (Kim *et al.* 2005). PINK1 localizes to the mitochondrial matrix and is proposed to protect against apoptosis (Petit *et al.* 2005). Furthermore, it has been shown that about 10% of LRRK2 also localizes to the mitochondria where it functions as a kinase (West *et al.* 2005). The role played by these different genes has reinforced the importance of mitochondrial dysfunction and oxidative stress as key mechanisms in PD pathogenesis.

1.6.2 The ubiquitin proteasome system and the autophagy-lysosomal pathway

The UPS is the major proteolytic system for the degradation of cytosolic, secretory and membrane proteins (Hershko & Ciechanover 1998). This system removes unwanted proteins that are no longer required by the cell. The UPS is important for several basic cellular processes, such as regulation of cell cycle and division, cellular differentiation and development, morphogenesis of neuronal networks, cellular responses to stress and extracellular effectors, and DNA repair (Ciechanover & Brundin 2003).

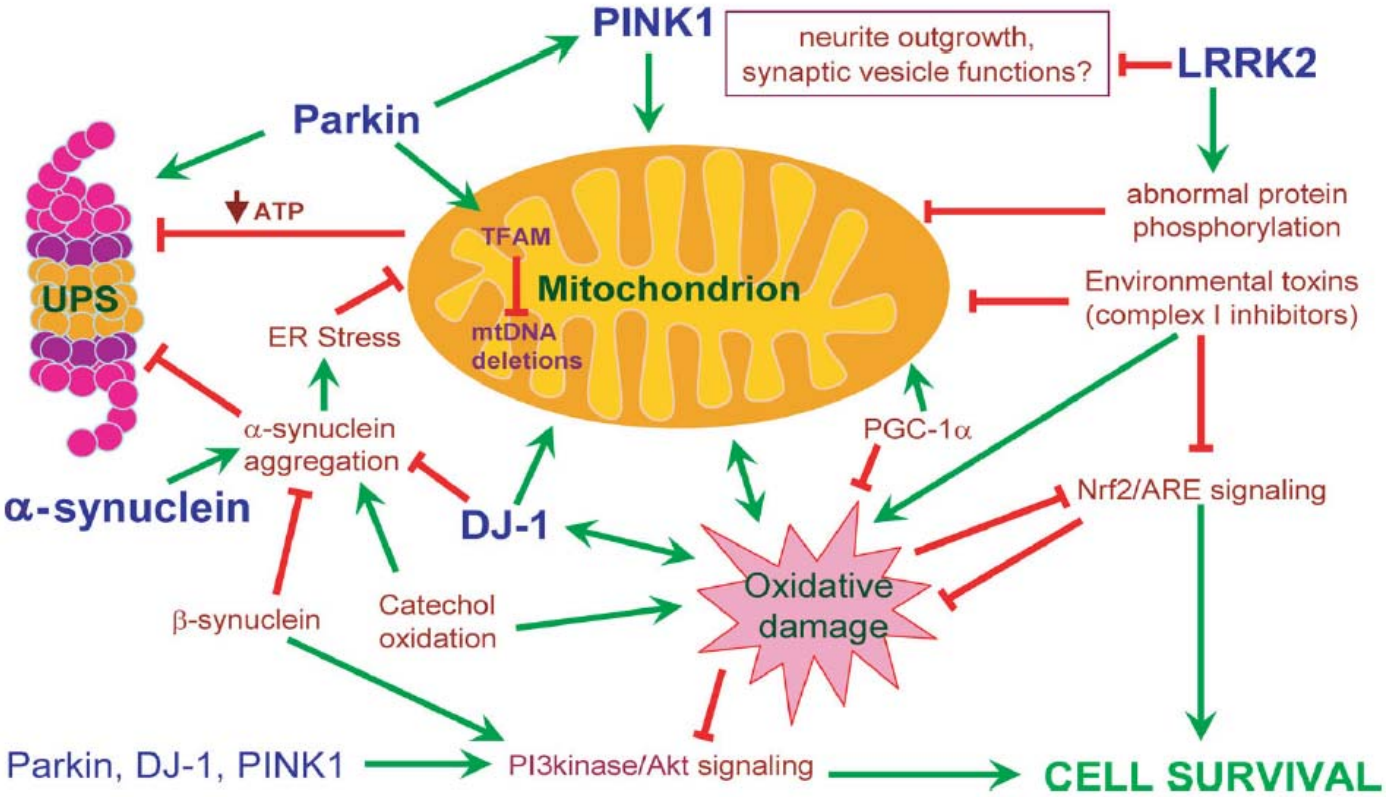


Figure 1.14 Molecular pathways involved in PD pathogenesis. The role played by parkin, PINK1, DJ-1, α-synuclein and LRRK2 is illustrated. The green arrows indicate promoting or activating effects. The red lines with blunt ends indicate inhibitory effects (taken from (Thomas & Beal 2007).

Target proteins that must be removed by the UPS are covalently tagged with ubiquitin, which is a 76 amino acid residue protein (Hershko & Ciechanover 1998). This occurs through the formation of an iso-peptide bond between the ε-amino group of a lysine residue of the substrate and the C-terminal carboxylate of ubiquitin, and normally leads to the formation of a polyubiquitin chain on the substrate (Hershko & Ciechanover 1998; Lim & Tan 2007). The E1 ubiquitin activating enzymes, E2 conjugating enzymes and E3 ligating enzymes are involved in this process of tagging proteins for degradation by the UPS and defects in any of them could lead to accumulation of protein degradation products.

The involvement of the UPS in PD was strengthened after the identification of mutations in the *parkin* gene (Kitada *et al.* 1998) which, as previously mentioned, plays an important role in the UPS as an E3-type, E2 enzyme-dependent ubiquitin ligase. Inactivation of parkin results in decreased UPS mediated degradation of target proteins, and the accumulation of these leads to selective toxicity of dopaminergic neurons (Shimura *et al.* 2000; Yang *et al.* 2003; Sriram *et al.* 2005). Parkin, PINK1 and DJ-1 interact to form a novel E3 complex, designated as the PPD complex, which regulates UPS-mediated protein degradation (Xiong *et al.* 2009). The complex was shown to play an important role in ubiquitination and degradation of un-/misfolded parkin substrates. The involvement of the PPD complex in PD is supported by the finding that the PD-pathogenic mutants of PINK1 and parkin impair the E3 ligase activity of the complex (Xiong *et al.* 2009).

Furthermore, there is evidence that the α-synuclein gene is also involved with dysfunction of the UPS, and PD pathogenesis. Studies in which α-synuclein was overexpressed showed a decrease in the function of the proteasome, resulting in abnormal protein aggregation and neurotoxicity in dopaminergic neurons (Tanaka *et al.* 2001). In addition, α-synuclein has been shown to be a major component of LBs (intracytoplasmic proteinaceous inclusions) in PD (Spillantini *et al.* 1997). Interestingly, MPTP, rotenone and paraquat have also been shown to cause abnormalities in the UPS by impairing proteasome function in rats and mice (Fornai *et al.* 2005; Betarbet *et al.* 2006).

The autophagy-lysosome pathway (ALP) is another system by which unwanted proteins are removed from the cell. Impairment of this process also leads to the accumulation and aggregation of proteins, resulting in cellular toxicity and neurodegeneration (Pan *et al.* 2008). ALP is suggested to be the major mechanism involved in the degradation of long-lived, stable proteins as well as the mechanism by which organelles (e.g. mitochondria) are recycled. ALP

dysfunction may result from failure of autophagosome formation or autophagosome fusion with lysosomes, deficiency of enzymes in lysosomes, dysfunction of the molecular chaperone or lysosomal membrane receptor. It has been shown that α-synuclein is also cleared through ALP (Webb *et al.* 2003). Impaired clearance of α-synuclein by abnormal functioning ALP could therefore lead to toxic accumulation of the protein and neurodegeneration. The role played by ALP in PD has been further strengthened by the identification of mutations in the *ATP13A2* gene which leads to insufficient lysosomal protein degradation (Ramirez *et al.* 2006; Di Fonzo *et al.* 2007).

1.7 Current therapeutic approaches for PD

There is no cure for PD and current therapies are aimed primarily to offer relief from the motor symptoms. The therapeutics includes medications such as Levodopa (L-DOPA), Dopamine agonists, Monoamine-oxidase-B (MAO-B) inhibitors, Catechol-O-methyl transferase (COMT) inhibitors and deep brain stimulation (DBS). Recent advances have shown stem cell therapy and gene silencing to be experimental therapeutic approaches.

1.7.1 Levodopa (L-DOPA)

In 1967 it was discovered that L-DOPA can promptly and dramatically reduce PD motor symptoms and it is currently still the most effective treatment (Cotzias *et al.* 1967). L-DOPA is converted into dopamine in the dopaminergic neurons by the dopa-decarboxylase enzyme causing an increase in the levels of dopamine and improvement in PD motor symptoms. Increased dosing and prolonged use of this drug are reliably predicted to result in motor fluctuations, and abnormal movements referred to as dopamine-induced dyskinesias.

1.7.2 Dopamine agonists

Dopamine agonists stimulate the post-synaptic dopamine receptors and are useful agents, particularly since they have a longer half-life than dopamine, and a lower potential for inducing dyskinesias. Currently, dopamine agonists are used as initial therapy in order to reduce motor complications, particularly in young onset PD. Commonly used dopamine antagonists include, Ropinirole, Pramipexole, Rotigotine, and Bromocriptine (Davie 2008). Side effects include hallucinations, impulse control disorders, insomnia, nausea, low blood pressure, and sleepiness.

1.7.3 MAO-B inhibitors

The function of monoamine-oxidase-B (MAO-B) is to break down dopamine that is secreted by dopaminergic neurons. MAO-B inhibitors act to reduce the breakdown of dopamine in order to prolong its half-life in the brain (Espay 2010). The dosage of L-DOPA is lower in combination with the MAO-B inhibitors. Examples of these drugs are Selegiline and Rasagiline, and some of their side effects include insomnia, hallucinations, dizziness, and headaches.

1.7.4 COMT inhibitors

Catechol-O-methyl transferase (COMT) is an enzyme that metabolizes dopamine. COMT inhibitors reduce the action of this enzyme, resulting in greater and more sustained plasma and central nervous system dopamine levels. If motor fluctuations are present, COMT inhibitors have been shown to result in a decrease in 'off time'. COMT inhibitors increase the half-life of L-DOPA by ~45% after each dose (Davie 2008). Examples include Entacapone and Tolcapone. Some of the side effects of this medication are constipation, nausea, diarrhea, blood in urine and liver failure.

1.7.5 Deep brain stimulation (DBS)

DBS is a surgical procedure and is offered to patients in a number of settings, in particular, marked dyskinesias or tremor, and where medication has not materially affected motor fluctuations (Davie 2008). The DBS system consists of a lead, extension and the neurostimulator. The lead is inserted through a small opening in the skull and is implanted in the brain with the tip of the electrodes positioned in the target area where it delivers the necessary stimulation. The extension (an insulated wire) is passed under the skin of the head, neck and shoulder to connect the lead to the neurostimulator, a battery-operated medical device which is implanted under the skin near the collarbone. Electrical impulses are sent from this device to the electrodes implanted in the target brain area in order to block the abnormal electrical nerve signals that cause PD symptoms. Breakage of wires, infection and hematoma (a collection of blood outside a blood vessel) are complications that may arise from DBS. Side effects include impairment of verbal fluency and depression.

1.7.6 Experimental therapeutic approaches

Stem cell therapy

The above-mentioned therapeutic strategies provide symptomatic relief and do not change the course of PD. Therefore, methods that can change the course of the disease by restoring or regenerating dopaminergic neurons are necessary. Stem cells have been the focus for such therapy due to the fact that they are pluripotent cells that can spontaneously differentiate, or can be induced to differentiate, into mature dopaminergic cells. Embryonic stem cells (derived from the inner cell mass of the blastula of the embryo), neural stem cells (endogenous pluripotent cells in fetal or adult brain) and non-neural stem cells (derived from bone marrow and umbilical cord) are types of cells that have been investigated for stem cell therapy (Snyder & Olanow 2005). In order to produce beneficial effects in treating PD, a sufficient number of stem cell-derived dopaminergic cells (>100 000 cells) that are implanted into the brain must survive, reinnervate the host striatum, integrate into the host nigrostriatal system, establish functional contacts, and should not lead to tumor formation. Also, stem cells must first be demonstrated to reverse motor deficits in animal PD models (Arenas 2010). There have unfortunately been difficulties that have delayed the routine use of stem cell therapy in treating PD. Furthermore, an objective of stem cell therapy is to target the cells to produce dopamine. Non-motor symptoms are generally not related to dopamine deficiency and therefore will not necessarily be improved by stem cell therapy.

Gene silencing

The use of RNA interference (RNAi) to suppress α-synuclein expression has recently been identified as a new therapeutic approach to treat PD. It has been shown that α-synuclein expression can be reduced by means of RNAi in neuronal cells susceptible to neurodegeneration in primates (McCormack *et al.* 2010). In this study, small interfering RNA directed against α-synuclein were directly infused into the SNpc of primates and resulted in a 40-50% reduction in α-synuclein expression. No adverse side effects were observed such as systemic or tissue-specific toxicity, loss of dopaminergic cells in the SNpc or behavioral changes. However, further studies are necessary in order to safely apply this therapy in humans.

Current treatment strategies focus mainly on relief from symptoms and none halt or lessen dopaminergic neuron degeneration and the progression of the disease. It is hoped that research into the molecular pathways involved in PD pathogenesis will lead to the development of improved and effective drugs that can target the underlying cause rather than only treating the

symptoms. These new treatment strategies may be pathway or mutation-specific if no unifying pathway for PD is found. As approximately, seventy percent of dopaminergic neurons are lost by the time motor symptoms start to manifest, one of the main aims of molecular research is to identify presymptomatic individuals who will be candidates for neuroprotective strategies to prevent further neurodegeneration, thereby preventing development of PD.

1.8 PD molecular research in South Africa

By September 2010, DNA from 250 families diagnosed with PD has been extracted and archived in our laboratory at the Division of Molecular Biology and Human Genetics at Stellenbosch University in Cape Town, South Africa. These patients have been recruited from the Movement Disorders clinic at Tygerberg Hospital, Cape Town, as well as from the Parkinson's Association of South Africa with informed written consent according to the guidelines of the institutional review board (ethics approval number 2002/C059). The patients recruited at Tygerberg Hospital all met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD (Gibb & Lees 1988). Other inclusion criteria for recruitment included early onset and/or a family history of the disorder.

A breakdown of the mode of inheritance of PD is indicated in Table 1.2. The term 'unknown/ indeterminate' is used to describe either a single PD patient with no family history of the disorder or cases where, although there was more than one affected PD member, the mode of transmission could not be deduced from the pedigree.

Table 1.2 Genetic categorization of 250 South African PD patients from which DNA has been archived

Mode of inheritance	Number of probands/ families (n=250)	Number of individuals (n=645)	Average AAO of probands (yrs)	Number of probands with AAO ≤50
Autosomal dominant	72 (28.8%)	224	53.6	32
Autosomal recessive	14 (5.6%)	55	50.1	8
Unknown/ indeterminate	164 (65.6%)	366	54.8	54

AAO: age at onset

In South Africa the population can be divided into various ethnic groups including Black, Caucasian, Afrikaner, mixed ancestry and Indian.

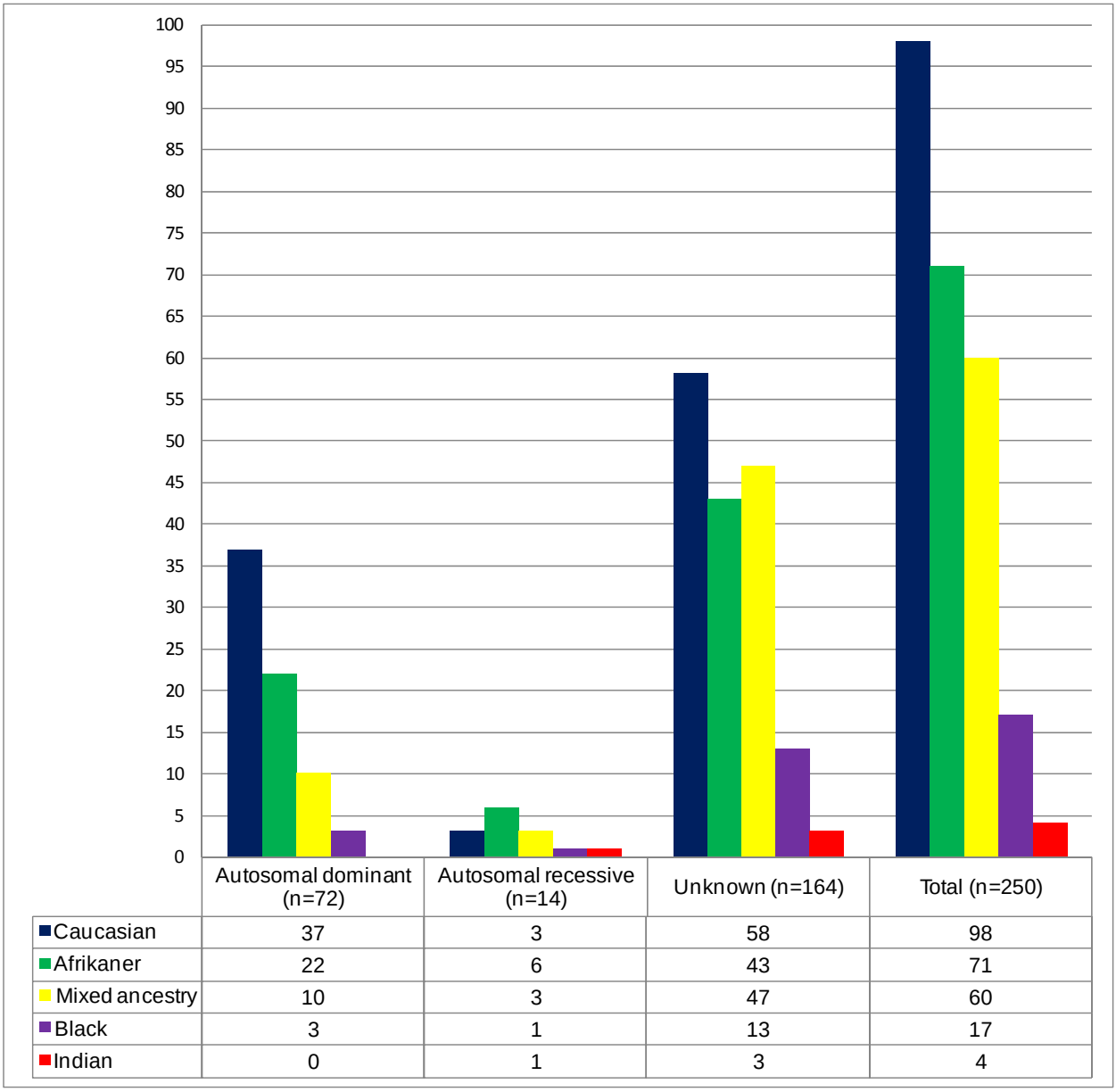
- The Black population comprises individuals with indigenous African ancestry who speak traditional African languages such as Ndebele, Northern Sotho, Sotho, Swazi, Tswana, Tsonga, Venda, Xhosa and Zulu. Xhosa is the predominant language in the Western Cape Province, where the study participants had been recruited.
- The Caucasian population consists of individuals of European descent.
- The Afrikaner population is known to have undergone a genetic bottleneck in the 19th century. They are comprised of Afrikaans speaking individuals mainly of Dutch descent but also of French and German ancestry (Greeff 2007).
- The mixed ancestry population has been defined as an admixture of indigenous African populations (San, Khoikhoi or Bantu-speaking) and immigrants from Western Europe, Madagascar, the Malaysian archipelago and India (Patterson *et al.* 2010).
- The Indian population group is descended from individuals that migrated to South Africa from colonial India in the 19th century.

The numbers of PD patients from each of the ethnic groups are indicated in Figure 1.15. About a third (28.4%) of the families belongs to the Afrikaner population group, which is most likely due to ascertainment bias, although founder effect in this population cannot be excluded.

In addition, at least 50 individuals from each of these ethnic groups were recruited as controls from unrelated healthy blood donors at the South African Western Province Blood Transfusion Service blood collection clinics. Both Caucasian and Afrikaner controls were recruited and the ethnic group was self-identified by the donors. The controls had been ‘de-identified’ and had not been clinically assessed for signs of PD.

1.8.1 The present study

The aim of the present study was to establish the first investigation on the genetic etiology of PD in South Africa. The genetic basis of PD in patients from the various South African sub-population groups was investigated and the study aimed to produce results which would, in the long term, be beneficial for the clinical management and improved genetic counseling of PD patients and at-risk family members. Predominantly early onset PD was investigated as this has been shown to have a significant genetic component. It is anticipated, however, that the findings may contribute to an understanding of the molecular mechanisms underlying the more common late onset idiopathic form of PD.



	Autosomal dominant (n=72)	Autosomal recessive (n=14)	Unknown (n=164)	Total (n=250)
Caucasian	37	3	58	98
Afrikaner	22	6	43	71
Mixed ancestry	10	3	47	60
Black	3	1	13	17
Indian	0	1	3	4

Figure 1.15 Distribution of AD and AR forms of PD across the different South African ethnic groups.

The objectives of the study were as follows:

1. To screen for mutations in the known PD genes using the high-resolution melt (HRM) and the multiplex ligation-dependent probe amplification (MLPA) techniques.
2. To identify novel PD candidate genes for mutation screening using the bioinformatics tool, CAESAR (CAndidatE Search And Rank).
3. To determine the functional significance of a sequence variant detected in the *DJ-1* gene using the Dual-Luciferase Reporter assay.

Figure 1.16 shows an outline of the present study. In the different parts of this study, varying numbers of study participants were screened depending on the patient availability at the time of the study and on the specific clinical criteria applied.

1.8.2 The experimental approaches used

High-resolution melt (HRM)

The HRM method makes it possible to assess the nucleotide sequence of a Polymerase chain reaction (PCR) amplification product according to its melting behavior as the sample changes from double stranded DNA (dsDNA) (pre-melt phase) to single stranded DNA (ssDNA) (post-melt phase) during the rise in temperature as illustrated in figure 1.17. Initially, the DNA region of interest is PCR amplified during which an intercalating fluorescent dye (for example SYTO9) is incorporated into the dsDNA molecule. dsDNA produce higher fluorescence than ssDNA because the dye is more concentrated. The rise in temperature (typically 75ºC to 95ºC rising by 0.1ºC each step) causes the molecule to become ssDNA resulting in the release of the fluorescent dye and subsequent decrease in its concentration. By measuring this melt behavior of the PCR product, it is possible to draw a thermal denaturation profile in which fluorescence (Y-axis) is plotted against temperature (X-axis). This profile is characteristic of the specific PCR product and is dependent on its sequence length, base and GC content. Therefore, changes in the nucleotide sequence will affect the melting behavior of a DNA molecule and will produce a different profile compared to the wild-type sample. In the present study, the Eppendorf® ep*Motion*™ 5070 (Brinkmann Instruments, Canada) was used for automated liquid handling, and HRM analysis was performed on a Rotor-Gene 6000 analyzer (Corbett Life Science, Australia). All samples exhibiting altered thermal denaturation profiles were sequenced to characterize the sequence variant.

Outline of the present study

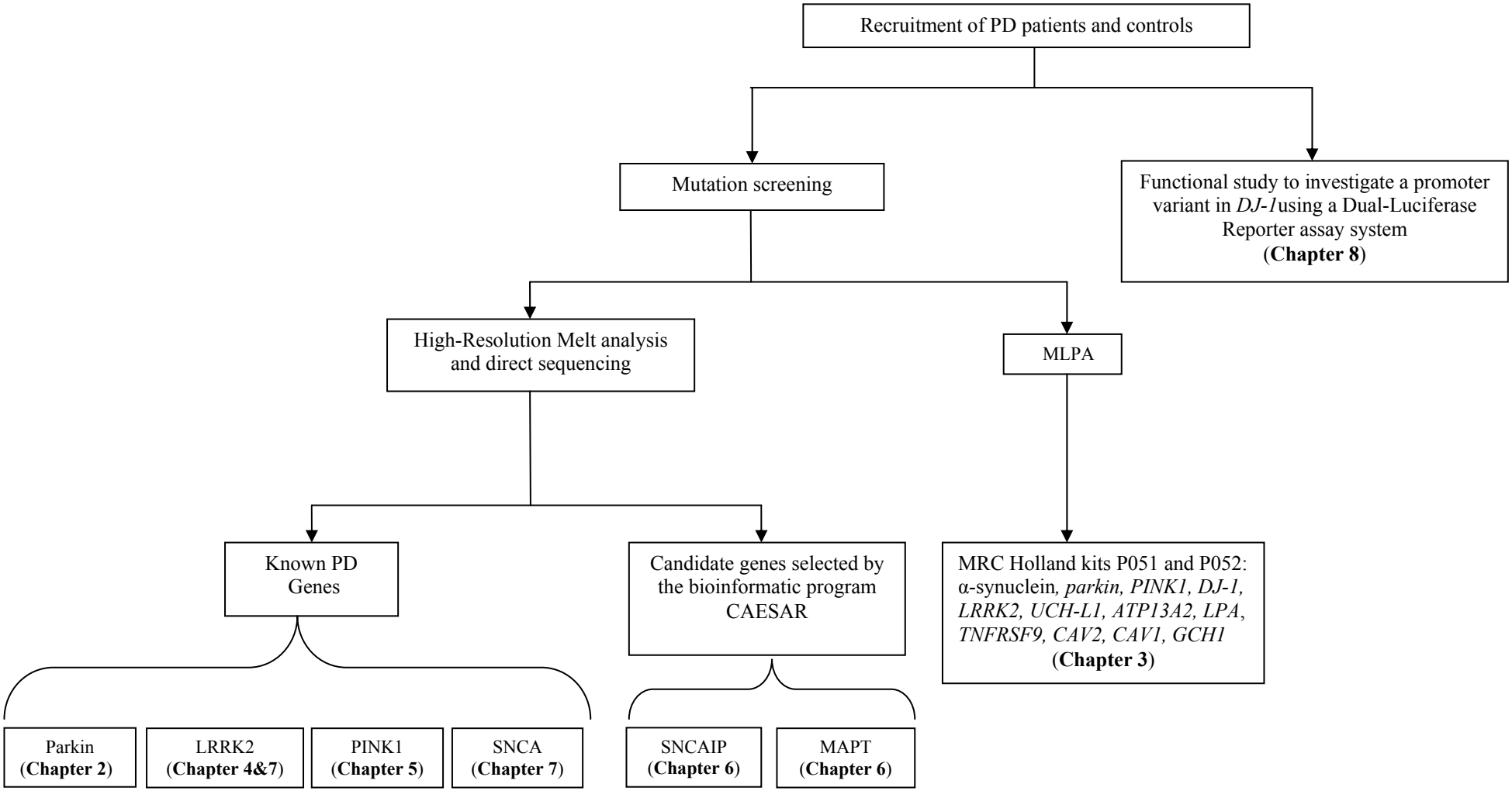


Figure 1.16 Flowchart showing the outline of the present study

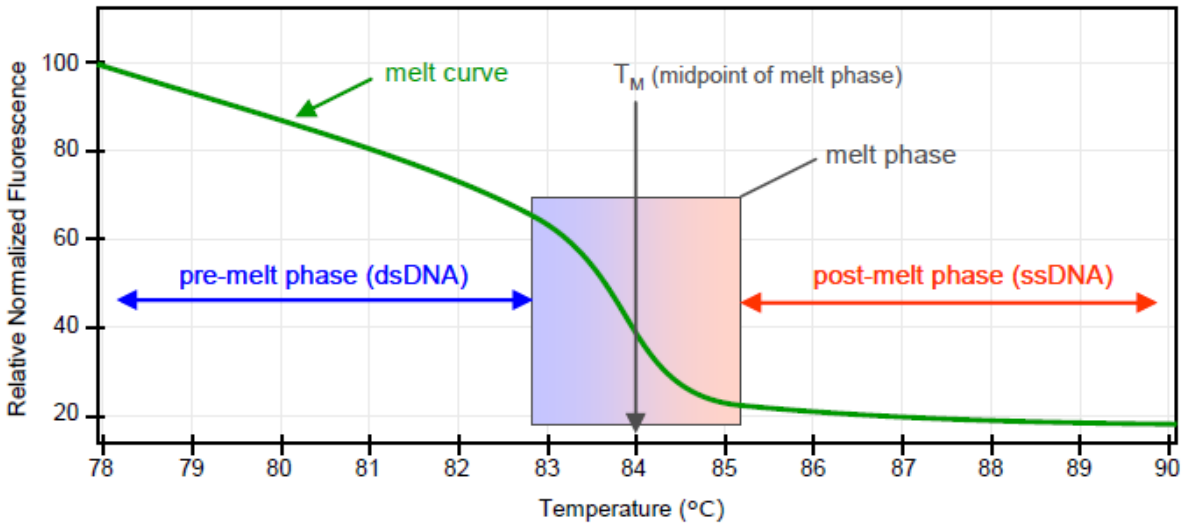


Figure 1.17 Illustration of the High-resolution melt technique. Double stranded DNA (dsDNA) melt to become single stranded DNA (ssDNA). During this process the fluorescent intercalating dye is released, and the fluorescence is measured to produce a thermal denaturation profile that is unique for a specific DNA sequence. Fluorescence (Y-axis) is plotted against temperature in degrees Celsius (X-axis) (taken from HRM Assay Design and Analysis Booklet, http://www.gene-quantification.de/hrm-protocol-cls.pdf)

Multiplex ligation-dependent probe amplification method (MLPA)

This method makes it possible to assess genomic DNA for changes such as exon insertions and deletions which cannot be detected by HRM or direct sequencing. It detects alterations in a number of different exons and genes in a single reaction. In the present study two commercially available MLPA kits, SALSA P051 and P052 (MRC Holland, Amsterdam), were used and together the kits consisted of probes for exons of α-synuclein, *parkin, PINK1, DJ-1, LRRK2, UCH-L1, ATP13A2, LPA*, *TNFRSF9, CAV2, CAV1, GCH1* and two point mutations (A30P in *SNCA* and G2019S in *LRRK2*). The MLPA probes are made up of two separate oligonucleotides containing a hybridization sequence and universal fluorescently-labeled primer sequences as shown in figure 1.18. In addition, one of the oligonucleotides also contain a stuffer fragment in order to produce PCR products of specific size which will make it possible to differentiate amplified exons after electrophoresis.

Initially, denatured genomic DNA is hybridized with the MLPA probes (either P051 or P052). The probes are then ligated to form one complete probe fragment which is amplified in a multiplex PCR reaction using one universal primer pair. Amplification products of specific lengths are produced and electrophoresed using capillary electrophoresis which produces a peak profile unique for the DNA sample. For the present study MLPA products were electrophoresed on 3130xl Genetic Analyzer (Applied Biosystems, Foster City, USA) and analyzed using

GeneMapper software v3.7 (Applied Biosystems). The amount of fluorescence detected reflects the relative copy number of the exons being investigated. An in-house method was designed for copy number calculations. The peak height results for patients and reference wild-type samples were exported from GeneMapper software to a Microsoft office Excel spreadsheet which had been set up for automated calculations.

The following steps were followed for the copy number calculations. The relative peak height (RPH) of each exon was obtained by dividing the single peak height by the sum of peak heights of all the internal control probes for the patient sample (intra-sample normalization). The same was done for the reference wild-type sample. A ratio was then generated by comparing each patient sample RPH to the value of the corresponding RPHs from the reference wild-type sample (inter-sample normalization). Ratios between 0.7-1.3 were considered to be normal (sample contains no exon rearrangements); a heterozygous deletion was expected at a ratio between 0.3-0.6, a heterozygous duplication between 1.4-1.6, and a triplication at a ratio of ≥1.7. Absence of a peak would indicate either a homozygous deletion or possibly a point mutation at that probe recognition site (ratio 0.0). Possible mutation positive results were repeated in an independent MLPA experiment and then subjected to quantitative PCR on a LightCycler 1.5 (Roche Diagnostics, Mannheim, Germany) using the LightCycler® FastStart DNA Master^{PLUS} SYBR Green I kit (Roche Diagnostics). The samples were also sequenced to determine whether false positive results were due to point mutations occurring at the probe recognition site.

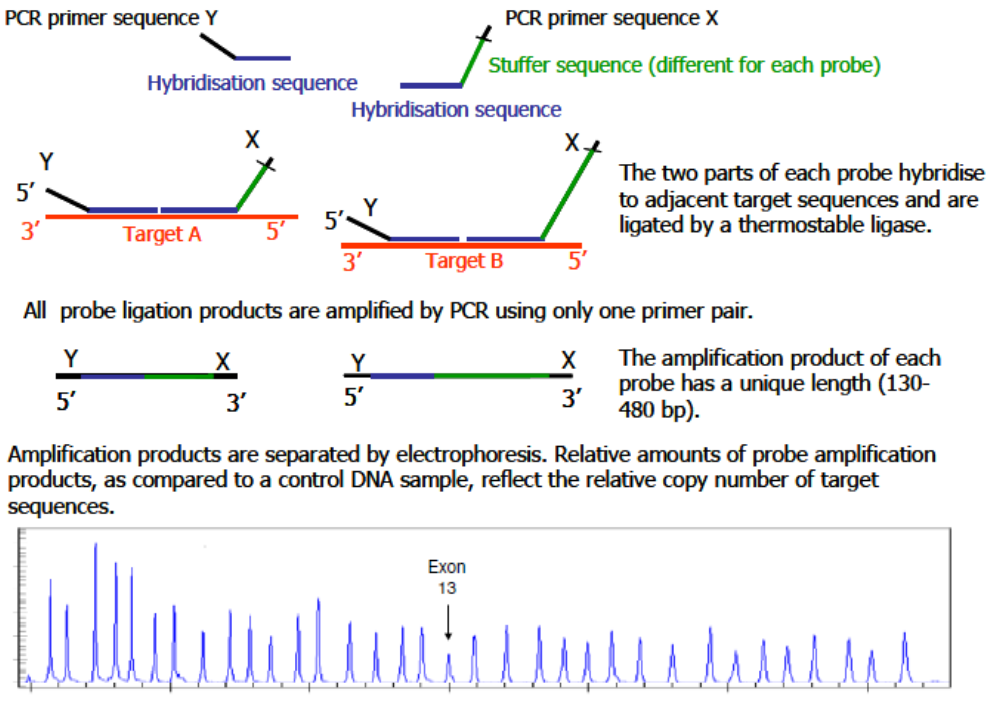
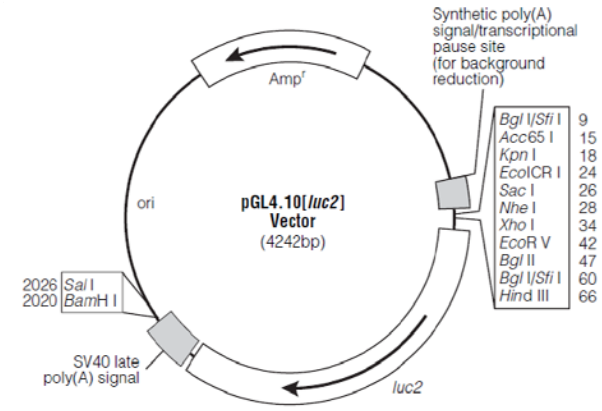


Figure 1.18 Illustration of the MLPA method (taken from http://www.mlpa.com).

Dual-Luciferase Reporter Assay

The functional significance of a novel 16bp deletion variant (g.-6_+10del), which was identified in the promoter region of *DJ-1* in a PD patient, was investigated using a Dual-Luciferase Reporter Assay System (http://www.promega.com/pGL4). For this assay, the *DJ-1* 5'-UTR region containing the sequence flanking the 16bp deletion was cloned into a pGL4.10-Basic luciferase-reporter vector (figure 1.19) and co-transfected with a pRL-SV40 internal control vector into HEK293 as well as BE(2)-M17 neuroblastoma cells. Promoter activity under hydrogen peroxide-induced oxidative stress conditions was also investigated. In addition, computational (*in silico*) *cis*-regulatory analysis of the *DJ-1* promoter sequence was conducted to identify possible transcription factor binding sites. This analysis was performed using the transcription factor-binding site database, TRANSFAC via the PATCH™ (http://www.gene-regulation.com) and rVISTA (http://genome.lbl.gov/vista/index.shtml) platforms.

a)



b)

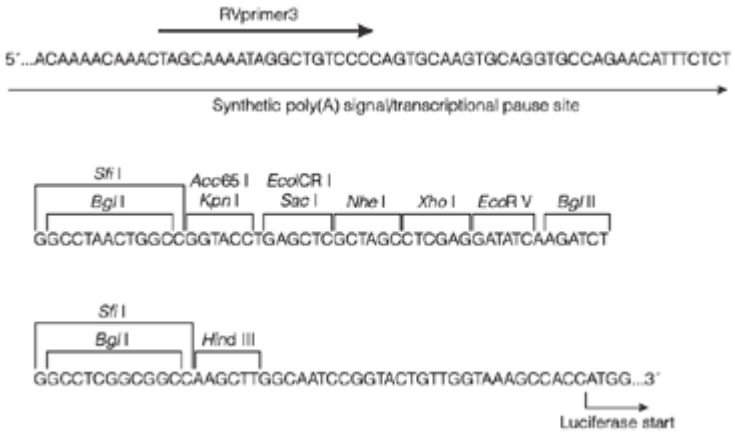


Figure 1.19 Restriction map and multiple cloning site (MCS) of the pGL4.10[*luc2*] vector.
a) The positions of the ampicillin resistance gene (*Amp*^{r}), plasmid origin of replication, luc2, SV40 late poly(A) signal, synthetic poly(A) signal, and MCS are indicated on the map. b) Nucleotide sequence of the pGL4.10[*luc2*] MCS. The positions of all unique restriction enzymes and recognition sequences, the position of the RVprimer3 primer and the synthetic poly(A) signal are indicated on the sequence (Taken from http://www.promega.com/vectors/reporter_vectors.htm#b07).

1.8.3 Outline of the dissertation

This dissertation is divided into nine chapters. Chapter 1 is the literature review and provides the rationale for the study. Chapter 2 describes analysis of the *parkin* gene. Chapter 3 reports the findings for exon dosage analysis in a number of PD genes using the MLPA technique. In Chapter 4 the frequency of the *LRRK2* common G2019S mutation, and the associated haplotype, is investigated. Chapter 5 describes the analysis of *PINK1*. Chapter 6 describes the use of the bioinformatics tool CAESAR (Gaulton *et al.* 2007) to identify novel PD candidate genes for mutation screening in the study participants. Chapter 7 reports unpublished data on mutation screening of selected exons of the *SNCA* and *LRRK2* genes. In Chapter 8 the results of a study to determine the functional significance of a novel 16bp deletion variant in the promoter region of *DJ-1* are reported. Finally, Chapter 9 summarizes the main findings of the present study, and states the study limitations and the proposed future work.

The findings from the present study have either been published (five manuscripts) or submitted for publication (one manuscript) in international peer reviewed journals, as listed below:

Chapter 2 - Bardien S, **Keyser RJ**, Yako Y, Lombard D, Carr J. Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease. *Parkinsonism and Related Disorders* 2009;15(2):116-121.

Chapter 3 - **Keyser RJ**, Lombard D, Veikondis R, Carr J, Bardien S. Analysis of exon dosage using MLPA in South African Parkinson's disease patients. *Neurogenetics* 2010; 11(3):305-312.

Chapter 4 - Bardien S, Marsberg A, **Keyser R**, Lombard D, Lesage S, Brice A, Carr J. LRRK2 G2019S mutation: frequency and haplotype data in South African Parkinson's disease patients. *J Neural Transm*. 2010; 117(7):847-853.

Chapter 5 - **Keyser RJ**, Lesage S, Brice A, Carr J, Bardien S. Assessing the prevalence of PINK1 genetic variants in South African patients diagnosed with early- and late-onset Parkinson's disease. *Biochem Biophys Res Commun*. 2010; 398(1):125-129.

Chapter 6 - **Keyser RJ**, Oppon E, Carr J, Bardien S. Identification of Parkinson's disease candidate genes using CAESAR and screening in South African Parkinson's disease patients (*submitted to Journal of Neural Transmission, September 2010*)

Chapter 7 – Additional unpublished results

Chapter 8 - **Keyser RJ**, van der Merwe L, Venter M, Kinnear C, Lombard D, Warnich L, Carr J, Bardien S. Identification of a novel functional deletion variant in the 5'-UTR of the DJ-1 gene. *BMC Medical Genetics* 2009; 10:105.

1.9 Reference list

Abeliovich, A., Y. Schmitz, I. Farinas, D. Choi-Lundberg, W. H. Ho *et al.* 2000 Mice lacking alpha-synuclein display functional deficits in the nigrostriatal dopamine system. *Neuron* **25**: 239-252.

Abou-Sleiman, P. M., M. M. Muqit, N. Q. McDonald, Y. X. Yang, S. Gandhi *et al.* 2006 A heterozygous effect for PINK1 mutations in Parkinson's disease? *Ann.Neurol.* **60**: 414-419.

Aharon-Peretz, J., H. Rosenbaum, and R. Gershoni-Baruch, 2004 Mutations in the glucocerebrosidase gene and Parkinson's disease in Ashkenazi Jews. *N.Engl.J.Med.* **351**: 1972-1977.

Alves, G., E. B. Forsaa, K. F. Pedersen, G. M. Dreetz, and J. P. Larsen, 2008 Epidemiology of Parkinson's disease. *J.Neurol.* **255 Suppl 5**: 18-32.

Arenas, E., 2010 Towards stem cell replacement therapies for Parkinson's disease. *Biochem.Biophys.Res.Commun.* **396**: 152-156.

Ashok, P. P., K. Radhakrishnan, R. Sridharan, and M. E. Mousa, 1986 Epidemiology of Parkinson's disease in Benghazi, North-East Libya. *Clin.Neurol.Neurosurg.* **88**: 109-113.

Beilina, A., B. M. van der, R. Ahmad, S. Kesavapany, D. W. Miller *et al.* 2005 Mutations in PTEN-induced putative kinase 1 associated with recessive parkinsonism have differential effects on protein stability. *Proc.Natl.Acad.Sci.U.S.A* **102**: 5703-5708.

Berg, D., K. Schweitzer, P. Leitner, A. Zimprich, P. Lichtner *et al.* 2005 Type and frequency of mutations in the LRRK2 gene in familial and sporadic Parkinson's disease*. *Brain* **128**: 3000-3011.

Betarbet, R., R. M. Canet-Aviles, T. B. Sherer, P. G. Mastroberardino, C. McLendon *et al.* 2006 Intersecting pathways to neurodegeneration in Parkinson's disease: effects of the pesticide rotenone on DJ-1, alpha-synuclein, and the ubiquitin-proteasome system. *Neurobiol.Dis.* **22**: 404-420.

Betarbet, R., T. B. Sherer, G. MacKenzie, M. Garcia-Osuna, A. V. Panov *et al.* 2000 Chronic systemic pesticide exposure reproduces features of Parkinson's disease. *Nat.Neurosci.* **3**: 1301-1306.

Bonifati, V., B. A. Oostra, and P. Heutink, 2004 Linking DJ-1 to neurodegeneration offers novel insights for understanding the pathogenesis of Parkinson's disease. *J.Mol.Med.* **82**: 163-174.

Bonifati, V., P. Rizzu, M. J. van Baren, O. Schaap, G. J. Breedveld *et al.* 2003 Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science* **299**: 256-259.

Bonifati, V., C. F. Rohe, G. J. Breedveld, E. Fabrizio, M. M. De *et al.* 2005 Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes. *Neurology* **65**: 87-95.

Canet-Aviles, R. M., M. A. Wilson, D. W. Miller, R. Ahmad, C. McLendon *et al.* 2004 The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfinic acid-driven mitochondrial localization. *Proc.Natl.Acad.Sci.U.S.A* **101**: 9103-9108.

Chartier-Harlin, M. C., J. Kachergus, C. Roumier, V. Mouroux, X. Douay *et al.* 2004 Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet* **364**: 1167-1169.

Chung, K. K., Y. Zhang, K. L. Lim, Y. Tanaka, H. Huang *et al.* 2001 Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease. *Nat.Med.* **7**: 1144-1150.

Ciechanover, A., and P. Brundin, 2003 The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. *Neuron* **40**: 427-446.

Conway, K. A., J. D. Harper, and P. T. Lansbury, 1998 Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Nat.Med.* **4**: 1318-1320.

Cookson, M. R., and O. Bandmann, 2010 Parkinson's disease: insights from pathways. *Hum.Mol.Genet.* **19**: R21-R27.

Cookson, M. R., P. J. Lockhart, C. McLendon, C. O'Farrell, M. Schlossmacher *et al.* 2003 RING finger 1 mutations in Parkin produce altered localization of the protein. *Hum.Mol.Genet.* **12**: 2957-2965.

Cotzias, G. C., M. H. Van Woert, and L. M. Schiffer, 1967 Aromatic amino acids and modification of parkinsonism. *N.Engl.J.Med.* **276**: 374-379.

Coulom, H., and S. Birman, 2004 Chronic exposure to rotenone models sporadic Parkinson's disease in Drosophila melanogaster. *J.Neurosci.* **24**: 10993-10998.

Darios, F., O. Corti, C. B. Lucking, C. Hampe, M. P. Muriel *et al.* 2003 Parkin prevents mitochondrial swelling and cytochrome c release in mitochondria-dependent cell death. *Hum.Mol.Genet.* **12**: 517-526.

Dauer, W., and S. Przedborski, 2003 Parkinson's disease: mechanisms and models. *Neuron* **39**: 889-909.

Davie, C. A., 2008 A review of Parkinson's disease. *Br.Med.Bull.* **86**: 109-127.

de Rijk, M. C., L. J. Launer, K. Berger, M. M. Breteler, J. F. Dartigues *et al.* 2000 Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology* **54**: S21-S23.

Di Fonzo, A., H. F. Chien, M. Socal, S. Giraudo, C. Tassorelli *et al.* 2007 ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease. *Neurology* **68**: 1557-1562.

Di Fonzo, A., C. Tassorelli, M. M. De, H. F. Chien, J. Ferreira *et al.* 2006a Comprehensive analysis of the LRRK2 gene in sixty families with Parkinson's disease. *Eur.J.Hum.Genet.* **14**: 322-331.

Di Fonzo, A., Y. H. Wu-Chou, C. S. Lu, D. M. van, E. J. Simons *et al.* 2006b A common missense variant in the LRRK2 gene, Gly2385Arg, associated with Parkinson's disease risk in Taiwan. *Neurogenetics.* **7**: 133-138.

Dluzen, D. E., and J. L. McDermott, 2000 Gender differences in neurotoxicity of the nigrostriatal dopaminergic system: implications for Parkinson's disease. *J.Gend.Specif.Med.* **3**: 36-42.

Edwards, T. L., W. K. Scott, C. Almonte, A. Burt, E. H. Powell *et al.* 2010 Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann.Hum.Genet.* **74**: 97-109.

Elbaz, A., and C. Tranchant, 2007 Epidemiologic studies of environmental exposures in Parkinson's disease. *J.Neurol.Sci.* **262**: 37-44.

Engelender, S., Z. Kaminsky, X. Guo, A. H. Sharp, R. K. Amaravi *et al.* 1999 Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. *Nat.Genet.* **22**: 110-114.

Engelender, S., T. Wanner, J. J. Kleiderlein, K. Wakabayashi, S. Tsuji *et al.* 2000 Organization of the human synphilin-1 gene, a candidate for Parkinson's disease. *Mamm.Genome* **11**: 763-766.

Ephraty, L., O. Porat, D. Israeli, O. S. Cohen, O. Tunkel *et al.* 2007 Neuropsychiatric and cognitive features in autosomal-recessive early parkinsonism due to PINK1 mutations. *Mov Disord.* **22**: 566-569.

Espay, A. J., 2010 Management of motor complications in Parkinson disease: current and emerging therapies. *Neurol.Clin.* **28**: 913-925.

Fahn, S., 2003 Description of Parkinson's disease as a clinical syndrome. *Ann.N.Y.Acad.Sci.* **991**: 1-14.

Farrer, M., L. Skipper, M. Berg, G. Bisceglio, M. Hanson *et al.* 2002 The tau H1 haplotype is associated with Parkinson's disease in the Norwegian population. *Neurosci.Lett.* **322**: 83-86.

Feany, M. B., and W. W. Bender, 2000 A Drosophila model of Parkinson's disease. *Nature* **404**: 394-398.

Fornai, F., O. M. Schluter, P. Lenzi, M. Gesi, R. Ruffoli *et al.* 2005 Parkinson-like syndrome induced by continuous MPTP infusion: convergent roles of the ubiquitin-proteasome system and alpha-synuclein. *Proc.Natl.Acad.Sci.U.S.A* **102**: 3413-3418.

Foroud, T., S. K. Uniacke, L. Liu, N. Pankratz, A. Rudolph *et al.* 2003 Heterozygosity for a mutation in the parkin gene leads to later onset Parkinson disease. *Neurology* **60**: 796-801.

Gandhi, S., M. M. Muqit, L. Stanyer, D. G. Healy, P. M. bou-Sleiman *et al.* 2006 PINK1 protein in normal human brain and Parkinson's disease. *Brain* **129**: 1720-1731.

Gaulton, K. J., K. L. Mohlke, and T. J. Vision, 2007 A computational system to select candidate genes for complex human traits. *Bioinformatics.* **23**: 1132-1140.

Giasson, B. I., J. E. Duda, S. M. Quinn, B. Zhang, J. Q. Trojanowski *et al.* 2002 Neuronal alpha-synucleinopathy with severe movement disorder in mice expressing A53T human alpha-synuclein. *Neuron* **34**: 521-533.

Giasson, B. I., M. S. Forman, M. Higuchi, L. I. Golbe, C. L. Graves *et al.* 2003 Initiation and synergistic fibrillization of tau and alpha-synuclein. *Science* **300**: 636-640.

Gibb, W. R., and A. J. Lees, 1988 A comparison of clinical and pathological features of young- and old-onset Parkinson's disease. *Neurology* **38**: 1402-1406.

Gibb, W. R., and A. J. Lees, 1991 Anatomy, pigmentation, ventral and dorsal subpopulations of the substantia nigra, and differential cell death in Parkinson's disease. *J.Neurol.Neurosurg.Psychiatry* **54**: 388-396.

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. *Horm.Behav.* **57**: 23-34.

Gitler, A. D., A. Chesi, M. L. Geddie, K. E. Strathearn, S. Hamamichi *et al.* 2009 Alpha-synuclein is part of a diverse and highly conserved interaction network that includes PARK9 and manganese toxicity. *Nat.Genet.* **41**: 308-315.

Goris, A., C. H. Williams-Gray, G. R. Clark, T. Foltynie, S. J. Lewis *et al.* 2007 Tau and alpha-synuclein in susceptibility to, and dementia in, Parkinson's disease. *Ann.Neurol.* **62**: 145-153.

Gorman, A. M., S. Ceccatelli, and S. Orrenius, 2000 Role of mitochondria in neuronal apoptosis. *Dev.Neurosci.* **22**: 348-358.

Greeff, J. M., 2007 Deconstructing Jaco: genetic heritage of an Afrikaner. *Ann.Hum.Genet.* **71**: 674-688.

Gu, W. J., O. Corti, F. Araujo, C. Hampe, S. Jacquier *et al.* 2003 The C289G and C418R missense mutations cause rapid sequestration of human Parkin into insoluble aggregates. *Neurobiol.Dis.* **14**: 357-364.

Guo, L., W. Wang, and S. G. Chen, 2006 Leucine-rich repeat kinase 2: relevance to Parkinson's disease. *Int.J.Biochem.Cell Biol.* **38**: 1469-1475.

Hayashi, S., K. Wakabayashi, A. Ishikawa, H. Nagai, M. Saito *et al.* 2000 An autopsy case of autosomal-recessive juvenile parkinsonism with a homozygous exon 4 deletion in the parkin gene. *Mov Disord.* **15**: 884-888.

Healy, D. G., M. Falchi, S. S. O'Sullivan, V. Bonifati, A. Durr *et al.* 2008 Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol.* **7**: 583-590.

Hedrich, K., C. Eskelson, B. Wilmot, K. Marder, J. Harris *et al.* 2004 Distribution, type, and origin of Parkin mutations: review and case studies. *Mov Disord.* **19**: 1146-1157.

Hedrich, K., S. Winkler, J. Hagenah, K. Kabakci, M. Kasten *et al.* 2006 Recurrent LRRK2 (Park8) mutations in early-onset Parkinson's disease. *Mov Disord.* **21**: 1506-1510.

Hershko, A., and A. Ciechanover, 1998 The ubiquitin system. *Annu.Rev.Biochem.* **67**: 425-479.

Hilker, R., C. Klein, M. Ghaemi, B. Kis, T. Strotmann *et al.* 2001 Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene. *Ann.Neurol.* **49**: 367-376.

Hirokawa, N., 1994 Microtubule organization and dynamics dependent on microtubule-associated proteins. *Curr.Opin.Cell Biol.* **6**: 74-81.

Huai, Q., Y. Sun, H. Wang, L. S. Chin, L. Li *et al.* 2003 Crystal structure of DJ-1/RS and implication on familial Parkinson's disease. *FEBS Lett.* **549**: 171-175.

Hutton, M., C. L. Lendon, P. Rizzu, M. Baker, S. Froelich *et al.* 1998 Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393**: 702-705.

Huynh, D. P., D. R. Scoles, D. Nguyen, and S. M. Pulst, 2003 The autosomal recessive juvenile Parkinson disease gene product, parkin, interacts with and ubiquitinates synaptotagmin XI. *Hum.Mol.Genet.* **12**: 2587-2597.

Ibanez, P., A. M. Bonnet, B. Debarges, E. Lohmann, F. Tison *et al.* 2004 Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**: 1169-1171.

Jo, E., J. McLaurin, C. M. Yip, P. St George-Hyslop, and P. E. Fraser, 2000 alpha-Synuclein membrane interactions and lipid specificity. *J.Biol.Chem.* **275**: 34328-34334.

Johnson, J., C. Paisan-Ruiz, G. Lopez, C. Crews, A. Britton *et al.* 2007 Comprehensive screening of a North American Parkinson's disease cohort for LRRK2 mutation. *Neurodegener.Dis.* **4**: 386-391.

Kahle, P. J., J. Waak, and T. Gasser, 2009 DJ-1 and prevention of oxidative stress in Parkinson's disease and other age-related disorders. *Free Radic.Biol.Med.* **47**: 1354-1361.

Keeney, P. M., J. Xie, R. A. Capaldi, and J. P. Bennett, Jr., 2006 Parkinson's disease brain mitochondrial complex I has oxidatively damaged subunits and is functionally impaired and misassembled. *J.Neurosci.* **26**: 5256-5264.

Khan, N. L., S. Jain, J. M. Lynch, N. Pavese, P. bou-Sleiman *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**: 2786-2796.

Kim, R. H., M. Peters, Y. Jang, W. Shi, M. Pintilie *et al.* 2005 DJ-1, a novel regulator of the tumor suppressor PTEN. *Cancer Cell* **7**: 263-273.

Kitada, T., S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura *et al.* 1998 Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**: 605-608.

Klein, C., A. Djarmati, K. Hedrich, N. Schafer, C. Scaglione *et al.* 2005 PINK1, Parkin, and DJ-1 mutations in Italian patients with early-onset parkinsonism. *Eur.J.Hum.Genet.* **13**: 1086-1093.

Klein, C., and K. Lohmann-Hedrich, 2007 Impact of recent genetic findings in Parkinson's disease. *Curr.Opin.Neurol.* **20**: 453-464.

Kruger, R., W. Kuhn, T. Muller, D. Woitalla, M. Graeber *et al.* 1998 Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nat.Genet.* **18**: 106-108.

Lang, A. E., and A. M. Lozano, 1998 Parkinson's disease. First of two parts. *N.Engl.J.Med.* **339**: 1044-1053.

Langston, J. W., P. Ballard, J. W. Tetrud, and I. Irwin, 1983 Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. *Science* **219**: 979-980.

Lannuzel, A., P. P. Michel, G. U. Hoglinger, P. Champy, A. Jousset *et al.* 2003 The mitochondrial complex I inhibitor annonacin is toxic to mesencephalic dopaminergic neurons by impairment of energy metabolism. *Neuroscience* **121**: 287-296.

Lee, G., R. L. Neve, and K. S. Kosik, 1989 The microtubule binding domain of tau protein. *Neuron* **2**: 1615-1624.

Lesage, S., and A. Brice, 2009 Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Hum.Mol.Genet.* **18**: R48-R59.

Lev, N., D. Ickowicz, E. Melamed, and D. Offen, 2008 Oxidative insults induce DJ-1 upregulation and redistribution: implications for neuroprotection. *Neurotoxicology* **29**: 397-405.

Lewis, S. A., A. Villasante, P. Sherline, and N. J. Cowan, 1986 Brain-specific expression of MAP2 detected using a cloned cDNA probe. *J.Cell Biol.* **102**: 2098-2105.

Li, X., Z. Liu, K. Tamashiro, B. Shi, D. D. Rudnicki *et al.* 2010 Synphilin-1 exhibits trophic and protective effects against Rotenone toxicity. *Neuroscience* **165**: 455-462.

Lim, K. L., and J. M. Tan, 2007 Role of the ubiquitin proteasome system in Parkinson's disease. *BMC.Biochem.* **8 Suppl 1**: S13.

Lin, M. T., and M. F. Beal, 2006 Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* **443**: 787-795.

Lockwood, A. H., 2000 Pesticides and parkinsonism: is there an etiological link? *Curr.Opin.Neurol.* **13**: 687-690.

Lohmann, E., M. Periquet, V. Bonifati, N. W. Wood, M. G. De *et al.* 2003 How much phenotypic variation can be attributed to parkin genotype? *Ann.Neurol.* **54**: 176-185.

Lucking, C. B., and A. Brice, 2000 Alpha-synuclein and Parkinson's disease. *Cell Mol.Life Sci.* **57**: 1894-1908.

Maraganore, D. M., A. M. de, A. Elbaz, M. J. Farrer, J. P. Ioannidis *et al.* 2006 Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. *JAMA* **296**: 661-670.

Marion, S. A., 2001 The epidemiology of Parkinson's disease. Current issues. *Adv.Neurol.* **86**: 163-172.

Marx, F. P., C. Holzmann, K. M. Strauss, L. Li, O. Eberhardt *et al.* 2003 Identification and functional characterization of a novel R621C mutation in the synphilin-1 gene in Parkinson's disease. *Hum.Mol.Genet.* **12**: 1223-1231.

Masliah, E., E. Rockenstein, I. Veinbergs, M. Mallory, M. Hashimoto *et al.* 2000 Dopaminergic loss and inclusion body formation in alpha-synuclein mice: implications for neurodegenerative disorders. *Science* **287**: 1265-1269.

Mata, I. F., J. M. Kachergus, J. P. Taylor, S. Lincoln, J. Aasly *et al.* 2005 Lrrk2 pathogenic substitutions in Parkinson's disease. *Neurogenetics.* **6**: 171-177.

Mata, I. F., P. J. Lockhart, and M. J. Farrer, 2004 Parkin genetics: one model for Parkinson's disease. *Hum.Mol.Genet.* **13 Spec No 1**: R127-R133.

McCormack, A. L., S. K. Mak, J. M. Henderson, D. Bumcrot, M. J. Farrer *et al.* 2010 Alpha-synuclein suppression by targeted small interfering RNA in the primate substantia nigra. *PLoS.One.* **5**: e12122.

Melcon, M. O., D. W. Anderson, R. H. Vergara, and W. A. Rocca, 1997 Prevalence of Parkinson's disease in Junin, Buenos Aires Province, Argentina. *Mov Disord.* **12**: 197-205.

Mochizuki, H., K. Goto, H. Mori, and Y. Mizuno, 1996 Histochemical detection of apoptosis in Parkinson's disease. *J.Neurol.Sci.* **137**: 120-123.

Moore, D. J., A. B. West, V. L. Dawson, and T. M. Dawson, 2005 Molecular pathophysiology of Parkinson's disease. *Annu.Rev.Neurosci.* **28**: 57-87.

Myhre, R., H. Klungland, M. J. Farrer, and J. O. Aasly, 2008 Genetic association study of synphilin-1 in idiopathic Parkinson's disease. *BMC.Med.Genet.* **9**: 19.

Nagakubo, D., T. Taira, H. Kitaura, M. Ikeda, K. Tamai *et al.* 1997 DJ-1, a novel oncogene which transforms mouse NIH3T3 cells in cooperation with ras. *Biochem.Biophys.Res.Commun.* **231**: 509-513.

Naimark, D., E. Jackson, E. Rockwell, and D. V. Jeste, 1996 Psychotic symptoms in Parkinson's disease patients with dementia. *J.Am.Geriatr.Soc.* **44**: 296-299.

Nichols, W. C., V. E. Elsaesser, N. Pankratz, M. W. Pauciulo, D. K. Marek *et al.* 2007 LRRK2 mutation analysis in Parkinson disease families with evidence of linkage to PARK8. *Neurology* **69**: 1737-1744.

Nichols, W. C., N. Pankratz, D. K. Marek, M. W. Pauciulo, V. E. Elsaesser *et al.* 2009 Mutations in GBA are associated with familial Parkinson disease susceptibility and age at onset. *Neurology* **72**: 310-316.

Ning, Y. P., K. Kanai, H. Tomiyama, Y. Li, M. Funayama *et al.* 2008 PARK9-linked parkinsonism in eastern Asia: mutation detection in ATP13A2 and clinical phenotype. *Neurology* **70**: 1491-1493.

Okubadejo, N. U., J. H. Bower, W. A. Rocca, and D. M. Maraganore, 2006 Parkinson's disease in Africa: A systematic review of epidemiologic and genetic studies. *Mov Disord.* **21**: 2150-2156.

Onyango, I. G., 2008 Mitochondrial dysfunction and oxidative stress in Parkinson's disease. *Neurochem.Res.* **33**: 589-597.

Paisan-Ruiz, C., S. Jain, E. W. Evans, W. P. Gilks, J. Simon *et al.* 2004 Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron* **44**: 595-600.

Paisan-Ruiz, C., P. Nath, N. Washecka, J. R. Gibbs, and A. B. Singleton, 2008 Comprehensive analysis of LRRK2 in publicly available Parkinson's disease cases and neurologically normal controls. *Hum.Mutat.* **29**: 485-490.

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**: 1969-1978.

Pankratz, N., and T. Foroud, 2007 Genetics of Parkinson disease. *Genet.Med.* **9**: 801-811.

Park, J., S. B. Lee, S. Lee, Y. Kim, S. Song *et al.* 2006 Mitochondrial dysfunction in Drosophila PINK1 mutants is complemented by parkin. *Nature* **441**: 1157-1161.

Partinen, M., 1997 Sleep disorder related to Parkinson's disease. *J.Neurol.* **244**: S3-S6.

Patterson, N., D. C. Petersen, R. E. van der Ross, H. Sudoyo, R. H. Glashoff *et al.* 2010 Genetic structure of a unique admixed population: implications for medical research. *Hum.Mol.Genet.* **19**: 411-419.

Periquet, M., M. Latouche, E. Lohmann, N. Rawal, M. G. De *et al.* 2003 Parkin mutations are frequent in patients with isolated early-onset parkinsonism. *Brain* **126**: 1271-1278.

Petit, A., T. Kawarai, E. Paitel, N. Sanjo, M. Maj *et al.* 2005 Wild-type PINK1 prevents basal and induced neuronal apoptosis, a protective effect abrogated by Parkinson disease-related mutations. *J.Biol.Chem.* **280**: 34025-34032.

Pirkevi, C., S. Lesage, C. Condroyer, H. Tomiyama, N. Hattori *et al.* 2009 A LRRK2 G2019S mutation carrier from Turkey shares the Japanese haplotype. *Neurogenetics.* **10**: 271-273.

Polymeropoulos, M. H., C. Lavedan, E. Leroy, S. E. Ide, A. Dehejia *et al.* 1997 Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**: 2045-2047.

Rademakers, R., M. Cruts, and B. C. van, 2004 The role of tau (MAPT) in frontotemporal dementia and related tauopathies. *Hum.Mutat.* **24**: 277-295.

Ramirez, A., A. Heimbach, J. Grundemann, B. Stiller, D. Hampshire *et al.* 2006 Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nat.Genet.* **38**: 1184-1191.

Rogaeva, E., J. Johnson, A. E. Lang, C. Gulick, K. Gwinn-Hardy *et al.* 2004 Analysis of the PINK1 gene in a large cohort of cases with Parkinson disease. *Arch.Neurol.* **61**: 1898-1904.

Ross, O. A., A. T. Braithwaite, L. M. Skipper, J. Kachergus, M. M. Hulihan *et al.* 2008a Genomic investigation of alpha-synuclein multiplication and parkinsonism. *Ann.Neurol.* **63**: 743-750.

Ross, O. A., Y. R. Wu, M. C. Lee, M. Funayama, M. L. Chen *et al.* 2008b Analysis of Lrrk2 R1628P as a risk factor for Parkinson's disease. *Ann.Neurol.* **64**: 88-92.

Satake, W., Y. Nakabayashi, I. Mizuta, Y. Hirota, C. Ito *et al.* 2009 Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat.Genet.* **41**: 1303-1307.

Schultheis, P. J., T. T. Hagen, K. K. O'Toole, A. Tachibana, C. R. Burke *et al.* 2004 Characterization of the P5 subfamily of P-type transport ATPases in mice. *Biochem.Biophys.Res.Commun.* **323**: 731-738.

Seaton, T. A., J. M. Cooper, and A. H. Schapira, 1997 Free radical scavengers protect dopaminergic cell lines from apoptosis induced by complex I inhibitors. *Brain Res.* **777**: 110-118.

Shadrina, M. I., E. V. Semenova, P. A. Slominsky, G. H. Bagyeva, S. N. Illarioshkin *et al.* 2007 Effective quantitative real-time polymerase chain reaction analysis of the parkin gene (PARK2) exon 1-12 dosage. *BMC.Med.Genet.* **8**: 6.

Shimura, H., N. Hattori, S. Kubo, Y. Mizuno, S. Asakawa *et al.* 2000 Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat.Genet.* **25**: 302-305.

Shulman, L. M., 2007 Gender differences in Parkinson's disease. *Gend.Med.* **4**: 8-18.

Siderowf, A., 2001 Parkinson's disease: clinical features, epidemiology and genetics. *Neurol.Clin.* **19**: 565-78, vi.

Sim, C. H., D. S. Lio, S. S. Mok, C. L. Masters, A. F. Hill *et al.* 2006 C-terminal truncation and Parkinson's disease-associated mutations down-regulate the protein serine/threonine kinase activity of PTEN-induced kinase-1. *Hum.Mol.Genet.* **15**: 3251-3262.

Simon-Sanchez, J., C. Schulte, J. M. Bras, M. Sharma, J. R. Gibbs *et al.* 2009 Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat.Genet.* **41**: 1308-1312.

Singleton, A. B., M. Farrer, J. Johnson, A. Singleton, S. Hague *et al.* 2003 alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**: 841.

Snider, S. R., S. Fahn, W. P. Isgreen, and L. J. Cote, 1976 Primary sensory symptoms in parkinsonism. *Neurology* **26**: 423-429.

Snyder, B. J., and C. W. Olanow, 2005 Stem cell treatment for Parkinson's disease: an update for 2005. *Curr.Opin.Neurol.* **18**: 376-385.

Solano, S. M., D. W. Miller, S. J. Augood, A. B. Young, and J. B. Penney, Jr., 2000 Expression of alpha-synuclein, parkin, and ubiquitin carboxy-terminal hydrolase L1 mRNA in human brain: genes associated with familial Parkinson's disease. *Ann.Neurol.* **47**: 201-210.

Song, D. D., C. W. Shults, A. Sisk, E. Rockenstein, and E. Masliah, 2004 Enhanced substantia nigra mitochondrial pathology in human alpha-synuclein transgenic mice after treatment with MPTP. *Exp.Neurol.* **186**: 158-172.

Spillantini, M. G., M. L. Schmidt, V. M. Lee, J. Q. Trojanowski, R. Jakes *et al.* 1997 Alpha-synuclein in Lewy bodies. *Nature* **388**: 839-840.

Spittaels, K., H. C. Van den, D. J. Van, K. Bruynseels, K. Vandezande *et al.* 1999 Prominent axonopathy in the brain and spinal cord of transgenic mice overexpressing four-repeat human tau protein. *Am.J.Pathol.* **155**: 2153-2165.

Sriram, S. R., X. Li, H. S. Ko, K. K. Chung, E. Wong *et al.* 2005 Familial-associated mutations differentially disrupt the solubility, localization, binding and ubiquitination properties of parkin. *Hum.Mol.Genet.* **14**: 2571-2586.

Stamer, K., R. Vogel, E. Thies, E. Mandelkow, and E. M. Mandelkow, 2002 Tau blocks traffic of organelles, neurofilaments, and APP vesicles in neurons and enhances oxidative stress. *J.Cell Biol.* **156**: 1051-1063.

Sun, M., J. C. Latourelle, G. F. Wooten, M. F. Lew, C. Klein *et al.* 2006 Influence of heterozygosity for parkin mutation on onset age in familial Parkinson disease: the GenePD study. *Arch.Neurol.* **63**: 826-832.

Tanaka, Y., S. Engelender, S. Igarashi, R. K. Rao, T. Wanner *et al.* 2001 Inducible expression of mutant alpha-synuclein decreases proteasome activity and increases sensitivity to mitochondria-dependent apoptosis. *Hum.Mol.Genet.* **10**: 919-926.

Tekle-Haimanot, R., M. Abebe, A. Gebre-Mariam, L. Forsgren, J. Heijbel *et al.* 1990 Community-based study of neurological disorders in rural central Ethiopia. *Neuroepidemiology* **9**: 263-277.

Thomas, B., and M. F. Beal, 2007 Parkinson's disease. *Hum.Mol.Genet.* **16 Spec No. 2**: R183-R194.

Tobin, J. E., J. C. Latourelle, M. F. Lew, C. Klein, O. Suchowersky *et al.* 2008 Haplotypes and gene expression implicate the MAPT region for Parkinson disease: the GenePD Study. *Neurology* **71**: 28-34.

Touchman, J. W., A. Dehejia, O. Chiba-Falek, D. E. Cabin, J. R. Schwartz *et al.* 2001 Human and mouse alpha-synuclein genes: comparative genomic sequence analysis and identification of a novel gene regulatory element. *Genome Res.* **11**: 78-86.

Ueda, K., H. Fukushima, E. Masliah, Y. Xia, A. Iwai *et al.* 1993 Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc.Natl.Acad.Sci.U.S.A* **90**: 11282-11286.

Um, J. W., D. S. Min, H. Rhim, J. Kim, S. R. Paik *et al.* 2006 Parkin ubiquitinates and promotes the degradation of RanBP2. *J.Biol.Chem.* **281**: 3595-3603.

Um, J. W., C. Stichel-Gunkel, H. Lubbert, G. Lee, and K. C. Chung, 2009 Molecular interaction between parkin and PINK1 in mammalian neuronal cells. *Mol.Cell Neurosci.* **40**: 421-432.

Unoki, M., and Y. Nakamura, 2001 Growth-suppressive effects of BPOZ and EGR2, two genes involved in the PTEN signaling pathway. *Oncogene* **20**: 4457-4465.

Valente, E. M., P. M. bou-Sleiman, V. Caputo, M. M. Muqit, K. Harvey *et al.* 2004a Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science* **304**: 1158-1160.

Valente, E. M., S. Salvi, T. Ialongo, R. Marongiu, A. E. Elia *et al.* 2004b PINK1 mutations are associated with sporadic early-onset parkinsonism. *Ann.Neurol.* **56**: 336-341.

Van Den Eeden, S. K., C. M. Tanner, A. L. Bernstein, R. D. Fross, A. Leimpeter *et al.* 2003 Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. *Am.J.Epidemiol.* **157**: 1015-1022.

Vilarino-Guell, C., A. I. Soto, S. J. Lincoln, Y. S. Ben, M. Kefi *et al.* 2009 ATP13A2 variability in Parkinson disease. *Hum.Mutat.* **30**: 406-410.

von Campenhausen S., B. Bornschein, R. Wick, K. Botzel, C. Sampaio *et al.* 2005 Prevalence and incidence of Parkinson's disease in Europe. *Eur.Neuropsychopharmacol.* **15**: 473-490.

Wagenfeld, A., C. H. Yeung, K. Strupat, and T. G. Cooper, 1998 Shedding of a rat epididymal sperm protein associated with infertility induced by ornidazole and alpha-chlorohydrin. *Biol.Reprod.* **58**: 1257-1265.

Wakabayashi, K., S. Engelender, M. Yoshimoto, S. Tsuji, C. A. Ross *et al.* 2000 Synphilin-1 is present in Lewy bodies in Parkinson's disease. *Ann.Neurol.* **47**: 521-523.

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. *Neurosci.Lett.* **410**: 121-125.

Webb, J. L., B. Ravikumar, J. Atkins, J. N. Skepper, and D. C. Rubinsztein, 2003 Alpha-Synuclein is degraded by both autophagy and the proteasome. *J.Biol.Chem.* **278**: 25009-25013.

Weingarten, M. D., A. H. Lockwood, S. Y. Hwo, and M. W. Kirschner, 1975 A protein factor essential for microtubule assembly. *Proc.Natl.Acad.Sci.U.S.A* **72**: 1858-1862.

West, A. B., D. J. Moore, S. Biskup, A. Bugayenko, W. W. Smith *et al.* 2005 Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. *Proc.Natl.Acad.Sci.U.S.A* **102**: 16842-16847.

Wider, C., T. Foroud, and Z. K. Wszolek, 2010a Clinical implications of gene discovery in Parkinson's disease and parkinsonism. *Mov Disord.* **25 Suppl 1**: S15-S20.

Wider, C., C. Vilarino-Guell, B. Jasinska-Myga, M. G. Heckman, A. I. Soto-Ortolaza *et al.* 2010b Association of the MAPT locus with Parkinson's disease. *Eur.J.Neurol.* **17**: 483-486.

Wszolek, Z. K., R. F. Pfeiffer, M. H. Bhatt, R. L. Schelper, M. Cordes *et al.* 1992 Rapidly progressive autosomal dominant parkinsonism and dementia with pallido-ponto-nigral degeneration. *Ann.Neurol.* **32**: 312-320.

Xiong, H., D. Wang, L. Chen, Y. S. Choo, H. Ma *et al.* 2009 Parkin, PINK1, and DJ-1 form a ubiquitin E3 ligase complex promoting unfolded protein degradation. *J.Clin.Invest* **119**: 650-660.

Xiromerisiou, G., G. M. Hadjigeorgiou, V. Gourbali, J. Johnson, I. Papakonstantinou *et al.* 2007 Screening for SNCA and LRRK2 mutations in Greek sporadic and autosomal dominant Parkinson's disease: identification of two novel LRRK2 variants. *Eur.J.Neurol.* **14**: 7-11.

Yang, Y., I. Nishimura, Y. Imai, R. Takahashi, and B. Lu, 2003 Parkin suppresses dopaminergic neuron-selective neurotoxicity induced by Pael-R in Drosophila. *Neuron* **37**: 911-924.

Yavich, L., H. Tanila, S. Vepsalainen, and P. Jakala, 2004 Role of alpha-synuclein in presynaptic dopamine recruitment. *J.Neurosci.* **24**: 11165-11170.

Zabetian, C. P., C. M. Hutter, S. A. Factor, J. G. Nutt, D. S. Higgins *et al.* 2007 Association analysis of MAPT H1 haplotype and subhaplotypes in Parkinson's disease. *Ann.Neurol.* **62**: 137-144.

Zabetian, C. P., C. M. Hutter, D. Yearout, A. N. Lopez, S. A. Factor *et al.* 2006a LRRK2 G2019S in families with Parkinson disease who originated from Europe and the Middle East: evidence of two distinct founding events beginning two millennia ago. *Am.J.Hum.Genet.* **79**: 752-758.

Zabetian, C. P., H. Morino, H. Ujike, M. Yamamoto, M. Oda *et al.* 2006b Identification and haplotype analysis of LRRK2 G2019S in Japanese patients with Parkinson disease. *Neurology* **67**: 697-699.

Zarranz, J. J., J. Alegre, J. C. Gomez-Esteban, E. Lezcano, R. Ros *et al.* 2004 The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. *Ann.Neurol.* **55**: 164-173.

Zhang, L., M. Shimoji, B. Thomas, D. J. Moore, S. W. Yu *et al.* 2005a Mitochondrial localization of the Parkinson's disease related protein DJ-1: implications for pathogenesis. *Hum.Mol.Genet.* **14**: 2063-2073.

Zhang, Y., J. Gao, K. K. Chung, H. Huang, V. L. Dawson *et al.* 2000 Parkin functions as an E2-dependent ubiquitin- protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1. *Proc.Natl.Acad.Sci.U.S.A* **97**: 13354-13359.

Zhang, Z. X., and G. C. Roman, 1993 Worldwide occurrence of Parkinson's disease: an updated review. *Neuroepidemiology* **12**: 195-208.

Zhang, Z. X., G. C. Roman, Z. Hong, C. B. Wu, Q. M. Qu *et al.* 2005b Parkinson's disease in China: prevalence in Beijing, Xian, and Shanghai. *Lancet* **365**: 595-597.

Zimprich, A., S. Biskup, P. Leitner, P. Lichtner, M. Farrer *et al.* 2004 Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron* **44**: 601-607.

URL reference list

Diseases & Conditions (AllRefer.com): [http://health.allrefer.com/health/parkinsons-disease-substantia-nigra-and-parkinsons-disease.html] date accessed: 08/10/2010

HRM Assay Design and Analysis Booklet: [http://www.gene-quantification.de/hrm-protocol-cls.pdf] date accessed: 08/10/2010

MRC-Holland: [http://www.mlpa.com] date accessed: 08/10/2010

Parkinson's disease Foundation: [http://www.pdf.org/en/symptoms] date accessed: 08/10/2010

Parkinson's disease Mutation Database: [http://grenada.lumc.nl/LOVD2/TPI/home.php] date accessed: 11/10/2010

PATCH™: [http://www.gene-regulation.com] date accessed: 08/10/2010

Promega gene expression and reporter assays: [http://www.promega.com/pGL4] date accessed: 08/10/2010

Promega pGL4 reporter vectors: [http://www.promega.com/vectors/reporter_vectors.htm#b07] date accessed: 08/10/2010

rVISTA: [http://genome.lbl.gov/vista/index.shtml] date accessed: 08/10/2010

Chapter 2

Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease

Soraya Bardien[a], Rowena Keyser[a], Yandiswa Yako[a], Debbie Lombard[a, b] and Jonathan Carr[b]

[a]Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, University of Stellenbosch, PO Box 19063, Tygerberg 7505, Cape Town, South Africa
[b]Division of Neurology, Faculty of Health Sciences, University of Stellenbosch, PO Box 19063, Tygerberg 7505, Cape Town, South Africa

Published in: Parkinsonism and Related Disorders 2009 Feb;15(2):116-121

Impact factor of journal: 2.406

My contribution to this project: PCR and HRM analysis
Sequencing
Analysis and interpretation of data
Assisted with writing of manuscript

Abstract

Parkinson's disease (PD) is a common movement disorder which may arise from mutations in the *parkin* gene. To date, more than 100 different *parkin* mutations have been reported. The aim of the present study was to determine the frequency of point mutations and homozygous exon deletions in the *parkin* gene in a group of 91 South African patients diagnosed with PD. Mutation screening of the 12 exons of *parkin* was performed using single strand conformation polymorphism analysis and the high-resolution melt technique. Six different mutations were identified: four putative disease-causing missense heterozygous changes (H200Q, D280N, E310D and R402C) and two homozygous exon deletions (exons 3 and 4, and exon 4). The D280N and R402C variants have both previously been described but their pathogenic status has been equivocal. In the present study, the D280N variant was observed in three early-onset PD-affected siblings and was not present in a 63-year-old unaffected sibling. This data provide further support for the pathogenicity of this variant which is situated within the first RING finger of the RING-box. None of the four missense variants were detected in over 100 ethnic-matched control chromosomes. We conclude that point mutations and homozygous exon deletions in the *parkin* gene are not a major cause of PD in the South African population. Further studies on this group of patients are needed to determine the contribution of heterozygous exon deletions and insertions in *parkin*. The present study is the first report on the molecular etiology of PD in South African patients.

Introduction

Parkinson's disease (PD) is characterized clinically by the disabling motor manifestations of rigidity, resting tremor, bradykinesia, and postural instability, as well as 'non-motor' disturbances including cognitive and autonomic dysfunction. PD is estimated to affect roughly 1–2% of the population over the age of 65 years [1].

To date, at least five genes have been shown to cause PD. These include the *PARK2* (*parkin*), *PINK1* and *PARK7* genes for autosomal recessive PD, and the *SNCA* and the *LRRK2* genes for autosomal dominant forms of the disorder [2]. An understanding of the roles of these genes has provided important insights into the molecular patho-physiology of PD and thereby, ultimately, may provide possible targets for therapeutic intervention. It is proposed that the various biological pathways or processes involved include mitochondrial dysfunction, oxidative damage, abnormal protein accumulation, and protein phosphorylation, which may all play a role in dopamine neuronal function and survival [3,4].

A large number of PD-causing mutations have been found in the *parkin* gene, which is situated on chromosome 6q25.2–27, and comprises 12 exons (OMIM 602544). Mutations in *parkin* were first identified about a decade ago in Japanese families with autosomal recessive juvenile Parkinsonism [5]. Since then, *parkin* mutations have been identified in diverse populations worldwide and include whole exon deletions/insertions, small deletions/insertions and point mutations (http://www.thepi.org/altruesite/files/parkinson/Mutations/new_page_1.html). More than 100 different *parkin* mutations have been identified that are spread throughout the gene and currently there do not appear to be any major mutation hot-spots or founder effects in this gene.

Parkin mutations have been found not only in early-onset PD but also in the late-onset (>60 years) forms of PD [6,7,8]. It has therefore been proposed that individuals with late-onset PD should not be excluded from *parkin* mutation screening, particularly if they have a positive family history of the disorder. To date, there have been only 12 reports of genetic studies of Parkinson's disease in African populations [9]. The aim of the present study was to investigate the frequency of point mutations in the *parkin* gene in a group of South African patients. The population studied included unique sub-populations comprising individuals of Afrikaner descent, and Xhosa-speaking individuals from the Southern African Nguni group. The Afrikaner group is known to have had a genetic bottleneck at the end of the 19th century, and a number of

uncommon diseases are found at high frequency in this group, such as variegate porphyria and familial hypercholesterolemia.

Subjects and methods

Study participants

This study was approved by the Committee for Human Research at the University of Stellenbosch (Protocol number 2002/C059). The South African group of patients consisted of 91 unrelated individuals and was recruited from the Movement Disorders clinic at Tygerberg Hospital in Cape Town as well as through the Parkinson's Association of South Africa. Patients recruited from Tygerberg Hospital were examined by a movement disorders specialist (JC) and met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD [10]. Patients recruited through the Parkinson's Association had been diagnosed with Parkinson's disease by other neurologists. Inclusion criteria for recruitment included early-onset and/or a family history of PD.

The average age at onset of the PD patients was 51.8 years ± 28.0, ranging from 17 years to 77 years. A total of 67% was male. The ethnic breakdown was as follows: 57.1% (52/91) Caucasian, 16.5% (15/91) Black, 2.2% (2/91) Indian and 24.2% (22/91) Mixed ancestry. South African Mixed ancestry can be defined as an admixture of the indigenous African populations (San, Khoikhoi or Bantu-speaking) and immigrants from Western Europe, Madagascar, the Malaysian archipelago and India [11]. Ethnic-matched control samples were recruited from healthy unrelated blood donors at a blood donation clinic of the South African Western Province Blood Transfusion Service. These controls were not clinically assessed for any signs of PD and a family history was not taken. A total of 33% and 38% of the Caucasian and Mixed ancestry control samples were male, respectively. As the samples had been ‘de-identified’, no information was available on the ages of these controls.

Genetic analyses

After obtaining written informed consent from the study participants, peripheral blood samples were collected and genomic DNA was extracted according to established methods [12]. Polymerase chain reaction (PCR) primers were designed using Primer3 software for each of the 12 exons of the *parkin* gene [13]. Primer sequences are available from the authors on request. PCR reactions were performed in a 2720 Thermal Cycler (Applied Biosystems, USA). Each 50 µl reaction contained 200 ng template DNA, 20 pmol of each primer, 75 µM dNTPs (Promega,

USA), 1.5 mM $MgCl_2$, 1 × NH_4 buffer (Bioline, UK), 5% formamide (for selected primer sets) and 0.5 U of BIOTAQ DNA polymerase (Bioline). PCR cycling conditions comprised of an initial denaturation step of 94 °C for 5 min, 30 cycles of denaturation at 94 °C for 30 s, annealing at different temperatures according to each primer set for 30 s, extension at 72 °C for 45 s, and a final extension step of 72 °C for 7 min.

Homozygous whole exon deletions were identified by the consistent absence, in four separate attempts, of the PCR product using agarose gel analysis and multiplex PCR. Mutation screening was performed using either single strand conformation polymorphism (SSCP) analysis or the high-resolution melt (HRM) technique. For SSCP analysis, PCR-amplified products were electrophoresed overnight at 22 W at 4 °C in 8% and 10% mildly-denaturing polyacrylamide gels containing 5% glycerol and 15% urea. DNA bands were visualized by silver staining. For the HRM analysis, PCR was performed with incorporation of 1 μM of the fluorescent dye, SYTO9 (Invitrogen, USA), and thereafter the products underwent a melt from 75 °C to 95 °C rising by 0.1 °C each step on a Rotor-Gene 6000 analyzer (Corbett Life Science, Australia). HRM is a mutation scanning method that monitors the shift in fluorescence as a PCR product dissociates from double stranded DNA (dsDNA) to single stranded DNA with increasing temperature. Initially, the fluorescence is high but this diminishes as the dsDNA intercalating dye is released. The observed thermal denaturation profile is characteristic of a specific PCR product and is dependent on its sequence length, base and GC content (HRM Assay Design and Analysis Booklet; http://www.corbettlifescience.com/shared/Rotor-Gene%206000/hrm_corprotocol.pdf).

Samples exhibiting altered mobilities on SSCP gels or different thermal denaturation profiles on HRM were sequenced in order to characterize the sequence variation. Direct sequencing was performed using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems) and analyzed on a 3130 × l Genetic Analyzer (Applied Biosystems). BioEdit version 7.0.1 software was used for the analysis of the sequencing electropherograms [14]. The frequency of selected sequence variants was determined in ethnically-matched control chromosomes using either HRM analysis or by digestion of the PCR product with an appropriate restriction endonuclease.

All sequence variations detected were named according to the nomenclature proposed by den Dunnen and Antonarakis [15] (http://www.hgvs.org/mutnomen/). The genomic sequence for *parkin* was obtained from NCBI's GenBank using accession number NT_007422.

Results

In the present study, six sequence variants were identified in a group of 91 South African patients diagnosed with PD (Table 1). Two novel heterozygous variants (H200Q and E310D) were identified and were absent in over 100 control chromosomes. Four of the six variants identified were heterozygous point mutations and two were homozygous whole exon deletions.

In the case of the four heterozygous missense variants, only the individual who harbored the D280N variant had a confirmed family history of early-onset PD. This variant was also present in two of the proband's PD-affected siblings (AAO 45 years and 47 years), and not in an unaffected 63-year-old sibling. The homozygous deletion of exon 4 was detected in two affected siblings and none of the unaffected family members. The deletion encompassing exons 3 and 4 was found in only one individual who had no family history of the disease.

Using the SSCP and HRM techniques, a number of different polymorphisms were identified in this study that are found spread throughout the gene (Table 2). This illustrates the efficacy of our mutation screening techniques and how highly polymorphic the *parkin* gene is. HRM was able to detect all the sequence variants previously identified by SSCP analysis. Some of the common variants detected include S167N, M192L and R334C which have previously been recorded as pathogenic mutations (http://www.thepi.org/altruesite/files/parkinson/Mutations /new_page_1.html). Most of these polymorphisms have previously been described but we have also detected novel changes including the Q34R, IVS2 + 10C > T and three variants in the 3′-UTR (**16G > A, *94A > G* and **103C > T*).

The D280N (Fig. 1) and H200Q variants were detected using the HRM technique. This is a relatively new technique that has been used mainly to detect known sequence variations. In the present study, we have used it to screen for known and novel disease-causing mutations in a group of South African patients who have previously not been investigated at a molecular level.

Table 1 Mutations identified in the parkin gene in the South African group of PD patients.

	Exon	**Variant**	**No. of patients**	**Patient ethnicity***	**AAO (yrs) & phenotype**	**Family history of PD**	**In ethnic-matched control chromosomes**	**Protein domain**	**Previously reported**
Heterozygous point mutations	5	p.H200Q (CA**C**>CA**G**)	1	Caucasian (65.79)	55 yrs Typical PD with motor fluctuations	yes-probable	0/106 (HRM)	In the unique parkin domain	No
	7	p.D280N (**G**AC>**A**AC)	3 affected siblings	Indian (68.10)	37 yrs 45 yrs 47 yrs Typical PD with motor fluctuations	yes	0/118 (HRM)	In RING1 domain	Oliveira et al. [7]; Lücking et al. [22];
	8	p.E310D (GA**G**>GA**C**)	1	Caucasian (Afrikaner) (51.70)	42 yrs Typical PD with motor fluctuations	no	0/110 (MvaI digest)	Between the RING1 and IBR domains	No
	11	p.R402C (**C**GT>**T**GT)	1	Caucasian (Afrikaner) (42.06)	37 yrs Typical PD with motor fluctuations	no	0/100 (SSCP)	Between the IBR and RING2 domains	Bertoli-Avella et al. [19]; Schlitter et al. [20]; Chaudhary et al. [21]
Homozygous deletions	4		2 affected siblings	Caucasian (Afrikaner) (33.66)	27 yrs (both siblings) Prominent dystonia	yes	ND		yes
	3 & 4		1	Mixed Ancestry (53.44)	27 yrs Foot dystonia, mild symptoms	no	ND		yes

* patient sample number in brackets; AAO, age at onset; ND, not determined

Table 2 Parkin polymorphisms identified in the present study in both PD patients and controls.

Position in gene				Sequence variant	Ethnicity of patients with variant	Frequency in controls (%)
promoter	**Exon**	**Intron**	**3' UTR**			
Core promoter region				-227 A>G	Caucasian & Mixed ancestry	ND
Core promoter region				-258 T>G	Caucasian, Mixed ancestry & Black	14.0 (Caucasian) 18.0 (Mixed)
	Exon 2			**Q34R** (CA**G**>C**G**G)	Mixed ancestry	3.2
	Exon 2			P37P (CC**G**>CC**A**)	Mixed ancestry & Black	ND
		Intron 2		IVS2 +10C>T	Mixed ancestry	0
		Intron 2		IVS2 +25T>C	Mixed ancestry & Black	ND
		Intron 2		IVS2 +35G>A	Caucasian & Mixed ancestry	ND
		Intron 3		IVS3 -20T>C	Caucasian	ND
	Exon 4			**S167N** (A**G**C>A**A**C)	Mixed ancestry	7.9
	Exon 5			**M192L** (**A**TG>**C**TG)	Mixed ancestry	6.7
	Exon 6			C238C (TG**C**>TG**T**)	Black	ND
	Exon 7			L261L (TT**A**>TT**G**)	Black	ND
		Intron 7		IVS7 -35G>A	Caucasian & Black	ND
		Intron 7		IVS7 -68C>G	Caucasian	ND
		Intron 8		IVS8 +43A>G	Mixed ancestry & Black	ND
		Intron 8		IVS8 +48C>T	Caucasian & Black	ND
		Intron 8		IVS8 -21_-17del	Mixed ancestry	ND
	Exon 9			**R334C** (**C**GC>**T**GC)	Mixed ancestry	2.0
	Exon 10			**V380L** (**G**TA>**C**TA)	Caucasian, Mixed ancestry & Black	27.2 (Caucasian)
	Exon 11			**D394N** (**G**AT>**A**AT)	Caucasian, Mixed ancestry, Black & Indian	6.0 (Caucasian)
	Exon 11			R402R (CG**T**>CG**C**)	Mixed ancestry & Black	ND
			3' UTR	**16G >A*	Mixed ancestry	ND
			3' UTR	**94A>G*	Caucasian	ND
			3' UTR	**103C>T*	Black	ND

ND, not determined; Non-synonymous sequence variants are indicated in bold text.

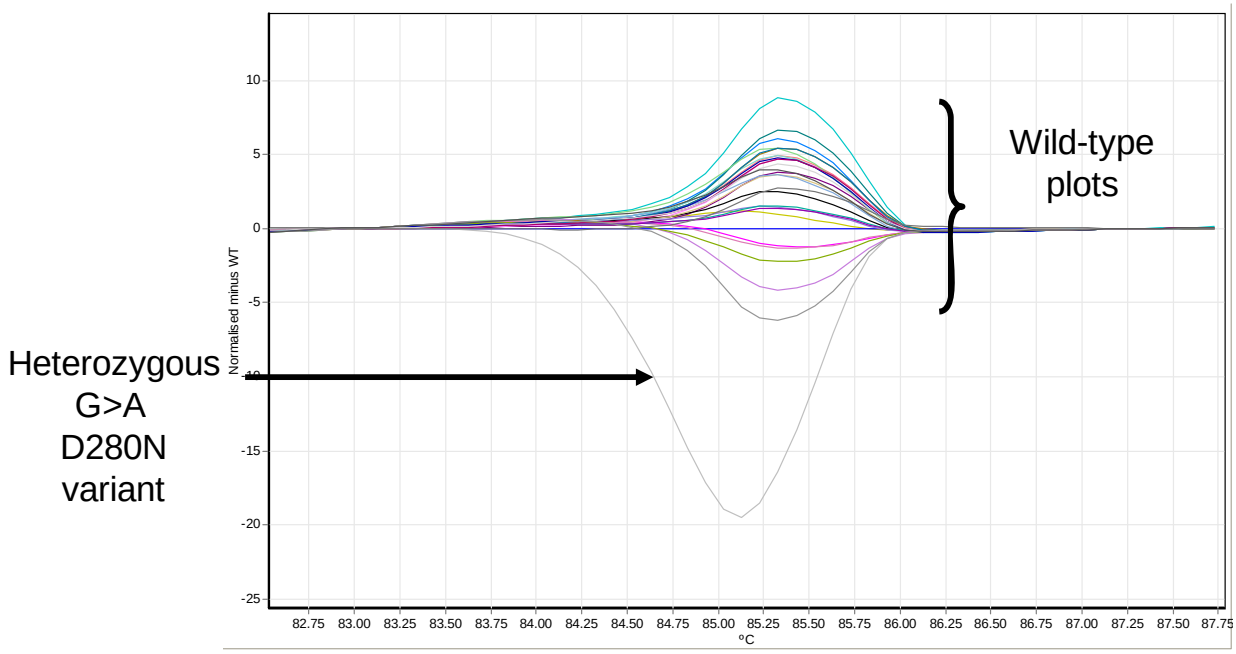


Figure 1 High-resolution melt analysis: difference plot exhibiting the heterozygous D280N variant in exon 7 that could easily be differentiated from that of wild type samples. Normalized fluorescence subtracted from a reference sample is plotted on the Y-axis and temperature in degrees Celsius is plotted on the X-axis.

Discussion

To date there have been a very limited number of studies on PD in African populations, particularly genetic studies [9], whereas, by comparison, European and American PD populations have been extensively investigated. This is the first report on the molecular etiology of PD in a South African group of patients. Only six sequence variants (four heterozygous missense variants and two homozygous deletions) were detected in the *parkin* gene in 91 individuals. No mutations were identified in any of the Black South African patients included in our study. PD in Black individuals has been reported to be atypical with a rigid-akinetic variant being more common [16] and it is therefore plausible that Black patients may harbor mutations in novel PD genes.

In the present study, two novel putative pathogenic variants were identified. One is the H200Q variant in exon 5 which was identified in one PD individual with a probable family history of the disorder. The pathogenicity of this sequence variant is unknown but the variant was not found in 106 control chromosomes and is situated in the unique *parkin* domain, in which other PD-causing mutations have been reported [17,18]. Similarly, the novel E310D variant was observed only in one PD patient but not in 110 control chromosomes. This variant was also observed in

the patient's 51-year-old brother who, at this stage, does not exhibit any signs of PD. The E310D variant is situated between the RING1 and IBR domains of parkin.

The R402C and D280N variants have both previously been described although their pathogenicity has been questioned. The R402C variant was previously observed in one patient of unknown ethnicity (AAO = 35 years) [19], one German PD patient (AAO = 65 years) [20] and one late-onset Indian patient [21]. The present study is the fourth report of the R402C variant in a PD patient with apparently no family history of PD. In all four patients harboring R402C, the variant was found in the heterozygous state and a second mutation was not found. It is therefore interesting to speculate that either the second mutation in all these cases has yet to be found or that the R402C variant on its own is sufficient to cause the disorder. The fact that it is situated between the IBR domain and the second RING finger, and that it affects an evolutionarily conserved amino acid suggests that it may be pathogenic. The D280N variant was previously reported in one family with early-onset PD (ancestry unknown) [22] and one Caucasian PD patient (AAO = 51 years) [7]. In the former study, the pathogenicity of D280N was questioned as both patients harboring this variant also had other *parkin* variants and in the latter study, it was speculated that this variant in the heterozygous state could act as a susceptibility allele for the development of late-onset PD. D280N is positioned within the first RING finger domain of parkin and it has been shown that both RING finger motifs are necessary for binding to the E2 ubiquitin-conjugating enzyme, UbcH7, another component of the ubiquitin proteasome system [23]. Further support for the pathogenicity of this variant is that in the present study it was observed in three PD-affected siblings and not in an unaffected family member or in 118 control chromosomes. It is also possible, however, that the four heterozygous *parkin* variants identified in the present study are rare non-pathogenic polymorphisms [24].

It can be concluded that point mutations or homozygous exon deletions in the *parkin* gene are not a major cause of PD in these South African patients. Importantly, the lack of mutations identified is not due to the mutation screening method since a number of non-pathogenic sequence variants were identified in this study (Table 2). Therefore, either heterozygous *parkin* deletions or mutations in other genes account for the disorder in these South African PD patients. In the present study, we show that high-resolution melt curve analysis can be used successfully to screen for unknown sequence variants in the *parkin* gene. This technique is sufficiently sensitive to detect 1-bp differences in PCR-amplified products and is less time and labor-intensive than other methods such as SSCP analysis. The novel variant identified in intron 2, IVS2 + 10C > T, was detected in one PD patient and was not present in 100 control

chromosomes. This variant could possibly have an effect on splicing of the gene but further studies are necessary to investigate this. Also, four other novel heterozygous variants, Q34R, **16G > A, *94A > G* and **103C > T* were identified and may prove useful for genetic association studies involving the *parkin* gene. Furthermore, the S167N, M192L and R334C variants have all been previously reported and have been speculated to be disease-causing but the present study shows that the variants are common in controls and are therefore unlikely to be associated with PD.

More studies are needed to determine whether the molecular etiology of PD across diverse populations is accounted for by the handful of the same PD-causative genes. In this regard, South Africa's unique population, particularly the Black and the Mixed ancestry sub-populations, have the potential of harboring either novel mutations or novel PD-causative genes.

Acknowledgments

We thank the PD patients and their family members for their participation in this study. We also gratefully acknowledge the Western Province Blood Transfusion Service for providing the control samples. This work was financially supported by the South African Medical Research Council, the Harry and Doris Crossley Foundation and the University of Stellenbosch, South Africa.

References

[1] M.C. de Rijk, L.J. Launer, K. Berger, M.M. Breteler, J.F. Dartigues and M. Baldereschi et al., Prevalence of Parkinson's disease in Europe: a collaborative study of population-based cohorts. Neurologic diseases in the elderly research group, Neurology 54 (11 Suppl. 5) (2000), pp. S21–S23.

[2] B. Thomas and M.F. Beal, Parkinson's disease, Hum Mol Genet 16 (Spec No. 2) (2007), pp. R183–R194.

[3] J.T. Greenamyre and T.G. Hastings, Biomedicine. Parkinson's-divergent causes, convergent mechanisms, Science 304 (5674) (2004), pp. 1120–1122.

[4] M.B. Feany, New genetic insights into Parkinson's disease, N Engl J Med 351 (19) (2004), pp. 1937–1940.

[5] T. Kitada, S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura and S. Minoshima et al., Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism, Nature 392 (6676) (1998), pp. 605–608.

[6] T. Foroud, S.K. Uniacke, L. Liu, N. Pankratz, A. Rudolph and C. Halter et al., Heterozygosity for a mutation in the parkin gene leads to later onset Parkinson disease, Neurology 60 (5) (2003), pp. 796–801.

[7] S.A. Oliveira, W.K. Scott, E.R. Martin, M.A. Nance, R.L. Watts and J.P. Hubble et al., Parkin mutations and susceptibility alleles in late-onset Parkinson's disease, Ann Neurol 53 (5) (2003), pp. 624–629.

[8] M. Sun, J.C. Latourelle, F. Wooten, M.F. Lew, C. Klein and H.A. Shill et al., Influence of heterozygosity for parkin mutation on onset age in familial Parkinson disease: the GenePD study, Arch Neurol 63 (6) (2006), pp. 826–832.

[9] N.U. Okubadejo, An analysis of genetic studies of Parkinson's disease in Africa, Parkinsonism Relat Disord (2007 Sep 17) [Epub ahead of print PMID: 17881276].

[10] W.R.G. Gibb and A.J. Lees, A comparison of clinical and pathological features of young-onset and old-onset Parkinson's disease, Neurology 38 (9) (1988), pp. 1402–1406.

[11] G.T. Nurse, J.S. Weiner and T. Jenkins, The growth of hybrid communities. In: G.T. Nurse, J.S. Weiner and T. Jenkins, Editors, The peoples of Southern Africa and their affinities, Clarendon Press, Oxford (1985), pp. 218–224 [chapter 9].

[12] V.A. Corfield, J.C. Moolman, R. Martell and P.A. Brink, Polymerase chain reaction-based detection of MN blood group-specific sequences in the human genome, Transfusion 33 (2) (1993), pp. 119–124.

[13] S. Rozen and H.J. Skaletsky, Primer3 on the WWW for general users and for biologist programmers. In: S. Krawetz and S. Misener, Editors, Bioinformatics methods and protocols: methods in molecular biology, Humana Press, Totowa, NJ (2000), pp. 365–386.

[14] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp Ser 41 (1999), pp. 95–98.

[15] J.T. den Dunnen and E. Antonarakis, Nomenclature for the description of human sequence variations, Hum Genet 109 (1) (2001), pp. 121–124.

[16] K.R. Chaudhuri, M.T.M. Hu and D.J. Brooks, Atypical parkinsonism in Afro-Caribbean and Indian origin immigrants to the UK, Mov Disord 15 (1) (2000), pp. 18–23.

[17] P.J. Kahle, U. Leimer and C. Haass, Does failure of parkin-mediated ubiquitination cause juvenile parkinsonism?, Trends Biochem Sci 25 (11) (2000), pp. 524–527.

[18] A. Yamamoto, A. Friedlein, Y. Imai, R. Takahashi, P.J. Kahle and C. Haass, Parkin phosphorylation and modulation of its E3 ubiquitin ligase activity, J Biol Chem 280 (5) (2005), pp. 3390–3399.

[19] A.M. Bertoli-Avella, J.L. Giroud-Benitez, A. Akyol, E. Barbosa, O. Schaap and H.C. van der Linde et al., Novel parkin mutations detected in patients with early-onset Parkinson's disease, Mov Disord 20 (4) (2005), pp. 424–431.

[20] A.M. Schlitter, M. Kurz, J.P. Larsen, D. Woitalla, T. Müller and J.T. Epplen et al., Parkin gene variations in late-onset Parkinson's disease: comparison between Norwegian and German cohorts, Acta Neurol Scand 113 (1) (2006), pp. 9–13.

[21] S. Chaudhary, M. Behari, M. Dihana, P.V. Swaminath, S.T. Govindappa and S. Jayaram et al., Parkin mutations in familial and sporadic Parkinson's disease among Indians, Parkinsonism Relat Disord 12 (4) (2006), pp. 239–245.

[22] C.B. Lücking, A. Durr, V. Bonifati, J. Vaughan, G. De Michele and T. Gasser et al., Association between early-onset Parkinson's disease and mutations in the parkin gene, N Engl J Med 342 (21) (2000), pp. 1560–1567.

[23] H. Shimura, N. Hattori, S. Kubo, Y. Mizuno, S. Asakawa and S. Minoshima et al., Familial Parkinson disease gene product, parkin, is a ubiquitin–protein ligase, Nat Genet 25 (3) (2000), pp. 302–305.

[24] D.M. Kay, D. Moran, L. Moses, P. Poorkaj, C.P. Zabetian and J. Nutt et al., Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients, Ann Neurol 61 (1) (2007), pp. 47–54.

Chapter 3

Analysis of exon dosage using MLPA in South African Parkinson's disease patients

Rowena J. Keyser[1] Contact Information, Debbie Lombard[1, 2], Rene Veikondis[3], Jonathan Carr[2] and Soraya Bardien[1]

[1]Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, University of Stellenbosch, P.O. Box 19063, Tygerberg 7505, Cape Town, South Africa
[2]Division of Neurology, Faculty of Health Sciences, University of Stellenbosch, Cape Town, South Africa
[3]Central Analytical Facility, University of Stellenbosch, Cape Town, South Africa

Published in: Neurogenetics 2010 Jul; 11(3):305-312

Impact factor of journal: 3.486

My contribution to this project:
Planning of study
Setting up the MLPA method and optimizing
Conducting MLPA analysis
Sequencing
Analysis and interpretation of data
Writing of manuscript

Abstract

Genomic rearrangements (exon dosage) are common mutations reported in Parkinson's disease (PD) patients. In the present study, we aimed to investigate the prevalence of genomic rearrangements in 88 South African patients with predominantly early-onset PD (age-at-onset ≤50 years). The multiplex ligation-dependent probe amplification method was used to detect exon dosage changes. Two commercially available probe kits, SALSA P051 and P052, were used and together the kits consisted of probes for exons of α-synuclein, *parkin, PINK1, DJ-1, LRRK2, UCH-L1, ATP13A2, LPA, TNFRSF9, CAV2, CAV1, GCH1*, and two-point mutations. We identified exonic rearrangements in *parkin* and α-synuclein in 8% of South African patients from different ethnic groups. One patient had a whole-gene triplication of α-synuclein; representing only the fourth family with this mutation reported to date. We found six patients with *parkin* mutations who had either heterozygous duplications and deletions, or homozygous deletions. A false positive result of an exonic deletion detected in one patient turned out to be homozygous point mutation (Y258X) in *PINK1*. No exonic rearrangements were found in four of the PD genes; *LRRK2, PINK1, DJ-1*, and *ATP13A2*. Mutations in *parkin* were the predominant genetic cause; however, the frequency of exon dosage in our study group is low compared with previous studies. This indicates the possible involvement of other as yet unidentified PD genes in the development of the disease in the South African population.

Introduction

Parkinson's disease (PD; OMIM 168600) is a debilitating neurodegenerative disorder characterized by progressive loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc), and prominent involvement of cognition and the autonomic system. To date, genetic defects in six genes, namely, α-synuclein (*SNCA*), *parkin* (*PARK2*), PTEN-induced putative kinase 1 (*PINK1*), oncogene DJ-1 (*DJ-1*), leucine-rich repeat kinase 2 (*LRRK2*), and ATPase type 13A2 (*ATP13A2*), have been conclusively implicated in the pathogenesis of PD [1]. Monogenic forms of PD are thought to represent less than 10% of cases in most populations. PD occurs worldwide, but the prevalence and incidence appear to exhibit substantial geographic and ethnic variability [2], with generally lower rates reported in Africa.

Whole exon insertions or deletions (i.e., exon dosage), primarily in *parkin*, are a common mutation in PD patients, and have been identified in autosomal recessive and sporadic forms of the disease [3, 4]. Exonic deletions in *parkin* were first identified in Japanese families with autosomal recessive juvenile parkinsonism [5]. To date, more than 100 different mutations have been identified in *parkin* including point mutations and exonic rearrangements such as duplications and deletions [6–8], which can present in either homozygous or compound heterozygous states. Whole-gene multiplications (duplications or triplications) have been identified only in α-synuclein and lead to over-expression of the protein. It was found that individuals with duplications have a classical PD phenotype, whereas those with triplications have earlier onset, faster disease progression, marked dementia and frequent dysautonomia [9–12]. Genomic rearrangements have also previously been reported for *PINK1* [13] and *DJ-1* [14] but have not been found in *LRRK2* or *ATP13A2*. Due to the fact that mutation screening methods such as high resolution melt or DNA sequencing cannot detect exon dosage, the possibility exists that important disease causing mutations are missed using these traditional methods. It is therefore important to include exon dosage studies when conducting PD mutation screening in order to detect all possible pathogenic mutations.

In the present study, we conducted an exon dosage analysis using the multiplex ligation-dependent probe amplification (MLPA) (http://www.mrc-holland.com) method, in order to determine the prevalence of genomic rearrangements in the known PD genes in patients from the various South African sub-population groups. This method has been used successfully previously by other groups to identify exon dosage [15, 16]. MLPA is a rapid method for detection of exonic deletions, insertions, and multiplications and can also detect specific

recurrent point mutations. An advantage of MLPA is that it is possible to perform a multiplex polymerase chain reaction (PCR) reaction in which 33 exons (using kit P051) and 31 exons (using kit P052) of PD-causing genes are screened simultaneously using a single PCR primer-pair. MLPA is cost effective, technically simple, reproducible, and large numbers of samples can be tested simultaneously, and it was therefore made use of in the present study.

Materials and methods

Study participants

The study was approved by the Committee for Human Research at the University of Stellenbosch, South Africa (Protocol number 2002/C059). For the present study, 88 unrelated PD patients, from all South African ethnic groups, were recruited from the Movement Disorders clinic at Tygerberg Hospital as well as from the Parkinson's Association of South Africa. The patients recruited from Tygerberg Hospital had been examined by a movement disorder specialist (JC). The patients met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD [17]. The majority of patients had an apparently sporadic form of the disease, but in families with a positive family history (34.1%) both autosomal recessive and autosomal dominant patterns of inheritance were present. Inclusion criteria for recruitment of patients were early onset and/or a positive family history of PD.

The ethnic breakdown was as follows: 34.1% South African Afrikaner, 25% Caucasian, 23.9% mixed ancestry (defined in 18), 13.6% Black, and 3.4% Indian. The average age at onset (AAO) of the study group was 50 years and ranged from 17 to 77 years. The percentage of males in our study group was 64.8%.

MLPA assay

For each study participant, a blood sample was collected for genetic analysis and genomic DNA was extracted using established methods. The MLPA method was made use of in order to detect exon dosage changes, caused by genomic rearrangements, of the known PD genes. Two commercially available probe kits, SALSA P051-B1 and P052-B1 Parkinson MLPA kits (MRC Holland, Amsterdam, The Netherlands; http://www.mlpa.com) were used for this assay. Together, the kits consisted of probes for all exons of α-synuclein, *parkin*, and *PINK1*, and specific exons of *DJ-1* (exons 1,3,5,7); *LRRK2* (exons 1,2,10,15,27,41,49); *UCH-L1* (exons 1,4,5,9); *ATP13A2* (exons 2,9); *LPA* (exon 31); *TNFRSF9* (exon 2); *CAV2* (exon 3); *CAV1* (exon 3); *GCH1* (exons 1,2,3,5,6) and two-point mutations (A30P in α-synuclein and G2019S in

LRRK2). The probes were distributed between two sets of probe mixes. All exons of *parkin* were included in both kits, and the probe sequences of the same exon were different between the two kits. The assay was performed for all the 88 samples for both kits, according to the manufacturer's protocol.

Initially, 150 ng of genomic DNA was denatured at 98ºC for 5 min and then allowed to hybridize to the MLPA probe set P051 and P052 (in independent reactions for each probe set). Thereafter, ligation was performed at 54ºC for 15 min using a Ligase-65 enzyme. PCR reactions were performed in a 2720 Thermal Cycler (Applied Biosystems, Foster City, USA) for 33 cycles (95ºC for 30 s; 60ºC for 30 s; 72ºC for 1 min) and a final 20 min step at 72ºC. The primers used for the PCR reactions were the SALSA PCR forward primer (labeled) 5′-FAM-GGG TTC CCT AAG GGT TGG A-3′ and SALSA PCR reverse primer (unlabeled) 5′-GTG CCA GCA AGA TCC AAT CTA GA-3′.The PCR fragments were analyzed on a 3130x1 Genetic Analyzer (Applied Biosystems) with GeneScan™-500 LIZ™ Size Standard (Applied Biosystems) and GeneMapper software v3.7 (Applied Biosystems). The manufacturer's recommendations were followed for the electrophoresis and analysis of the fragments on the Genetic Analyzer. The size standard made it possible to identify the peaks corresponding to specific exons. Results were first visually analyzed to discard failed samples. Thereafter, peak height results were exported from GeneMapper to a Microsoft office Excel spreadsheet for copy number calculations.

The results were analyzed as follows. The relative peak height (RPH) of each exon was obtained by dividing the single peak height by the sum of peak heights of all the internal control probes. A ratio was then generated by comparing each RPH to the mean value of the corresponding RPHs from a reference wild-type sample. Ratios between 0.7–1.3 were considered to be normal (sample contains no exon rearrangements); a heterozygous deletion was expected at a ratio between 0.3–0.6, a heterozygous duplication between 1.4–1.6, and a triplication at a ratio of ≥1.7. Absence of a peak would indicate either a possible point mutation at that probe recognition site or a homozygous deletion (ratio 0.0). Possible mutation positive results were repeated for confirmation.

Verification using quantitative PCR

Possible exonic deletions or multiplications detected by the MLPA assay were verified by means of quantitative PCR (qPCR) on a LightCycler 1.5 (Roche Diagnostics, Mannheim, Germany) using the LightCycler® FastStart DNA MasterPLUS SYBR Green I kit (Roche Diagnostics). qPCR was performed in triplicate for each sample. The reagents used for amplification in 10 µl

reactions were: 2 µl Hybridization FastStart Mix (Roche Diagnostics), 0.5 µM of each primer and 25 ng of genomic DNA. The PCR conditions were: 95ºC for 10 min, 95ºC for 10 s, 55ºC for 10 s, 72ºC for 15 s (45 cycles). For this method, the beta globin gene (*HBB*) served as the internal standard (i.e., reference gene) for subsequent delta-delta-CT (cycle threshold) calculations. The PD gene and *HBB* gene ratios generated by delta-delta-CT were within the same ranges to that produced for the MLPA assay.

Results

In the present study, seven out of 88 (8%) South African PD affected probands were found to have genomic rearrangements (Table 1). Six patients had exonic deletions or duplications in the *parkin* gene and one patient had a whole-gene triplication of α-synuclein. One patient was found to have a point mutation in the *PINK1* gene.

Parkin mutations

In one male patient of Black ethnicity, a heterozygous duplication of exon 2 and a heterozygous deletion of exon 9 were detected in *parkin* using the MLPA method (Fig. 1a). The same mutations were observed in an affected female sibling. Both siblings had an AAO of <50 years and presented with typical tremor predominant PD.

A heterozygous duplication of *parkin* exon 2 was found in a female patient of Afrikaner ethnicity who had no family history of the disease and an AAO of 56 years. The clinical features in this patient were of tremor predominant PD with a typical progressive course over 19 years. This patient had previously been screened for point mutations in *parkin* and no mutations had been identified [18].

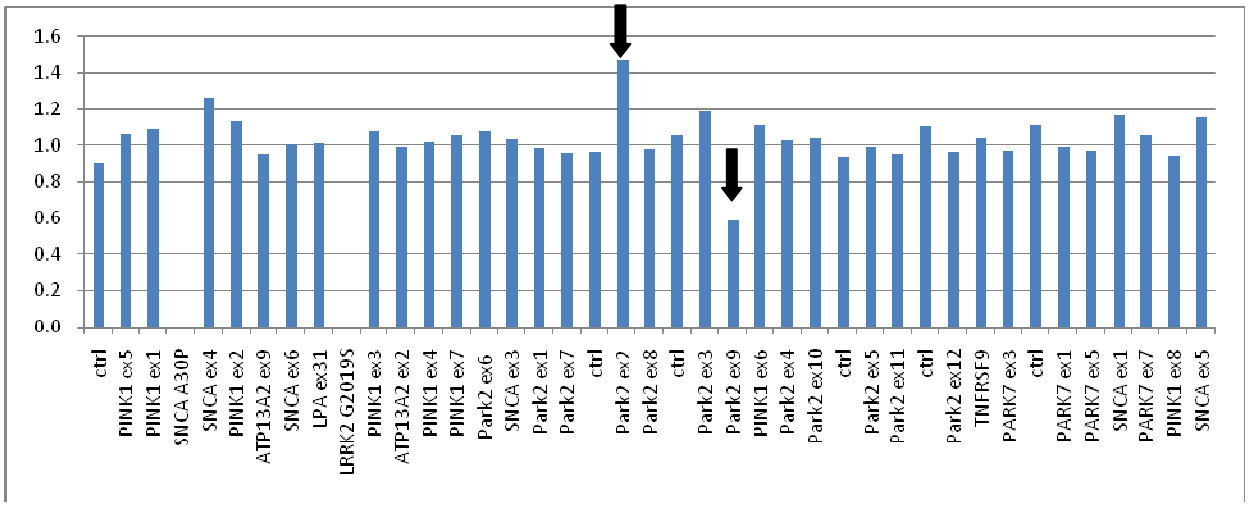
One male patient of mixed ancestry had heterozygous duplications of *parkin* exons 2 and 3. This patient was diagnosed with atypical PD (bilateral, symmetrical Parkinsonism) and schizophrenia and was treated with clozapine 400 mg daily. This individual had no family history of the disease and had an AAO of 50 years. At this stage it is uncertain whether this individual has the duplication of exons 2 and 3 in *cis* or *trans* as all three of his unaffected siblings have the same duplication of the two exons. Further haplotype studies using microsatellites and single nucleotide polymorphisms at the *parkin* locus are needed to resolve this matter.

Table 1 Mutations identified using MLPA kits P051 and P052

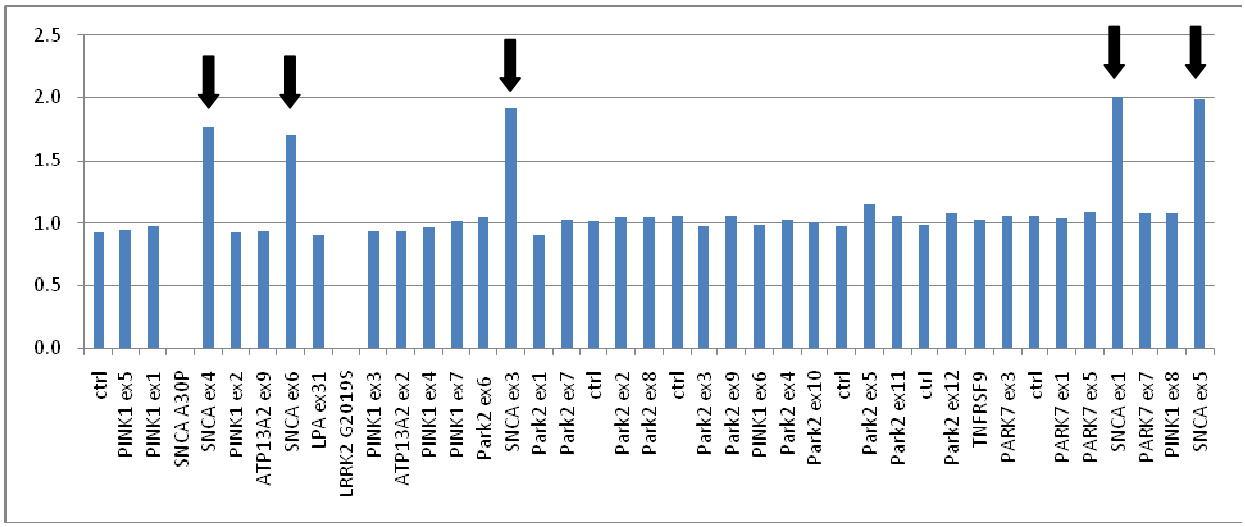
Gene	Zygosity	Mutation	AAO (years)	Family history	Ethnicity [a]	Present in affected family member(s)
Parkin	Heterozygous	Duplication of exon 2 + deletion of exon 9	45	+	Black (37.12)	One sibling (37.13)
	Heterozygous	Duplication of exon 2	56	-	Caucasian (Afrikaner) (55.54)	None
	Heterozygous	Duplication of exon 2 and exon 3	50	-	Mixed ancestry (68.22)	None
	Heterozygous	Deletion of exon 3	25	+	Caucasian (77.60)	None
	Homozygous [b]	Deletion of exon 3 and exon 4	27	-	Mixed ancestry (53.44)	None
	Homozygous [b]	Deletion of exon 4	27	+	Caucasian (Afrikaner) (56.45)	One sibling (56.43)
α-synuclein		Triplication	46	+	Caucasian (French-Italian) (42.35)	None
PINK1	Homozygous	Y258X	37	+	Indian (68.10)	Two siblings (71.93, 71.95)

AAO = age at onset
[a] = patient sample number shown in brackets
[b] = previously reported in Bardien *et al*., 2009[18]

a


1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
ctrl
PINK1 ex5
PINK1 ex1
SNCA A30P
SNCA ex4
PINK1 ex2
ATP13A2 ex9
SNCA ex6
LPA ex31
LRRK2 G2019S
PINK1 ex3
ATP13A2 ex2
PINK1 ex4
PINK1 ex7
Park2 ex6
SNCA ex3
Park2 ex1
Park2 ex7
ctrl
Park2 ex2
Park2 ex8
ctrl
Park2 ex3
Park2 ex9
PINK1 ex6
Park2 ex4
Park2 ex10
ctrl
Park2 ex5
Park2 ex11
ctrl
Park2 ex12
TNFRSF9
PARK7 ex3
ctrl
PARK7 ex1
PARK7 ex5
SNCA ex1
PARK7 ex7
PINK1 ex8
SNCA ex5

b


2.5
2.0
1.5
1.0
0.5
0.0
ctrl
PINK1 ex5
PINK1 ex1
SNCA A30P
SNCA ex4
PINK1 ex2
ATP13A2 ex9
SNCA ex6
LPA ex31
LRRK2 G2019S
PINK1 ex3
ATP13A2 ex2
PINK1 ex4
PINK1 ex7
Park2 ex6
SNCA ex3
Park2 ex1
Park2 ex7
ctrl
Park2 ex2
Park2 ex8
ctrl
Park2 ex3
Park2 ex9
PINK1 ex6
Park2 ex4
Park2 ex10
ctrl
Park2 ex5
Park2 ex11
ctrl
Park2 ex12
TNFRSF9
PARK7 ex3
ctrl
PARK7 ex1
PARK7 ex5
SNCA ex1
PARK7 ex7
PINK1 ex8
SNCA ex5

c

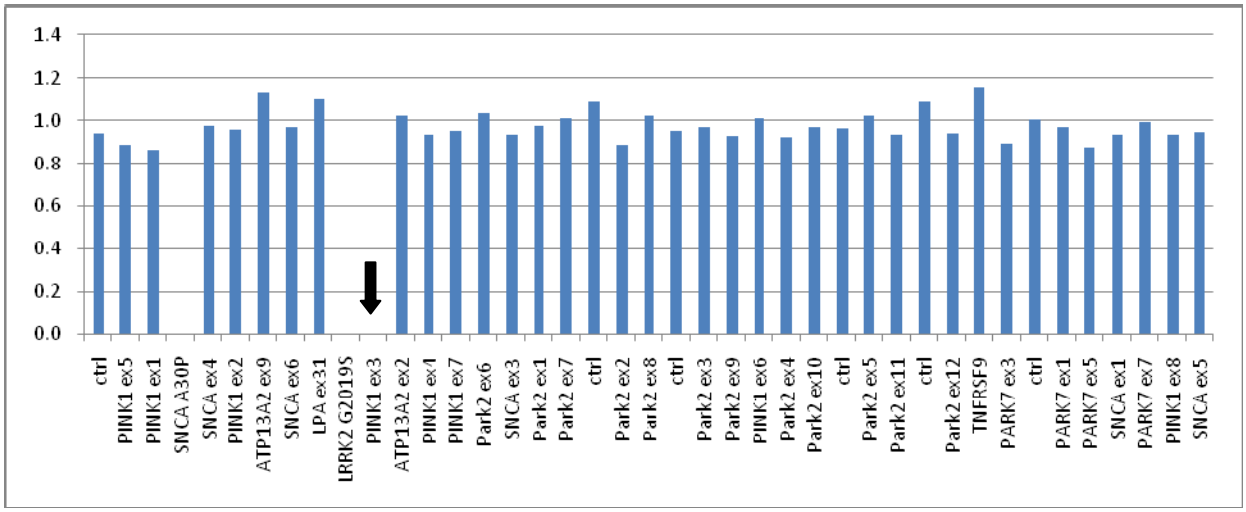

1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
ctrl
PINK1 ex5
PINK1 ex1
SNCA A30P
SNCA ex4
PINK1 ex2
ATP13A2 ex9
SNCA ex6
LPA ex31
LRRK2 G2019S
PINK1 ex3
ATP13A2 ex2
PINK1 ex4
PINK1 ex7
Park2 ex6
SNCA ex3
Park2 ex1
Park2 ex7
ctrl
Park2 ex2
Park2 ex8
ctrl
Park2 ex3
Park2 ex9
PINK1 ex6
Park2 ex4
Park2 ex10
ctrl
Park2 ex5
Park2 ex11
ctrl
Park2 ex12
TNFRSF9
PARK7 ex3
ctrl
PARK7 ex1
PARK7 ex5
SNCA ex1
PARK7 ex7
PINK1 ex8
SNCA ex5

Fig. 1 a MLPA analysis results of an individual with a heterozygous duplication of *parkin* exon 2 (ratio = 1.5) and heterozygous deletion of *parkin* exon 9 (ratio = 0.6). Arrows indicate the probes that were affected by the genomic rearrangements in the two exons of *parkin*. The vertical axis represents the ratios calculated, and the horizontal axis represents the different probes within the SALSA P051 Parkinson MLPA kit. b MLPA results of an individual with a whole-gene triplication of α-synuclein. Arrows indicate the probe amplification products of the α-synuclein exons that were affected by triplication of the gene: exon 1 (ratio = 2), exon 3 (ratio = 1.9), exon 4 (ratio = 1.8), exon 5 (ratio = 2), and exon 6 (ratio = 1.7). c MLPA results of an individual with an Y258X point mutation in *PINK1* exon 3 which prevented binding of the probe for this exon (ratio = 0.0)

One male Caucasian patient had a heterozygous deletion of *parkin* exon 3, with a possible positive family history and an AAO of 25 years. The patient presented with gradual onset of dystonia in the legs, followed by onset of mild features of PD 9 years later. In this patient also, only one *parkin* mutation was identified as a screen for point mutations in *parkin* [18] did not yield any mutations.

Furthermore, MLPA detected the previously reported homozygous deletions of *parkin* exons 3 and 4 in one female patient of mixed ancestry with an AAO of 27 years and no family history of the disease [18]. The previously reported homozygous deletion of *parkin* exon 4 in two siblings of Afrikaner ethnicity was also detected by this method. Both individuals had an AAO of 27 years and an autosomal recessive family history [18].

α-Synuclein mutations

Triplication of the entire α-synuclein gene (Fig. 1b) was identified in one patient of French-Italian ethnicity with an AAO of 46 years. The patient presented initially with typical features of PD, with levodopa responsiveness. He later developed dementia, associated with psychosis, and autonomic failure. He died at the age of 55 years with duration of illness of about 9 years. This individual had a family history of possibly autosomal dominant PD and had a PD-affected sister who had died at the age of 48 years. The sister had not been genetically tested.

PINK1 mutations

Using MLPA, we detected an apparent homozygous deletion of *PINK1* exon 3 in one Indian female patient with a positive family history of the disease and an AAO of 37 years (Fig. 1c). The qPCR verification method did not, however, confirm the presence of the deletion of *PINK1* exon 3. Subsequent DNA sequencing revealed the presence of a homozygous C>A point mutation (Y258X; Fig. 2) at the target site for the *PINK1* exon 3 MLPA probe. When assessed, the patient had had symptoms for 30 years; her disease had started asymmetrically, and remained so. Her symptoms were relatively well controlled on 500 mg levodopa daily. She reported

recently some difficulty with memory for recent events. This mutation was also observed in two affected siblings in a homozygous state and in two unaffected family members in a heterozygous state. The AAO of the other two affected family members were 45 and 47 years, respectively. The mutation leads to a premature stop codon which results in a significantly truncated PINK1 protein. The site of the mutation is within a highly conserved region of the gene and 13 amino acids upstream from the functional serine/threonine kinase domain of PINK1 (Fig. 3).

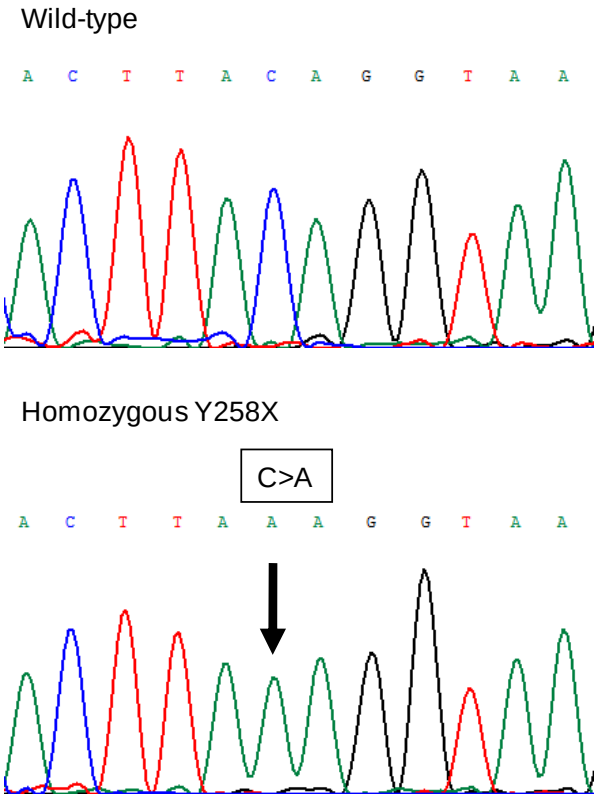


Fig. 2 Chromatogram illustrating the Y258X point mutation in *PINK1* exon 3 that was shown to co-segregate with PD in an Indian family

```
Human   NISAGSSSEAILNTMSQELVPASRVALAGEYGAVTYRKSKRGPKQLAPHPNIIRVLRAFT 282
Chimp   NISAGSSSEAILNTMSQELVPASRVALAGEYGAVTYRKSKRGPKQLAPHPNIIRVLRAFT 360
Mouse   NISAGSSSEAILSKMSQELVPASRVALAGEYGAVTYRRSRDGPKQLAPHPNIIRVFRAFT 281
Rat     NISAGSSSEAILSKMSQELVPASRMALDGEYGAVTYRRSRDGPKQLAPHPNIIRVFRAFT 281
        ***********..**********:** *********:*: **************:****
```

Fig. 3 Sequence alignment (using ClustalW) of PINK1 amino acid sequences of human (NP_115785.1), chimp (XP_001164912.1), mouse (NP_081156.2), and rat (XP_216565.2). The box indicates the position of the Y258X variant that leads to a premature stop codon. The serine/threonine kinase domain is shaded in grey. Identical amino acids are indicated by *asterisks*

The results for *parkin* and α-synuclein listed in Table 1 were detected by both MLPA kits and verified by qPCR on a LightCycler (Roche). Furthermore, in two patients, possible heterozygous deletions of *parkin* exon 5 were detected by kit P051 in two independent experiments. These deletions were, however, not detected with kit P052 and were also not confirmed by qPCR. DNA sequencing revealed that the common *parkin* polymorphism M192L [18] disrupts the ligation step in MLPA for kit P051 and therefore appears as a false deletion of exon 5 in the subsequent MLPA analysis. Kit P052 is not affected by the M192L variant as the probes are located at a different position.

Discussion

We report the results of an exon dosage study which was conducted on PD patients from all South African ethnic groups. In the present study, it was found that the frequencies of *parkin* and α-synuclein genomic rearrangements in a group of 88 patients were 6.8% and 1.1%, respectively. Genomic rearrangements were not restricted to any specific ethnic group. No genomic rearrangements were found in four of the major PD genes, namely, *LRRK2, PINK1, DJ-1*, and *ATP13A2*.

Parkin presented with the highest proportion of genomic rearrangements in our study group compared with the other known PD genes. *Parkin* is predominantly involved in early onset forms of PD, which possibly explains the higher frequency of mutations detected in our study group. Approximately 50% of all known *parkin* mutations are thought to fall in the category of genomic rearrangements [19]. These mutations are proposed to result in the absence of functional proteins because of frame shift or alterations of splicing [4]. Exonic rearrangements often occur due to recombination events between repetitive elements. The size, relative orientation, distance between copies, and shared percentage identities of these repetitive elements can cause the genomic region to be more susceptible to rearrangements. These factors can also affect the type of rearrangement that occurs [20–23]. It is suggested that the high frequency of exonic rearrangement mutations in *parkin* is due to its large introns and the presence of numerous repetitive elements that lead to instability.

The frequency of exonic rearrangements detected in the present study in *parkin*, however, is lower than that reported in previous studies in which up to 50% of the patients were found to have these mutations [19, 24]. No *parkin* mutations (point or exonic rearrangements) were detected in one of our patients who has a juvenile age at onset of 17 years. In two patients, only

single heterozygous exonic rearrangements were detected in *parkin* and no point mutations were found. It has been suggested that individuals heterozygous for *parkin* deletion mutations might suffer from haploinsufficiency because of reduced expression and enzymatic activity of the protein [25]. This might confer a greater risk for developing PD. It should be cautioned however that since *parkin* heterozygous deletions have been found in controls [26] the pathogenicity of these variants is not certain.

The MLPA method identified a triplication of α-synuclein in an individual of French-Italian ethnicity with a positive family history. Of note, this family is only the 4th world wide to be reported to have an α-synuclein triplication. Previously, triplications have been reported in an English-German family (Iowan kindred) [12], a Swedish-American family (Lister kindred) [27], and recently a French family [28]. According to unconfirmed information, the other affected individuals in this family in previous generations had later ages at onset and less severe phenotypes than the proband and his sibling. This points to a possible duplication to triplication event and it would be interesting to determine whether our family is related to either the Lister family, in which this phenomenon had occurred, or the French triplication family. Duplication mutations in α-synuclein have been reported in nine families. The mechanism underlying α-synuclein multiplication is currently unknown. The severity of the clinical phenotype is dependent on the level of exon dosage; duplication mutations are associated with typical, late-onset PD [9, 10] whereas, triplication mutations lead to a more severe form of early-onset PD with autonomic dysfunction and subsequent dementia [12, 27]. This was also observed in our patient. Therefore, the higher the expression levels of α-synuclein, the more malignant the PD phenotype. This suggests that there is more widespread neurodegeneration in patients with higher levels of α-synuclein expression. α-Synuclein is suggested to play a role in synaptic vesicle recycling, as well as storage and compartmentalization of neurotransmitters [29, 30].

In the present study, a point mutation (Y258X) in *PINK1* was found to cosegregate with PD in a family of Indian ethnicity. This mutation has previously been reported in an individual of Asian ethnicity (either Chinese, Malay, or Indian) with sporadic early-onset PD who presented with restless legs syndrome, resting tremors and bradykinesia [31]. The authors speculated that the restless leg syndrome in this patient was directly related to the specific mutation, however, none of the patients in our Y258X-positive family presented with restless legs syndrome. The Y258X mutation causes a significant truncation of the PINK1 protein, which results in loss of the functional serine/threonine kinase domain. PINK1 is a protein kinase that is localized to the mitochondrion and is ubiquitously expressed in the human brain [32]. Recent studies aimed at

elucidating the function of PINK1, have found that it has neuroprotective properties against mitochondrial dysfunction and proteasomally-induced apoptosis [33–36]. PINK1 phosphorylates specific mitochondrial proteins, and in so doing modulates their functions. Mutations in *PINK1* are the second most common cause after *parkin* of autosomal recessive early-onset PD, and it has been shown that the frequency of *PINK1* mutations ranges from 1% to 8% in patients of different ethnicities [37–39]. Furthermore, some *PINK1* mutation carriers are known to exhibit psychiatric symptoms [40], but none of the Y258X-positive individuals in the present study have psychiatric problems.

The present study validates the importance of including exon dosage studies as part of mutation screening for PD-causing mutations. The MLPA method proved effective in detecting a variety of genomic rearrangements in our study group as well as a point mutation. However, our results show that it is necessary to verify all MLPA positive results with another method as false positive results could arise due to point mutations located in the MLPA probe recognition site (e.g., Y258X in PINK1). Furthermore, we found that, *parkin* exon 5 in kit P051 was prone to false positive (deletion) results as a result of the M192L polymorphism.

Acknowledgments

We thank the study participants. Ms. Lindsey Adams is thanked for providing technical advice. This work was supported by the South African Medical Research Council, the Harry and Doris Crossley Foundation, and the University of Stellenbosch, South Africa.

Conflicts of interest

None declared.

References

1. Lesage S, Brice A (2009) Parkinson's disease: from monogenic forms to genetic susceptibility factors. Hum Mol Genet 18(R1):R48-59. Review.

2. Alves G, Forsaa EB, Pedersen KF, Dreetz Gjerstad M, Larsen JP (2008) Epidemiology of Parkinson's disease. J Neurol Sep;255 Suppl 5:18-32.

3. Hedrich K, Kann M, Lanthaler AJ, Dalski A, Eskelson C, Landt O et al (2001) The importance of gene dosage studies: mutational analysis of the parkin gene in early-onset parkinsonism. Hum Mol Genet 10(16):1649-1656.

4. Hedrich K, Marder K, Harris J, Kann M, Lynch T, Meija-Santana H et al (2002) Evaluation of 50 probands with early-onset Parkinson's disease for Parkin mutations. Neurology 58(8):1239-1246.

5. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S et al (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 392(6676):605-608.

6. Periquet M, Latouche M, Lohmann E, Rawal N, De Michele G, Ricard S (2003) French Parkinson's Disease Genetics Study Group; European Consortium on Genetic Susceptibility in Parkinson's Disease. Parkin mutations are frequent in patients with isolated early-onset parkinsonism. Brain 126(Pt 6):1271-1278.

7. Shadrina MI, Semenova EV, Slominsky PA, Bagyeva GH, Illarioshkin SN, Ivanova-Smolenskaia II, Limborska SA (2007) Effective quantitative real-time polymerase chain reaction analysis of the parkin gene (PARK2) exon 1-12 dosage. BMC Med Genet 26;8:6.

8. Sun M, Latourelle JC, Wooten GF, Lew MF, Klein C, Shill HA et al (2006) Influence of heterozygosity for parkin mutation on onset age in familial Parkinson disease: the GenePD study. Arch Neurol 63(6):826-832.

9. Chartier-Harlin MC, Kachergus J, Roumier C, Mouroux V, Douay X, Lincoln S et al (2004) Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. Lancet 364(9440):1167-1169.

10. Ibáñez P, Bonnet AM, Débarges B, Lohmann E, Tison F, Pollak P et al (2004) Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. Lancet 364(9440):1169-1171.

11. Ross OA, Braithwaite AT, Skipper LM, Kachergus J, Hulihan MM, Middleton FA et al (2008) Genomic investigation of alpha-synuclein multiplication and parkinsonism. Ann Neurol 63(6):743-750.

12. Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J et al (2003) alpha-Synuclein locus triplication causes Parkinson's disease. Science 302(5646):841.

13. Marongiu R, Brancati F, Antonini A, Ialongo T, Ceccarini C, Scarciolla O et al (2007) Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum. Hum Mutat 28(1):98.

14. Djarmati A, Hedrich K, Svetel M, Schäfer N, Juric V, Vukosavic S 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. Hum Mutat 23(5):525.

15. Djarmati A, Guzvić M, Grünewald A, Lang AE, Pramstaller PP, Simon DK et al (2007) Rapid and reliable detection of exon rearrangements in various movement disorders genes by multiplex ligation-dependent probe amplification. Mov Disord 22(12):1708-1714.

16. Scarciolla O, Brancati F, Valente EM, Ferraris A, De Angelis MV, Valbonesi S et al (2007) Multiplex ligation-dependent probe amplification assay for simultaneous detection of Parkinson's disease gene rearrangements. Mov Disord 22(15):2274-2278.

17. Gibb WRG, Lees AJ (1988) A comparison of clinical and pathological features of young-onset and old-onset Parkinson’s disease. Neurology 38(9):1402-1406.

18. Bardien S, Keyser RJ, Yako Y, Lombard D, Carr J (2009) Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson’s disease. Parkinsonism and Related Disorders 15(2):116-121.

19. Hedrich K, Eskelson C, Wilmot B, Marder K, Harris J, Garrels J et al (2004) Distribution, type, and origin of Parkin mutations: review and case studies. Mov Disord 19(10):1146-1157.

20. Ballabio A, Bardoni B, Guioli S, Basler E, Camerino G (1990) Two families of low-copy-number repeats are interspersed on Xp22.3: implications for the high frequency of deletions in this region. Genomics 8(2):263-270.

21. Chen KS, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault AC, Lee CC, Lupski JR (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. Nat Genet 17(2):154-163.

22. Yen PH, Li XM, Tsai SP, Johnson C, Mohandas T, Shapiro LJ (1990) Frequent deletions of the human X chromosome distal short arm result from recombination between low copy repetitive elements. Cell 61(4):603-610.

23. Lee JA, Lupski JR et al (2006) Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 52(1):103-21. Review.

24. Hattori N, Kitada T, Matsumine H, Asakawa S, Yamamura Y, Yoshino H et al (1998) 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. Ann Neurol 44(6):935-941.

25. Klein C, Lohmann-Hedrich K, Rogaeva E, Schlossmacher MG, Lang AE (2007) Deciphering the role of heterozygous mutations in genes associated with parkinsonism. Lancet Neurol 6(7):652-662. Review.

26. Brüggemann N, Mitterer M, Lanthaler AJ, Djarmati A, Hagenah J, Wiegers K et al (2009) Frequency of heterozygous Parkin mutations in healthy subjects: need for careful prospective follow-up examination of mutation carriers. Parkinsonism Relat Disord 15(6):425-429.

27. Farrer M, Kachergus J, Forno L, Lincoln S, Wang DS, Hulihan M (2004) Comparison of kindreds with parkinsonism and alpha-synuclein genomic multiplications. Ann Neurol 55(2):174-179.

28. Ibáñez P, Lesage S, Janin S, Lohmann E, Durif F, Destée A et al (2009) Alpha-synuclein gene rearrangements in dominantly inherited parkinsonism: frequency, phenotype, and mechanisms.Arch Neurol 66(1):102-108.

29. Abeliovich A, Schmitz Y, Fariñas I, Choi-Lundberg D, Ho WH, Castillo PE et al (2000) Mice lacking alpha-synuclein display functional deficits in the nigrostriatal dopamine system. Neuron 25(1):239-252.

30. Yavich L, Tanila H, Vepsäläinen S, Jäkälä P (2004) Role of alpha-synuclein in presynaptic dopamine recruitment. J Neurosci 24(49):11165-11170.

31. Tan EK, Yew K, Chua E, Puvan K, Shen H, Lee E et al (2006) PINK1 mutations in sporadic early-onset Parkinson's disease. Mov Disord 21(6):789-793.

32. Gandhi S, Muqit MM, Stanyer L, Healy DG, Abou-Sleiman PM, Hargreaves I et al (2006) PINK1 protein in normal human brain and Parkinson's disease. Brain 129(Pt7):1720-1731.

33. Clark IE, Dodson MW, Jiang C, Cao JH, Huh JR, Seol JH et al (2006) Drosophila pink1 is required for mitochondrial function and interacts genetically with parkin. Nature 441(7097):1162-1166.

34. Park J, Lee SB, Lee S, Kim Y, Song S, Kim S et al (2006) Mitochondrial dysfunction in Drosophila PINK1 mutants is complemented by parkin. Nature 441(7097):1157-1161.

35. Petit A, Kawarai T, Paitel E, Sanjo N, Maj M, Scheid M et al (2005) Wild-type PINK1 prevents basal and induced neuronal apoptosis, a protective effect abrogated by Parkinson disease-related mutations. J Biol Chem 280(40):34025-34032.

36. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, Gispert S et al (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 304(5674):1158-1160.

37. Bonifati V, Rohé CF, Breedveld GJ, Fabrizio E, De Mari M, Tassorelli C et al (2005) Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes. Neurology 65(1):87-95.

38. Klein C, Djarmati A, Hedrich K, Schäfer N, Scaglione C, Marchese R et al (2005) PINK1, Parkin, and DJ-1 mutations in Italian patients with early-onset parkinsonism. Eur J Hum Genet 13(9):1086-1093.

39. Rogaeva E, Johnson J, Lang AE, Gulick C, Gwinn-Hardy K, Kawarai T et al (2004) Analysis of the PINK1 gene in a large cohort of cases with Parkinson disease. Arch Neurol 61(12):1898-1904.

40. Ephraty L, Porat O, Israeli D, Cohen OS, Tunkel O, Yael S, Hatano Y et al (2007) Neuropsychiatric and cognitive features in autosomal-recessive early parkinsonism due to PINK1 mutations. Mov Disord 15;22(4):566-569.

Chapter 4

LRRK2 G2019S mutation: frequency and haplotype data in South African Parkinson's disease patients

Soraya Bardien[1], Angelica Marsberg[1], Rowena Keyser[1], Debbie Lombard[2], Suzanne Lesage[3, 4, 5], Alexis Brice[3, 4, 5] and Jonathan Carr[2]

[1]Division of Molecular Biology and Human Genetics, University of Stellenbosch, 4th Floor Fisan Building, PO Box 19063, Tygerberg, Cape Town, 7505, South Africa

[2]Division of Neurology, Stellenbosch University, Cape Town, South Africa

[3]Université Pierre et Marie Curie-Paris6, Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, UMR-S975, Paris, France

[4]INSERM, U975, Paris, France

[5]CNRS, UMR 7225, Paris, France

Published in: Journal of Neural Transmission 2010 Jul;117(7):847-853

Impact factor of journal: 2.259

My contribution to this project: Assisted with HRM analysis
Assisted with analyzing data
Assisted with writing of manuscript

Abstract

Mutations in the leucine-rich repeat kinase 2 gene (*LRRK2*) are the most significant genetic cause of Parkinson's disease (PD). The exact function of LRRK2 is currently unknown but the presence of multiple protein interaction domains including WD40 and ankyrin indicates that it may act a scaffold for assembly of a multi-protein signaling complex. The G2019S mutation in *LRRK2* represents the most clinically relevant PD-causing mutation and has been found in both familial and sporadic forms of the disorder. This mutation is situated in the highly conserved kinase MAPKKK domain, and has been found in up to 40% of PD patients from North African Arabic, 30% of Ashkenazi Jewish and ~10% of Portuguese and Spanish populations. Although extensively investigated in numerous European and North American populations, studies on the frequency of G2019S in African countries have been rare. The present study is the first on the South African population. High-resolution melt analysis was used to identify the G2019S mutation and it was found in 2% (4/205) of the patients studied. G2019S was not found in any of the Black PD patients screened. In all four G2019S-positive probands the mutation was shown to be present on the common haplotype referred to as haplotype 1. This reveals that the four South African G2019S-positive probands (three Caucasian and one of mixed ancestry) share a common ancestor with the other haplotype 1-associated families reported worldwide.

Introduction

Parkinson's disease (PD) is a common progressive neurodegenerative disorder affecting roughly 1–2% of the population aged 65 years or older (de Rijk et al. 2000). The lifetime risk of developing PD has been calculated to be 2.0% for men and 1.3% for women (Elbaz et al. 2002). PD typically occurs in individuals over the age of 60 years but approximately 15% of PD occurs in younger individuals. In this young-onset form of the disorder, disease-causing mutations have been identified in a number of genes. One of these, the leucine-rich repeat kinase 2 gene (*LRRK2*) has been implicated in both autosomal dominant and sporadic forms of PD, and is thought to be the most common genetic cause of PD identified to date (Giasson and Van Deerlin 2008; Correia et al. 2010).

LRRK2 encodes a protein consisting of multiple domains including Roc (Ras of complex proteins), COR (C-terminal of Roc), MAPKKK (mitogen-activated protein kinase) and WD40 (Bosgraaf and Van Haastert 2003; Marin et al. 2008). Although the exact function of *LRRK2* is currently unknown the presence of the Roc and MAPKKK domains suggests a role in the regulation of signal transduction cascades (Guo et al. 2006).

The G2019S mutation in the MAPKKK domain likely increases the kinase activity of the protein (West et al. 2005). This mutation is present on only a limited number of haplotypes (Lesage and Brice 2009), of which the commonest, referred to as haplotype 1, is thought to have originated in the Ashkenazi Jewish population (Lesage et al. 2010). Two other G2019S-carrying haplotypes have been found in three families of European descent (haplotype 2, Zabetian et al. 2006a) and in one Turkish and two Japanese patients (haplotype 3, Zabetian et al. 2006b; Pirkevi et al. 2009). Mainly due to founder effects, the prevalence of G2019S is 37–41% in North African Arab PD patients (Lesage et al. 2006, 2008) and 10–27% in Ashkenazi Jewish PD patients (Orr-Urtreger et al. 2007; Ozelius et al. 2006). The prevalence is estimated to be 3–7% of familial and 1–2% of sporadic PD in European and North American Caucasian patients (Di Fonzo et al. 2005; Healy et al. 2008; Lesage et al. 2007; Kachergus et al. 2005).

In contrast, G2019S is rare in Asian populations. Studies have found that the G2385R variant in the WD40 region is frequent in Asian populations but infrequent in non-Asian populations. There have been reports of a significant association of the G2385R variant with PD in Han Chinese from Taiwan, Singapore and mainland China (Fung et al. 2006; An et al. 2008; Farrer et al. 2007; Di Fonzo et al. 2006; Tan et al. 2007). Recently, a second risk factor for the

development of PD, the R1628P variant in the COR domain, has been found in the Chinese population in independent studies (Ross et al. 2008; Zhang et al. 2009; Tan et al. 2008). The distribution and frequency of mutations and genetic susceptibility factors in LRRK2 therefore shows ethnic-specific differences.

On the entire African continent, the only genetic studies on PD have been conducted on patients from the North African Arabic countries (Morocco, Algeria, and Tunisia) and Nigeria. G2019S was found to be frequent in the North African countries but, in contrast, neither this mutation nor any other mutations in the exon containing this mutation (exon 41) was present in a group of Nigerian patients (Okubadejo et al. 2008). Okubadejo et al. (2008) urged that the usefulness for Africa's populations of commercially available PD-predictive genetic tests or therapeutic modalities, as developed by first world countries, will only be revealed once we know the genetic basis of this disorder in Africa's diverse populations. In the present study we investigated the frequency of the G2019S mutation in a group of PD patients in South Africa of different ethnic groups.

Materials and methods

Study participants

This study was approved by the Committee for Human Research at the University of Stellenbosch (Protocol number 2002/C059). A total of 205 South African PD probands were recruited for genetic analysis either from the Movement Disorders clinic at Tygerberg Hospital in Cape Town or from the Parkinson's Association of South Africa. Patients recruited from Tygerberg Hospital had been examined by a movement disorders specialist (J.C.) and met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD (Gibb and Lees 1988). The majority (65%) of the probands were male. The average age at onset (AAO) of PD was 54.4 ± 12.4 years (range 17–80). About a third of the probands (35%) had a positive family history of the disorder and the remainder (65%) either had no family history of PD or there was insufficient information about the family. In 104 of the patients the *PARK2* and *PINK1* genes had previously been screened for mutations. The ethnic breakdown of the study group was as follows: 42% Caucasian, 31% Afrikaner, 17% mixed ancestry; 8% Black and 2% Indian. South African mixed ancestry can be defined as an admixture of the indigenous African populations (San, Khoikhoi or Bantu-speaking) and immigrants from Western Europe, Madagascar, the Malaysian archipelago and India (Patterson et al. 2010). A total of 79 Caucasian

controls were recruited from the Western Province Blood Transfusion Service clinics in Cape Town. These individuals had not been clinically assessed for the presence of PD.

Genetic analysis

After obtaining written informed consent from the study participants, peripheral blood samples were collected and genomic DNA was extracted according to established methods. To detect the *LRRK2* G2019S mutation, the following polymerase chain reaction (PCR) primers were designed: forward primer: 5′ gca cag aat ttt tga tgc ttg 3′ and reverse primer: 5′ gag gtc agt ggt tat cca tcc 3′. PCR reactions were performed in a 2720 Thermal Cycler (Applied Biosystems, Foster City, USA). Each 25 µl reaction contained 200 ng template DNA, 10 pmol of each primer, 75 µM dNTPs (Promega, USA), 1.5 mM $MgCl_2$, 1× NH_4 buffer (Bioline, UK), 0.25 units of BIOTAQ DNA polymerase (Bioline, UK) and 2 µM of the green fluorescent dye, SYTO9 (Invitrogen, USA). PCR cycling conditions comprised of an initial denaturation step of 94°C for 5 min, 30 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 45 s, and a final extension step of 72°C for 7 min.

High-resolution melt (HRM) analysis from 75 to 95°C rising by 0.1°C each step was performed on a Rotor-Gene 6000 analyzer (Corbett Life Science, Australia). HRM is a mutation scanning method that monitors the shift in fluorescence as a PCR product dissociates from double stranded DNA (dsDNA) to single stranded DNA (ssDNA) with increasing temperature. Third generation fluorescent dyes such as SYTO9 have low toxicity in an amplification reaction and therefore can be used at higher concentrations for greater saturation resulting in increased melt sensitivity and higher resolution melt profiles. The observed thermal denaturation profile is characteristic of a specific PCR product and is dependant on its sequence length, base and GC content (Erali et al. 2008; Wittwer et al. 2003). Samples exhibiting altered thermal denaturation profiles were sequenced in order to characterize the sequence variation. Direct sequencing was performed using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems) and analyzed on a 3130xl Genetic Analyzer (Applied Biosystems).

Haplotype analysis

Haplotypes were constructed using the following microsatellite markers and single nucleotide polymorphisms (SNPs): D12S2194, rs726624, D12S2514, rs28903073, D12S2516, D12S2518, rs3952596, rs7308560, D12S2519 and D12S2520, as described previously (Lesage et al. 2010). Briefly, genotyping of the four SNPs was performed using direct sequencing. Genotyping of the six microsatellite markers was performed using fluorescently labeled primers and the products

were analyzed on a 3130xl Genetic Analyzer (Applied Biosystems) using GeneMapper Software (version 3.7; Applied Biosystems). Haplotypes were constructed manually.

Results

Only 4 out of 205 (2%) (95% CI 0.8–4.9) South African PD probands screened were shown to harbor the *LRRK2* G2019S mutation, all in a heterozygous state. G2019S was not present in 158 Caucasian control chromosomes. The mutation was successfully detected using the cost-effective HRM technique (Fig. 1). As can be seen from both the normalized (Fig. 1a) and the difference graphs (Fig. 1b), the G2019S variant produces a different melt profile to that of the wild type. All G2019S-positive samples as determined by HRM were verified by direct sequencing.

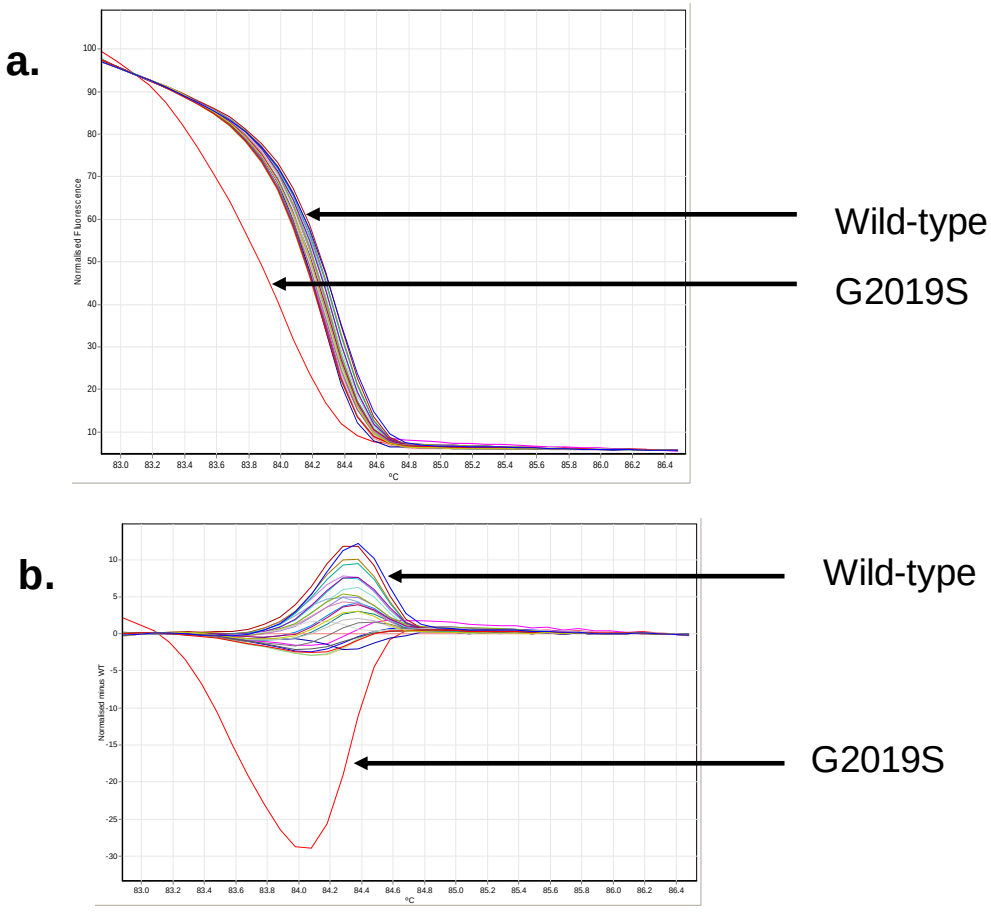
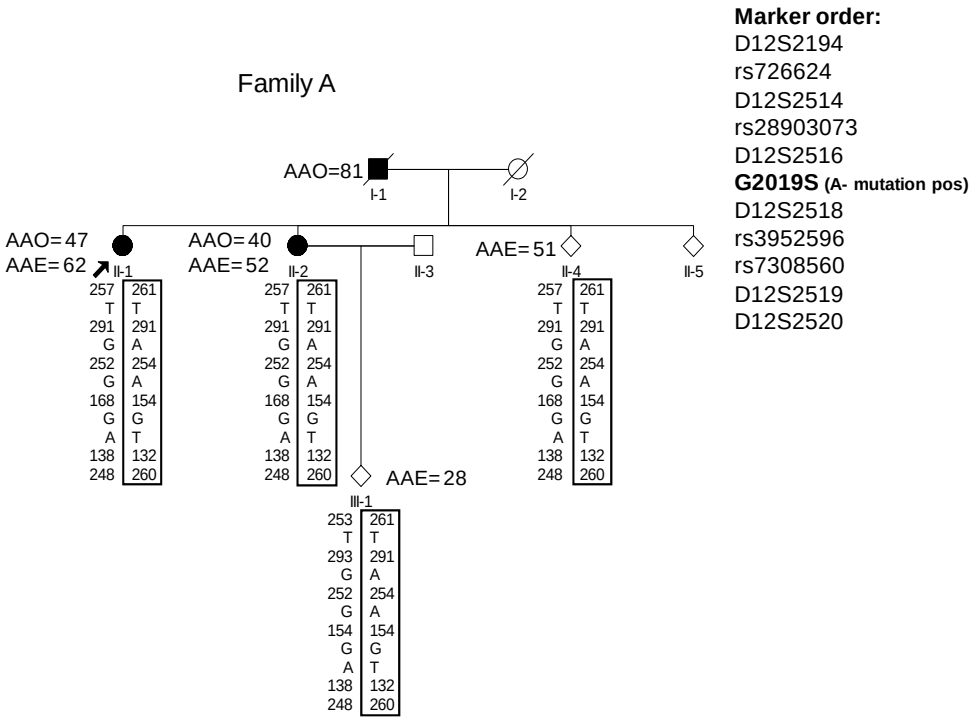
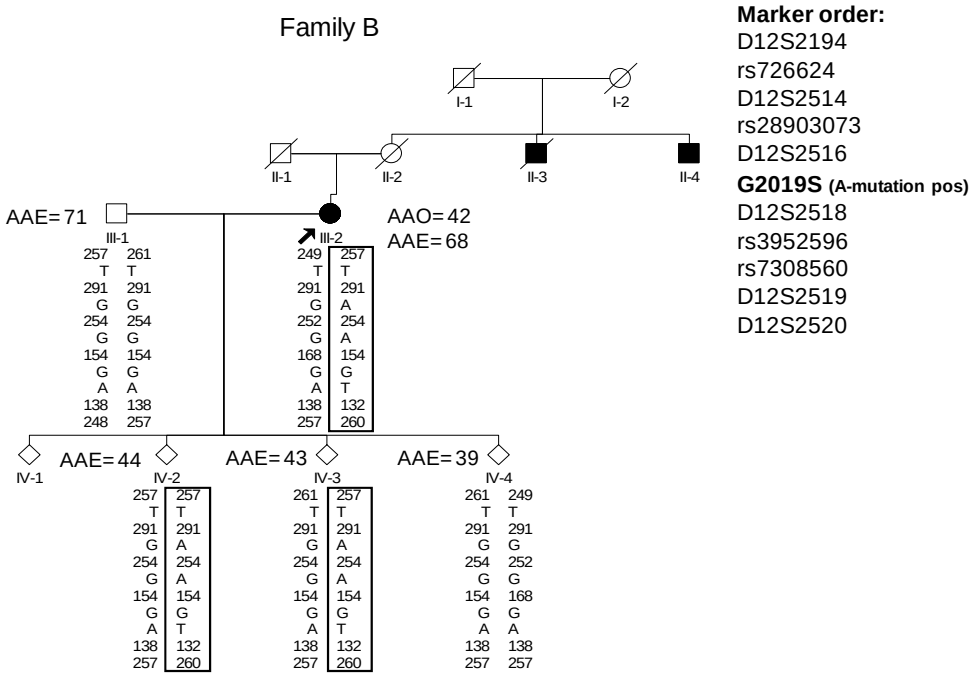


Fig. 1 High-resolution melt (HRM) analysis of the G2019S mutation illustrating that the mutation can be distinguished from the wild-type allele. a Normalized graph. b Difference curve

The one G2019S-positive individual (individual II-1; Fig. 2, family A) is of South African mixed ancestry and her affected father (individual I-1) was German. The proband developed typical features of PD at the age of 47. She showed a good response to dopamine therapy, complicated by marked motor fluctuations, with subsequent development of autonomic failure and psychosis.

The family exhibits an autosomal dominant inheritance pattern with reduced penetrance as individuals II-4 and III-1 also harbor the G2019S mutation but are currently asymptomatic. Individual II-2 who also has the G2019S mutation was diagnosed with schizophrenia in her early 20s. She presented at age 40 with marked bilateral tremor and bradykinesia which was possibly related to her antipsychotic medication depot fluphenazine. She was treated with carbilev/levodopa and demonstrated improvement in her tremor and bradykinesia.





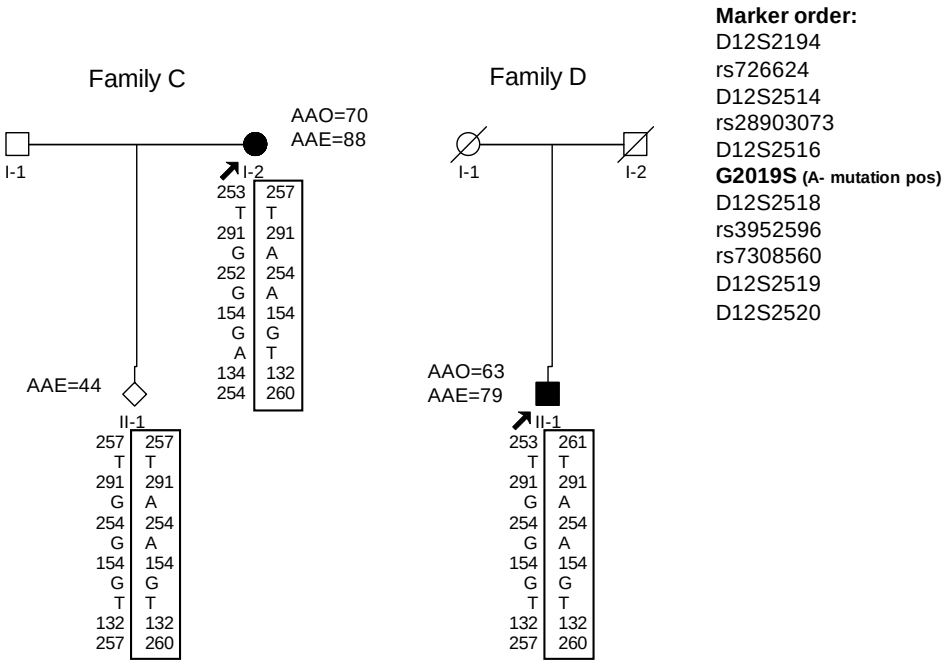


Fig. 2 Pedigrees of the South African families (family A, B, C and D) with the G2019S mutation. The disease-associated haplotype is depicted in a *box* beneath each individual. AAO age at onset, AAE age at examination. *Arrow* proband, *open square* male, *open circle* female, *filled square and circle* PD affected individual, *square and circle with a strike through* deceased individual, *open diamond* gender disguised

The second G2019S-positive family is of Russian Ashkenazi Jewish ancestry (Fig. 2, family B). The proband (individual III-2) had an AAO of 42 years and underwent a pallidotomy for tremor in Sweden at the age of 57 years. She displayed a good response to dopamine therapy, and has been troubled by only minor dyskinesias at last review at the age of 70. Other than constipation, she has had no symptoms of autonomic disturbance, REM sleep behavior disorder or loss of olfaction. This family also exhibits a possible autosomal dominant inheritance pattern with reduced penetrance. Her maternal uncle (individual II-4) had an AAO of 70 years but no further information could be obtained from him as he was not willing to take part in the study.

In addition, two Caucasian patients with sporadic PD were also found to carry the G2019S mutation. The two patients, one female (Fig. 2, family C) and one male (Fig. 2, family D) both presented with late-onset PD, with AAO of 70 and 63 years, respectively. The female patient has Polish Jewish ancestry. Her disease progression has not been severe, although she also suffers from depression and insomnia. She showed a good response to dopamine, and has had only mild motor fluctuations and mild dyskinesias when assessed 15 years after onset of her illness. The male patient has Jewish ancestry from Eastern Europe and he has had PD for approximately 15 years with a very slow disease progression. At time of last assessment he was 79 years old and

maintained a good response to dopamine, albeit with motor fluctuations. He had no autonomic impairment or cognitive impairment, but did have REM sleep behavior disorder, and severe freezing of speech. His mother had been diagnosed with motor neuron disease.

Haplotype analysis using ten markers spanning the *LRRK2* locus revealed that all four probands harbor the common G2019S haplotype referred to as haplotype 1 (Fig. 2; families A, B, C and D). Three of these probands are known to have Jewish ancestry and the fourth has German ancestry.

Discussion

In the present study we show that the frequency of the *LRRK2* G2019S mutation is relatively low (~2%) in a group of 205 South African PD patients and is absent in our Afrikaner and Black South African patients. Two of the G2019S-positive patients have onset of PD of <50 years and a possible autosomal dominant family history. The remaining two patients have sporadic late-onset PD. We also demonstrate that the HRM technique is a technically simple and cost-effective method for detecting the presence of this common PD-causing mutation.

In one of our G2019S-positive families (family A), the proband presented with both PD and psychosis, and her sister has PD and schizophrenia. It has previously been shown that the G2019S mutation can manifest in phenotypes other than PD: in one patient diagnosed with corticobasal syndrome and a family history of dementia (Chen-Plotkin et al. 2008) and another patient with Alzheimer's disease and PD (Santos-Reboucas et al. 2008). These findings suggest that the screening for G2019S and possibly other LRRK2 mutations should not be restricted to only PD patients. However, a limitation of the present study is the lack of autopsy material for the two siblings with PD and psychiatric problems, which precludes pathological examination and diagnostic confirmation.

This is the first report on the frequency of G2019S in the South African population and only the second sub-Saharan country in Africa to be studied. G2019S was not detected in any of the 16 Black patients screened and to our knowledge this mutation has never been found in a Black PD patient. Three of our four patients are of Eastern European Jewish origin, and the identification of the G2019S mutation in the South African PD population is likely to represent limited preferential immigration policies to South Africa from Eastern Europe in the late nineteenth and early twentieth century. Compared with many other countries the frequency of G2019S is

relatively low in the South African patients studied. In the present study it was detected at a frequency of 2.8% in patients with familial PD (2/72 patients) and 1.5% in apparently sporadic patients (2/133 patients). Our preliminary data therefore indicates that genetic screening for G2019S is possibly less important in the general South African setting, particularly in the Black PD population. The low frequency also indicates that there are no apparent founder effects in South African populations for this particular mutation. This is interesting, given that the South African Afrikaner population, in particular, is well known to have founder effects for a number of different inherited disorders. Furthermore, finding *LRRK2* mutations in apparently sporadic PD cases concurs with the findings of many other studies and effectively blurs the distinction between early-onset familial (classified as 'genetic') and late-onset sporadic ('non-genetic') forms of this disorder.

It was shown that the four South African G2019S-positive families share a common ancestor in that they all have the same disease-associated haplotype. This haplotype, referred to as haplotype 1, has been found in numerous families worldwide of diverse origins including Ashkenazi Jewish, North African Arabs as well as Caucasian European and North American populations. The fact that this haplotype is found in so many different populations on different continents indicates that the mutation is very old. Current age estimates date the mutation to have originated approximately 1,830–2,600 years ago (Bar-Shira et al. 2009; Warren et al. 2008). Further studies on G2019S and the disease-associated haplotypes in diverse populations are warranted in order to assess the contribution of *LRRK2* as a cause of PD worldwide. On a broader level, this collective data will assist with more accurate age estimates of the mutation and will provide insight into the migratory patterns of humans many thousands of years ago. To this end, studies to determine whether or not G2019S is present in the Black PD population are of major relevance.

Acknowledgments

We thank the study participants for their participation in the study and Melissa du Plessis for performing the DNA isolations. We also gratefully acknowledge the Western Province Blood Transfusion Service for providing the control samples. This work was financially supported by the South African Medical Research Council, the Harry and Doris Crossley Foundation, the University of Stellenbosch, South Africa, and by the Agence Nationale de la Recherche, France (ANR-05-NEUR-019).

Conflict of interest statement

The authors declare that they have no conflict of interest.

References

An,X.K., Peng,R., Li,T., Burgunder,J.M., Wu,Y., Chen,W.J., Zhang,J.H., Wang,Y.C., Xu,Y.M., Gou,Y.R., Yuan,G.G., and Zhang,Z.J. (2008). LRRK2 Gly2385Arg variant is a risk factor of Parkinson's disease among Han-Chinese from mainland China. Eur. J Neurol 15, 301-305.

Bar-Shira,A., Hutter,C.M., Giladi,N., Zabetian,C.P., and Orr-Urtreger,A. (2009). Ashkenazi Parkinson's disease patients with the LRRK2 G2019S mutation share a common founder dating from the second to fifth centuries. Neurogenetics. 10, 355-358.

Bosgraaf,L. and Van Haastert,P.J. (2003). Roc, a Ras/GTPase domain in complex proteins. Biochim. Biophys Acta 1643, 5-10.

Chen-Plotkin,A.S., Yuan,W., Anderson,C., McCarty,W.E., Hurtig,H.I., Clark,C.M., Miller,B.L., Lee,V.M., Trojanowski,J.Q., Grossman,M., and Van,D., V (2008). Corticobasal syndrome and primary progressive aphasia as manifestations of LRRK2 gene mutations. Neurology 70, 521-527.

Correia,G.L., Ferreira,J.J., Rosa,M.M., Coelho,M., Bonifati,V., and Sampaio,C. (2010). Worldwide frequency of G2019S LRRK2 mutation in Parkinson's disease: a systematic review. Parkinsonism. Relat Disord 16, 237-242.

de Rijk,M.C., Launer,L.J., Berger,K., Breteler,M.M., Dartigues,J.F., Baldereschi,M., Fratiglioni,L., Lobo,A., Martinez-Lage,J., Trenkwalder,C., and Hofman,A. (2000). Prevalence of Parkinson's disease in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. Neurology 54, S21-S23.

Di Fonzo,A., Rohe,C.F., Ferreira,J., Chien,H.F., Vacca,L., Stocchi,F., Guedes,L., Fabrizio,E., Manfredi,M., Vanacore,N., Goldwurm,S., Breedveld,G., Sampaio,C., Meco,G., Barbosa,E., Oostra,B.A., and Bonifati,V. (2005). A frequent LRRK2 gene mutation associated with autosomal dominant Parkinson's disease. Lancet 365, 412-415.

Di Fonzo,A., Wu-Chou,Y.H., Lu,C.S., van,D.M., Simons,E.J., Rohe,C.F., Chang,H.C., Chen,R.S., Weng,Y.H., Vanacore,N., Breedveld,G.J., Oostra,B.A., and Bonifati,V. (2006). A common missense variant in the LRRK2 gene, Gly2385Arg, associated with Parkinson's disease risk in Taiwan. Neurogenetics. 7, 133-138.

Elbaz,A., Bower,J.H., Maraganore,D.M., McDonnell,S.K., Peterson,B.J., Ahlskog,J.E., Schaid,D.J., and Rocca,W.A. (2002). Risk tables for parkinsonism and Parkinson's disease. J Clin Epidemiol. 55, 25-31.

Erali,M., Voelkerding,K.V., and Wittwer,C.T. (2008). High resolution melting applications for clinical laboratory medicine. Exp. Mol Pathol. 85, 50-58.

Farrer,M.J., Stone,J.T., Lin,C.H., Dachsel,J.C., Hulihan,M.M., Haugarvoll,K., Ross,O.A., and Wu,R.M. (2007). Lrrk2 G2385R is an ancestral risk factor for Parkinson's disease in Asia. Parkinsonism. Relat Disord 13, 89-92.

Fung,H.C., Chen,C.M., Hardy,J., Singleton,A.B., and Wu,Y.R. (2006). A common genetic factor for Parkinson disease in ethnic Chinese population in Taiwan. BMC. Neurol 6, 47.

Giasson,B.I. and Van Deerlin,V. (2008). Mutations in LRRK2 as a cause of Parkinson's disease. Neurosignals. 16, 99-105.

Gibb,W.R.G. and Lees,A.J. (1988). A Comparison of Clinical and Pathological Features of Young-Onset and Old-Onset Parkinsons-Disease
1. Neurology 38, 1402-1406.

Guo,L., Wang,W., and Chen,S.G. (2006). Leucine-rich repeat kinase 2: relevance to Parkinson's disease. Int. J Biochem Cell Biol 38, 1469-1475.

Healy,D.G., Falchi,M., O'Sullivan,S.S., Bonifati,V., Durr,A., Bressman,S., Brice,A., Aasly,J., Zabetian,C.P., Goldwurm,S., Ferreira,J.J., Tolosa,E., Kay,D.M., Klein,C., Williams,D.R., Marras,C., Lang,A.E., Wszolek,Z.K., Berciano,J., Schapira,A.H., Lynch,T., Bhatia,K.P., Gasser,T., Lees,A.J., and Wood,N.W. (2008). Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. Lancet Neurol 7, 583-590.

Kachergus,J., Mata,I.F., Hulihan,M., Taylor,J.P., Lincoln,S., Aasly,J., Gibson,J.M., Ross,O.A., Lynch,T., Wiley,J., Payami,H., Nutt,J., Maraganore,D.M., Czyzewski,K., Styczynska,M., Wszolek,Z.K., Farrer,M.J., and Toft,M. (2005). Identification of a novel LRRK2 mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations. Am. J Hum Genet 76, 672-680.

Lesage,S., Belarbi,S., Troiano,A., Condroyer,C., Hecham,N., Pollak,P., Lohman,E., Benhassine,T., Ysmail-Dahlouk,F., Durr,A., Tazir,M., and Brice,A. (2008). Is the common LRRK2 G2019S mutation related to dyskinesias in North African Parkinson disease? Neurology 71, 1550-1552.

Lesage,S. and Brice,A. (2009). Parkinson's disease: from monogenic forms to genetic susceptibility factors. Hum Mol Genet 18, R48-R59.

Lesage,S., Durr,A., Tazir,M., Lohmann,E., Leutenegger,A.L., Janin,S., Pollak,P., and Brice,A. (2006). LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs. N. Engl J Med 354, 422-423.

Lesage,S., Janin,S., Lohmann,E., Leutenegger,A.L., Leclere,L., Viallet,F., Pollak,P., Durif,F., Thobois,S., Layet,V., Vidailhet,M., Agid,Y., Durr,A., Brice,A., Bonnet,A.M., Borg,M., Broussolle,E., Damier,P., Destee,A., Martinez,M., Penet,C., Rasco,O., Tison,F., Tranchan,C., and Verin,M. (2007). LRRK2 exon 41 mutations in sporadic Parkinson disease in Europeans. Arch. Neurol 64, 425-430.

Lesage,S., Patin,E., Condroyer,C., Leutenegger,A.L., Lohmann,E., Giladi,N., Bar-Shira,A., Belarbi,S., Hecham,N., Pollak,P., Ouvrard-Hernandez,A.M., Bardien,S., Carr,J., Benhassine,T., Tomiyama,H., Pirkevi,C., Hamadouche,T., Cazeneuve,C., Basak,A.N., Hattori,N., Durr,A., Tazir,M., Orr-Urtreger,A., Quintana-Murci,L., and Brice,A. (2010). Parkinson's disease-related LRRK2 G2019S mutation results from independent mutational events in humans. Hum Mol Genet Epub ahead of print.

Marin,I., van Egmond,W.N., and Van Haastert,P.J. (2008). The Roco protein family: a functional perspective. FASEB J 22, 3103-3110.

Okubadejo,N., Britton,A., Crews,C., Akinyemi,R., Hardy,J., Singleton,A., and Bras,J. (2008). Analysis of Nigerians with apparently sporadic Parkinson disease for mutations in LRRK2, PRKN and ATXN3. PLoS. One. 3, e3421.

Orr-Urtreger,A., Shifrin,C., Rozovski,U., Rosner,S., Bercovich,D., Gurevich,T., Yagev-More,H., Bar-Shira,A., and Giladi,N. (2007). The LRRK2 G2019S mutation in Ashkenazi Jews with Parkinson disease: is there a gender effect? Neurology 69, 1595-1602.

Ozelius,L.J., Senthil,G., Saunders-Pullman,R., Ohmann,E., Deligtisch,A., Tagliati,M., Hunt,A.L., Klein,C., Henick,B., Hailpern,S.M., Lipton,R.B., Soto-Valencia,J., Risch,N., and Bressman,S.B. (2006). LRRK2 G2019S as a cause of Parkinson's disease in Ashkenazi Jews. N. Engl J Med 354, 424-425.

Patterson,N., Petersen,D.C., van der Ross,R.E., Sudoyo,H., Glashoff,R.H., Marzuki,S., Reich,D., and Hayes,V.M. (2010). Genetic structure of a unique admixed population: implications for medical research. Hum Mol Genet 19, 411-419.

Pirkevi,C., Lesage,S., Condroyer,C., Tomiyama,H., Hattori,N., Ertan,S., Brice,A., and Basak,A.N. (2009). A LRRK2 G2019S mutation carrier from Turkey shares the Japanese haplotype. Neurogenetics. 10, 271-273.

Ross,O.A., Wu,Y.R., Lee,M.C., Funayama,M., Chen,M.L., Soto,A.I., Mata,I.F., Lee-Chen,G.J., Chen,C.M., Tang,M., Zhao,Y., Hattori,N., Farrer,M.J., Tan,E.K., and Wu,R.M. (2008). Analysis of Lrrk2 R1628P as a risk factor for Parkinson's disease. Ann Neurol 64, 88-92.

Santos-Reboucas,C.B., Abdalla,C.B., Baldi,F.J., Martins,P.A., Correa,J.C., Goncalves,A.P., Cunha,M.S., Borges,M.B., Pereira,J.S., Laks,J., and Pimentel,M.M. (2008). Co-occurrence of sporadic parkinsonism and late-onset Alzheimer's disease in a Brazilian male with the LRRK2 p.G2019S mutation. Genet Test. 12, 471-473.

Tan,E.K., Tan,L.C., Lim,H.Q., Li,R., Tang,M., Yih,Y., Pavanni,R., Prakash,K.M., Fook-Chong,S., and Zhao,Y. (2008). LRRK2 R1628P increases risk of Parkinson's disease: replication evidence. Hum Genet 124, 287-288.

Tan,E.K., Zhao,Y., Skipper,L., Tan,M.G., Di,F.A., Sun,L., Fook-Chong,S., Tang,S., Chua,E., Yuen,Y., Tan,L., Pavanni,R., Wong,M.C., Kolatkar,P., Lu,C.S., Bonifati,V., and Liu,J.J. (2007). The LRRK2 Gly2385Arg variant is associated with Parkinson's disease: genetic and functional evidence. Hum Genet 120, 857-863.

Warren,L., Gibson,R., Ishihara,L., Elango,R., Xue,Z., Akkari,A., Ragone,L., Pahwa,R., Jankovic,J., Nance,M., Freeman,A., Watts,R.L., and Hentati,F. (2008). A founding LRRK2 haplotype shared by Tunisian, US, European and Middle Eastern families with Parkinson's disease. Parkinsonism. Relat Disord 14, 77-80.

West,A.B., Moore,D.J., Biskup,S., Bugayenko,A., Smith,W.W., Ross,C.A., Dawson,V.L., and Dawson,T.M. (2005). Parkinson's disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. Proc. Natl. Acad. Sci U. S. A 102, 16842-16847.

Wittwer,C.T., Reed,G.H., Gundry,C.N., Vandersteen,J.G., and Pryor,R.J. (2003). High-resolution genotyping by amplicon melting analysis using LCGreen. Clin Chem. 49, 853-860.

Zabetian,C.P., Hutter,C.M., Yearout,D., Lopez,A.N., Factor,S.A., Griffith,A., Leis,B.C., Bird,T.D., Nutt,J.G., Higgins,D.S., Roberts,J.W., Kay,D.M., Edwards,K.L., Samii,A., and Payami,H. (2006a). LRRK2 G2019S in families with Parkinson disease who originated from Europe and the Middle East: evidence of two distinct founding events beginning two millennia ago. Am. J Hum Genet 79, 752-758.

Zabetian,C.P., Morino,H., Ujike,H., Yamamoto,M., Oda,M., Maruyama,H., Izumi,Y., Kaji,R., Griffith,A., Leis,B.C., Roberts,J.W., Yearout,D., Samii,A., and Kawakami,H. (2006b). Identification and haplotype analysis of LRRK2 G2019S in Japanese patients with Parkinson disease. Neurology 67, 697-699.

Zhang,Z., Burgunder,J.M., An,X., Wu,Y., Chen,W., Zhang,J., Wang,Y., Xu,Y., Gou,Y., Yuan,G., Mao,X., and Peng,R. (2009). LRRK2 R1628P variant is a risk factor of Parkinson's disease among Han-Chinese from mainland China. Mov Disord 24, 1902-1905.

Chapter 5

Assessing the prevalence of PINK1 genetic variants in South African patients diagnosed with early- and late-onset Parkinson's disease

Rowena J. Keyser[a], Suzanne Lesage[b, c, d], Alexis Brice[b, c, d], Jonathan Carr[e] and Soraya Bardien[a]

[a]Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Cape Town, South Africa

[b]Université Pierre et Marie Curie-Paris6, Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, UMR-S975, Paris, France

[c]Inserm, U975, Paris, France

[d]CNRS, UMR 7225, Paris, France

[e]Division of Neurology, Faculty of Health Sciences, University of Stellenbosch, Cape Town, South Africa

Published in: Biochemical and Biophysical Research Communications. 2010 Jul 16;398(1):125-129

Impact factor of journal: 2.548

My contribution to this project:

- Planning of study
- PCR and HRM analysis
- Sequencing
- Analysis and interpretation of data
- Writing of manuscript

Abstract

Mutations in the *PINK1* gene are the second most common cause after parkin of autosomal recessive early-onset Parkinson's disease (PD). PINK1 is a protein kinase that is localized to the mitochondrion and is ubiquitously expressed in the human brain. Recent studies aimed at elucidating the function of PINK1, have found that it has neuroprotective properties against mitochondrial dysfunction and proteasomally-induced apoptosis. In the present study, we aimed to investigate the prevalence of *PINK1* genetic variants in 154 South African PD patients from all ethnic groups. Mutation screening was performed using the High-Resolution Melt technique and direct sequencing. A total of 16 sequence variants were identified: one known homozygous mutation (Y258X), two heterozygous missense variants (P305A and E476K), and 13 polymorphisms of which five were novel. No homozygous exonic deletions were detected. The novel P305A variant was found in a female patient of Black Xhosa ethnicity who has a positive family history of the disease and an age at onset of 30 years. This variant has the potential to modulate enzymatic activity due to its location in the kinase domain. This is the first report on mutation screening of *PINK1* in the South African population. Results from the present study showed that point mutations and homozygous exonic deletions in *PINK1* are not a common cause of PD in the South African population.

Introduction

The *PINK1* (PTEN-induced putative kinase 1) gene (*PARK6*; OMIM 608309) encodes a 581 amino acid protein kinase that is localized to the mitochondrion and is ubiquitously expressed in the human brain [1,2]. It contains an N-terminal mitochondrial-targeting motif and a highly conserved serine/threonine kinase domain. Mutations in *PINK1* are the second most common cause after parkin of autosomal recessive early-onset Parkinson's disease (PD), a common and incurable progressive neurodegenerative disorder that is characterized by motor impairments involving resting tremors, bradykinesia, postural instability and rigidity. Atypical PD clinical features, such as psychiatric disturbances, dystonia at onset and sleep benefit, have been observed in PD patients with *PINK1* mutations [3,4].

The involvement of *PINK1* in PD was first discovered in 2004 with the identification of homozygous point mutations in families of Italian and Spanish origin [1]. Since then, missense mutations, genomic rearrangements and truncating mutations have been identified in diverse populations. It has been shown that the frequency of *PINK1* mutations ranges from 1% to 8% in patients of different ethnicities [5,6,7]. The discovery of the involvement of *PINK1* in PD was the first evidence that a kinase signaling pathway may be important in the pathogenesis of dopaminergic nigral cell death. Also, it provided a molecular link between the mitochondria and neurodegeneration in PD [1,8]. Mitochondrial dysfunction is thought to play an integral role in the pathogenesis of PD [9].

PINK1-deficient Drosophila has been shown to display mitochondrial defects leading to degeneration of flight muscles and loss of dopaminergic neurons [10,11,12]. PINK1 is thought to have neuroprotective properties since overexpression in neuroblastoma cells is associated with a decrease in the susceptibility of the cells to neurotoxin-induced cell death [1,13]. The mechanism of PINK1's suggested ability to prevent cell death is by the inhibition of release of cytochrome C from mitochondria and maintenance of the mitochondrial membrane potential [1,13,14]. PINK1 phosphorylates mitochondrial proteins in response to cellular stress and in so doing protects against mitochondrial dysfunction [1]. It has been reported that PINK1 phosphorylates TNF receptor-associated protein 1, which causes the cells to be protected against oxidative stress-induced apoptosis [15]. Parkin and PINK1 are thought to function in the same pathway in maintaining mitochondrial integrity and function, with PINK1 functioning upstream from parkin [11,16].

Currently, the majority of *PINK1* mutations are distributed throughout the serine/threonine kinase domain. Some of these mutations have been reported to cause reduction in enzymatic activity [17,18] and could affect substrate recognition as well. The mutations that are located outside the domain might affect protein stability, which could indirectly influence kinase activity. In the present study, we conducted genetic mutation screening, in order to determine the prevalence of *PINK1* genetic variants in South African patients diagnosed with early- and late-onset PD.

Materials and methods

Study participants

The study was approved by the Committee for Human Research at the University of Stellenbosch, South Africa (Protocol number 2002/C059). A total of 154 unrelated PD patients from all South African ethnic groups were recruited with informed consent. They were recruited from the Movement Disorders clinic at Tygerberg Hospital, Cape Town, as well as from the Parkinson's Association of South Africa. The patients were examined by a movement disorder specialist (JC) and met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD [19]. Inclusion criteria of patients in this study were early-onset and/or a positive family history of PD.

The ethnic breakdown was as follows: 35.3% Caucasian, 30.9% South African Afrikaner, 19.8% mixed ancestry (defined in Bardien et al., 2009 [20]), 11.8% Black, and 2.2% Indian. The average age at onset (AAO) of the study group was 52 years ± 12.83. Among the patients, 65 (48%) were ≤50 years old at onset of the disease (age range 17–50 years) and 71 (52%) were >50 years old at onset (age range 51–77). The percentage of males in our study group was 62%. The percentage of patients with positive family history was 36% and both autosomal recessive and autosomal dominant patterns of inheritance were present. Seventy-two mixed ancestry control samples (male = 72%) and 54 black control samples (male = 57%) were recruited from unrelated blood donors at the South African Western Province Blood Transfusion Service blood collection clinics. The controls had been 'de-identified' and had not been clinically assessed for signs of PD.

Genetic analysis

For each study participant, a blood sample was collected for genetic analysis and genomic DNA was extracted using established methods. Polymerase chain reaction (PCR) primers were

designed for each of the eight exons of the *PINK1* gene using Primer3 software. Primer sequences are available from the authors on request. PCR reactions were performed in a 2720 Thermal Cycler (Applied Biosystems, USA). The 25 μl reactions contained 10 ng template DNA, 10 pmol of each primer, 75 μM dNTPs (Promega, USA), 1.5 mM $MgCl_2$, 1× NH_4 buffer (Bioline, UK), 5% DMSO (for selected primer sets), 2 μM SYTO9 fluorescent dye (Invitrogen, USA) and 0.25 U of BIOTAQ DNA polymerase (Bioline, UK). The PCR cycling parameters were as follows: initial denaturation at 94 °C for 5 min, followed by 30 cycles of denaturation at 94 °C for 30 sec, annealing at different temperatures (ranging from 55 to 61 °C) according to each primer set for 30 s, extension at 72 °C for 45 s, and a final extension step at 72 °C for 7 min.

Thereafter, the PCR products were subjected to high-resolution melt (HRM) analysis. For this, the samples underwent a melt from 75 to 95 °C rising by 0.1 °C each step on a Rotor-Gene 6000 analyzer (Corbett Life Sciences, Australia). In HRM, double stranded DNA dissociates into single stranded DNA as the temperature increases up to 95 °C, during which the shift in fluorescence is monitored. The thermal denaturation profile that is produced is characteristic of a specific PCR product and is dependent on its sequence length, base and GC content [21,22]. The samples that showed different HRM profiles when compared to profiles produced by the wild-type sample were sequenced in order to characterize the genetic variant. Direct sequencing was performed using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems) and analyzed on a 3130xl Genetic Analyzer (Applied Biosystems). The primers used for sequencing were the same as those used to generate the PCR products for HRM analysis. For eighteen of the samples, mutation screening was performed using direct sequencing of all eight exons and not the HRM method. The frequency of selected sequence variants was determined in ethnically matched control samples.

Results

The genomic DNA of 154 South African PD patients were screened for mutations in *PINK1* using the HRM method and direct sequencing. No homozygous exonic deletions were detected. A total of 16 sequence variants were identified (Table 1A and B) of which one (Y258X) is a known mutation that had previously been identified using the Multiplex Ligation-dependent Probe Amplification method (MLPA) [23]. The homozygous Y258X mutation (Fig. 1A) was found in one patient of Indian ethnicity and was also present in both of her affected siblings. Although her duration of illness at the time of assessment was 30 years, her symptoms were

relatively well controlled on only 500 mg of levodopa daily. She experienced occasional freezing, and had mild dyskinesias affecting the legs. There were no features of autonomic dysfunction, but she did self-report impairment of memory for recent events. Detection of Y258X by HRM in the present study showed the efficacy of this mutation screening method in detecting sequence variants.

Two other sequence variants (P305A, E476K) were also identified as well as 13 polymorphisms. The novel P305A variant (Fig. 1B) was identified in a heterozygous state in a female patient of Black Xhosa ethnicity. This patient had a positive family history of the disease and an AAO of 30 years. However, the variant was not present in the proband's affected sibling who had an AAO of 45 years. P305A is located within an evolutionarily conserved region of the gene (Fig. 2). PolyPhen analysis (http://genetics.bwh.harvard.edu/pph/) revealed that this variant is 'probably damaging' although it was found in 1.9% of control chromosomes.

The E476K variant (Fig. 1C) was identified in a heterozygous state in a female patient of mixed ancestry who had an AAO of 48 years and no family history of the disease. This variant has been previously reported in a number of studies [3,5,6,8]. PolyPhen analysis revealed that E476K is 'benign' and in the present study it was found in 1.4% of control chromosomes. This variant is not located in an evolutionarily conserved region of PINK1 (data not shown). All three of these variants occur in the functional serine/threonine kinase domain of PINK1. The P305A variant occurs within the third kinase insert domain [24], which is suggested to contain regulatory motifs important for kinase activity and substrate selectivity [25,26]. Y258X and E476K are located in the second kinase insert domain and eleventh kinase subdomain, respectively [24].

PINK1 was found to be polymorphic in the South African population although very few pathogenic mutations were found. Out of the 13 polymorphisms detected, 5 were novel variants and included 5′UTR-35 C > T, IVS4 + 56 G > T, IVS7-8 T > G, 3′UTR + 37 A > T, and 3′UTR + 43 G > A. The IVS1-7 A > G polymorphism had the highest frequency (22%) in the study group compared to the other variants.

Table 1A Known and putative mutations identified in *PINK1* in South African PD patients

Exon	Variant	Zygosity	Number of patients/Ethnicity[#]	AAO (yrs)	Family history	Protein domain	Frequency in control chromosomes	Previously reported/ Reference
3	Y258X (TAC-TAA)	Homozygous	1/Indian (68.10)	37	+	Kinase domain	0/130	[23,37]
4	P305A (CCT-GCT)	Heterozygous	1/Black (42.03)	30	+	Kinase domain	2/108 (1.9%)	Novel
7	E476K (GAG-AAG)	Heterozygous	1/Mixed ancestry (63.69)	48	-	Kinase domain	2/144 (1.4%)	[3,5,6,8]

[#] = patient sample number shown in brackets; AAO = age at onset

Table 1B Polymorphisms identified in *PINK1* in South African PD patients

Location	Variant	Ethnicity of patients with variant	Frequency in patients (n=154)
5'-UTR	-35 C>T	Afrikaner	1
Exon 1	L63L	Afrikaner	6
Exon 1	Q115L	Afrikaner	1
Intron 1	IVS1-7 A>G	Afrikaner, Caucasian, Mixed ancestry, Indian and Black	30
Intron 1	IVS1-65 C>G	Afrikaner	1
Intron 4	IVS4+56 G>T	Black	1
Intron 4	IVS4+72 G>C	Indian	1
Exon 5	A340T	Afrikaner, Caucasian, Mixed ancestry, Indian and Black	10
Intron 6	IVS6+43 C>T	Afrikaner and Caucasian	4
Intron 7	IVS7-8 T>G	Afrikaner, Caucasian, Mixed ancestry, Indian and Black	18
Exon 8	N521T	Afrikaner, Caucasian, Mixed ancestry, Indian and Black	10
3'-UTR	*37 A>T	Afrikaner, Caucasian and Mixed ancestry, Indian and Black	24
3'-UTR	*43 G>A	Mixed ancestry and Indian	2

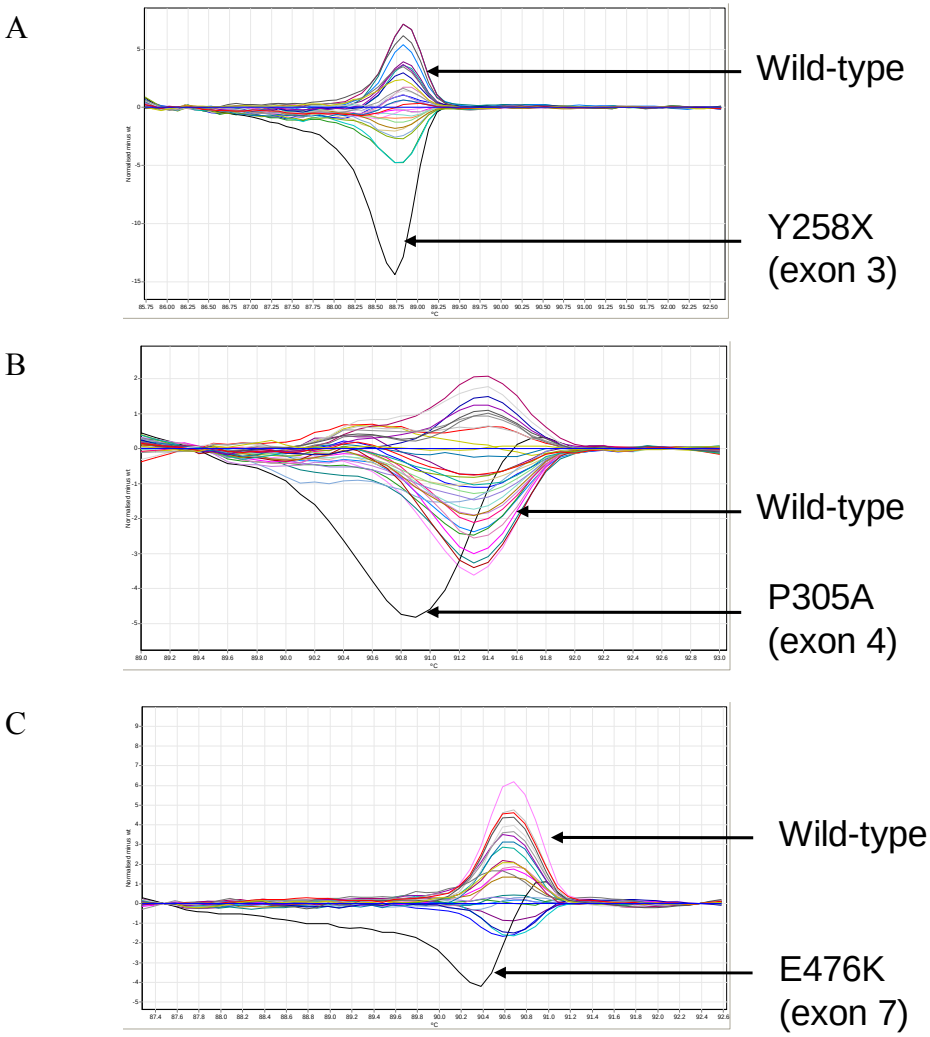


Fig. 1 Difference graphs produced by the high-resolution melt technique illustrating that sequence variants can be distinguished from wild-type alleles. Melt curves of the A) previously reported Y258X mutation, B) novel P305A variant, C) previously reported E476K variant.

```
Human       YPDVLPSRLHPEGLGHGRTLFLVMKNYPCTLRQYLCVNTPSPRLAAMMLLQLLEGVDHLV
Chimp       YPDVLPSRLHPEGLGHGRTLFLVMKNYPCTLRQYLCVNTPSPRLAAMMLLQLLEGVDHLV
Mouse       YPDMLPPHYYPEGLGHGRTLFLVMKNYPCTLRQYLEEQTPSSRLATMMTLQLLEGVDHLV
Rat         YPDMLPPHYYPEGLGHGRTLFLVMKNYPCTLRQYLEEQTPSSRLATMMTLQLLEGVDHLV
Cow         YPDVLPPRLHPAGLGHGRTLFLVMKNYPCTLRQYLRGNTPSPRLATVMTLQLLEGVDHLV
Chicken     YPDVLPVSLNPRGIGRSHTLFLVMKNYPCTLCQYLRDNSPDSRLSTMMILQLLEGVDHLV
Zebrafish   YPDVLPTRLNPHGLGSNRTLFLVMKNYPCTLRQYLEVCVPKRTQASLMFLQLLEGVDHLC
            ***:**     * *:* .:************ ***     *.   :::* **********
```

Fig. 2 Sequence alignment (using ClustalW) of PINK1 amino acid sequences of human (NP_115785.1), chimp (ENSPTRP00000000500), mouse (NP_081156.2), rat (XP_216565.2), cow (NP_001093171.1), chicken (XP_423139.2), and zebrafish (NP_001008628.1).The box indicates the position of the novel P305A variant. Identical amino acids are indicated by *asterisks*

Discussion

We report the results of genetic mutation screening of the *PINK1* gene which was conducted on PD patients from all South African ethnic groups. The study participants had previously been screened for mutations in the *parkin* gene [20] and the G2019S mutation in the *LRRK2* gene [27]. For *parkin*, only two patients with mutations were found (both homozygous whole exon deletions) and only four individuals were found to harbor the G2019S mutation. These results indicate that *parkin* and *LRRK2* are not a common cause of PD in South African patients and the *PINK1* gene was therefore investigated.

In the present study one previously identified mutation (Y258X) and two missense variants (P305A and E476K) were identified as well as 13 polymorphisms. The novel P305A variant was found in one PD patient of Black Xhosa ethnicity who had a positive family history of the disease. The clinical features of PD in Black individuals have been reported to be atypical and that it is a late-onset akinetic-rigid syndrome [28]. However, the individual harboring the P305A variant has typical clinical features of PD but she had an earlier onset and a more severe phenotype than her affected sibling. The P305A variant, due to its location in the serine/threonine kinase domain, has the potential to influence the function of PINK1 by affecting its autophosphorylation ability [29], and it is predicted to be 'probably damaging' by PolyPhen analysis. However, the fact that P305A was not present in the proband's affected sibling as well as the fact that it was found in 1.9% of ethnic matched control chromosomes suggests that it is a non-pathogenic polymorphism.

The E476K variant has previously been found in heterozygous form in four unrelated PD patients [3,5,6,30] but has also been found at very low frequencies in control subjects [5,8]. This variant has been shown to impair mitochondrial membrane potential after cellular stress induced by proteasomal inhibition [8]; however, its pathogenic status is currently equivocal. In the present study E476K was observed at a frequency of 1.4% in control chromosomes, which together with the fact that this residue is evolutionarily poorly-conserved and is also predicted by PolyPhen analysis to be 'benign', indicates that it may be a non-pathogenic polymorphism.

HRM detected 13 polymorphisms in the present study. The previously reported IVS1-7 A > G variant had the highest prevalence in the study group compared to the other polymorphisms. This variant might have an effect on splicing of the gene; however, further studies are necessary to investigate this. The presence of the IVS1-7 A > G variant has been shown to decrease the AAO

in patients exposed to various environmental risk factors [31]. Since we do not have comprehensive data on environmental exposures for all our study participants, this finding could not be verified in the present study. However, preliminary analysis showed that the average AAO for patients with (44.1 ± 9.9 years; 30/154) and without (52.7 ± 12.5 years; 124/154) the IVS1-7 A > G variant was significantly different ($p < 0.001$; t-test).

A few recent studies have reported either no *PINK1* mutations in PD patients from Brazil [32], China [33] and Portugal [34] or only a few heterozygous variants of unknown pathogenic significance in patients from Australia [35] and India [36]. These findings are in contrast to the high frequencies observed in earlier studies. This may reflect that the contribution of this gene varies according to ethnicity or that the frequency of *PINK1* mutations is not as common as was previously thought. Therefore, innovative bioinformatic *in silico* and wet-bench experimental strategies are urgently needed to identify novel disease-causing genes for PD.

Conclusion

Results from the present study showed that point mutations, small insertions or deletions, and homozygous exonic deletions in *PINK1* are not a common cause of PD (<1%) in the South African population. Therefore, it is proposed that as yet unknown genes are responsible for PD in this population. Further studies are necessary to determine whether the P305A variant could possibly alter or modulate the enzymatic activity of PINK1.

Conflict of interest

None declared.

Acknowledgments

We thank the study participants, Ms. D. Lombard and the South African Western Province Blood Transfusion Service for collection of blood samples. Dr. Craig Kinnear and Dr. Sian Hemmings are thanked for providing assistance with statistical analysis. This work was supported by the South African Medical Research Council, the Harry and Doris Crossley Foundation, the University of Stellenbosch, South Africa, and the Agence Nationale de la Recherche, France (ANR-05-NEUR-019).

References

1. E.M. Valente, P.M. Abou-Sleiman, V. Caputo, et al., Hereditary early-onset Parkinson's disease caused by mutations in PINK1, Science. 304(2004) 1158-1160.

2. S. Gandhi, M.M. Muqit, L. Stanyer, et al., PINK1 protein in normal human brain and Parkinson's disease, Brain. 129(2006) 1720-1731.

3. E.M. Valente, S. Salvi, T. Ialongo, et al., PINK1 mutations are associated with sporadic early-onset parkinsonism, Ann. Neurol. 56(2004) 336-341.

4. L. Ephraty, O. Porat, D. Israeli, et al., Neuropsychiatric and cognitive features in autosomal-recessive early parkinsonism due to PINK1 mutations, Mov. Disord. 22(2007) 566-569.

5. E. Rogaeva, J. Johnson, A.E. Lang, et al., Analysis of the PINK1 gene in a large cohort of cases with Parkinson disease, Arch. Neurol. 61(2004) 1898-1904.

6. V. Bonifati, C.F. Rohé, G.J. Breedveld, et al., Early-onset parkinsonism associated with PINK1 mutations: frequency, genotypes, and phenotypes, Neurology. 65(2005) 87-95.

7. C. Klein, A. Djarmati, K. Hedrich, et al., PINK1, Parkin, and DJ-1 mutations in Italian patients with early-onset parkinsonism, Eur. J. Hum. Genet. 13(2005) 1086-1093.

8. P.M. Abou-Sleiman, M.M. Muqit, N.Q. McDonald, et al., A heterozygous effect for PINK1 mutations in Parkinson's disease?, Ann. Neurol. 60(2006) 414-419.

9. I.G. Onyango, Mitochondrial dysfunction and oxidative stress in Parkinson's disease, Neurochem. Res. 2008 33(2008) 589-597.

10. I.E. Clark, M.W. Dodson, C. Jiang, et al., Drosophila pink1 is required for mitochondrial function and interacts genetically with parkin, Nature. 441(2006) 1162-1166.

11. J. Park, S.B. Lee, S. Lee, Y. Kim, S. Song, S. Kim, et al., (2006) Mitochondrial dysfunction in Drosophila PINK1 mutants is complemented by parkin, Nature. 441(2006) 1157-1161.

12. Y. Yang, S. Gehrke, Y. Imai, Z. Huang, et al., Mitochondrial pathology and muscle and dopaminergic neuron degeneration caused by inactivation of Drosophila Pink1 is rescued by Parkin, Proc. Natl. Acad. Sci. U S A. 103(2006) 10793-1078.

13. A. Petit, T. Kawarai, E. Paitel, et al., Wild-type PINK1 prevents basal and induced neuronal apoptosis, a protective effect abrogated by Parkinson disease-related mutations, J. Biol. Chem. 280(2005) 34025-34032.

14. M.R. Duchen, A. Surin, J. Jacobson, Imaging mitochondrial function in intact cells, Methods. Enzymol. 361(2003) 353-389.

15. J.W. Pridgeon, J.A. Olzmann, L.S. Chin, L. Li, PINK1 protects against oxidative stress by phosphorylating mitochondrial chaperone TRAP1, PLoS. Biol.5(2007):e172.

16. J.W. Um, C. Stichel-Gunkel, H. Lübbert, et al., Molecular interaction between parkin and PINK1 in mammalian neuronal cells, Mol. Cell. Neurosci. 40(2009) 421-432.

17. A. Beilina, M. Van Der Brug, R. Ahmad, et al., Mutations in PTEN-induced putative kinase 1 associated with recessive parkinsonism have differential effects on protein stability, Proc. Natl. Acad. Sci. U S A. 102(2005) 5703-5708.

18. C.H. Sim, D.S. Lio, S.S. Mok, et al., C-terminal truncation and Parkinson's disease-associated mutations down-regulate the protein serine/threonine kinase activity of PTEN-induced kinase-1, Hum. Mol. Genet. 15(2006) 3251-3262.

19. W.R.G. Gibb, A.J. Lees, A comparison of clinical and pathological features of young-onset and old-onset Parkinson's disease, Neurology. 38(1988) 1402-1406.

20. S. Bardien, R.J. Keyser, Y. Yako, D. Lombard, J. Carr, Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease, Parkinsonism. Relat. Disord. 15(2009) 116-121.

21. C.T. Wittwer, G.H. Reed, C.N. Gundry, High-resolution genotyping by amplicon melting analysis using LCGreen, Clin. Chem. 49(2003) 853-860.

22. M. Erali, K.V. Voelkerding, C.T. Wittwer, 2008 High resolution melting applications for clinical laboratory medicine, Exp. Mol. Pathol. 85(2008) 50-58.

23. R.J. Keyser, D. Lombard, R. Veikondis, et al., Analysis of exon dosage using MLPA in South African Parkinson's disease patients, Neurogenetics. 2009 Dec 15. [Epub ahead of print]

24. R.D. Mills, C.H. Sim, S.S. Mok, et al., Biochemical aspects of the neuroprotective mechanism of PTEN-induced kinase-1 (PINK1), J. Neurochem. 105(2008) 18-33. Review.

25. H. Tokumitsu, N. Takahashi, K. Eto, et al., Substrate recognition by Ca2+/Calmodulin-dependent protein kinase kinase. Role of the arg-pro-rich insert domain, J. Biol. Chem. 274(1999) 15803-15810.

26. M. Rafie-Kolpin, P.J. Chefalo, Z. Hussain, et al., Two heme-binding domains of heme-regulated eukaryotic initiation factor-2alpha kinase. N terminus and kinase insertion, J. Biol. Chem. 275(2000) 5171-5178.

27. S. Bardien, A. Marsberg, R. Keyser, et al., *LRRK2* G2019S mutation: Frequency and haplotype data in South African Parkinson's disease patients, J. Neural. Transm. (in press)

28. K.R. Chaudhuri, M.T. Hu, D.J Brooks, Atypical parkinsonism in Afro-Caribbean and Indian origin immigrants to the UK, Mov. Disord. 15(2000) 18-23.

29. L. Silvestri, V. Caputo, E. Bellacchio, et al., Mitochondrial import and enzymatic activity of PINK1 mutants associated to recessive parkinsonism, Hum. Mol. Genet. 14(2005) 3477-3492.

30. R. Marongiu, A. Ferraris, T. Ialongo, et al., PINK1 heterozygous rare variants: prevalence, significance and phenotypic spectrum, Hum. Mutat. 29(2008) 565-576.

31. C. Godeiro Jr, P.M. Aguiar, A.C. Felício, et al., PINK1 polymorphism IVS1-7 A-->G, exposure to environmental risk factors and anticipation of disease onset in Brazilian patients with early-onset Parkinson's Disease, Neurosci. Lett. 469(2010) 155-158.

32. C. Godeiro-Junior, P.M. de Carvalho-Aguiar, A.C. Felício, et al,. PINK1 mutations in a Brazilian cohort of early-onset Parkinson's disease patients, Mov. Disord. 24(2009) 1693-1696.

33. B.R. Zhang, Z.X. Hu, X.Z. Yin, et al., Mutation analysis of parkin and PINK1 genes in early-onset Parkinson's disease in China, Neurosci. Lett. 477(2010) 19-22.

34. J. Bras, R. Guerreiro, M. Ribeiro, et al., Analysis of Parkinson disease patients from Portugal for mutations in SNCA, PRKN, PINK1 and LRRK2, BMC. Neurol. 8(2008) 1.

35. G.D. Mellick, G.A. Siebert, M. Funayama, et al., Screening PARK genes for mutations in early-onset Parkinson's disease patients from Queensland, Australia, Parkinsonism. Relat. Disord. 15(2009) 105-109.

36. A. Biswas, T. Sadhukhan, S. Majumder, et al., Evaluation of PINK1 variants in Indian Parkinson's disease patients, Parkinsonism. Relat. Disord. 16(2010) 167-171.

37. E.K. Tan, K. Yew, E. Chua, et al., PINK1 mutations in sporadic early-onset Parkinson's disease, Mov. Disord. 21(2006) 789-793.

Chapter 6

Identification of Parkinson's disease candidate genes using CAESAR and screening in South African Parkinson's disease patients

Rowena J. Keyser, Ekow Oppon, Jonathan A. Carr, Soraya Bardien

R.J. Keyser, S. Bardien
Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Cape Town, South Africa

E. Oppon
The South African National Bioinformatics Institute (SANBI), University of the Western Cape, Cape Town, South Africa

J.A. Carr
Division of Neurology, Faculty of Health Sciences, University of Stellenbosch, Cape Town, South Africa

Accepted for publication in: Journal of Neural Transmission (25 January 2011)

Impact factor of journal: 2.259

My contribution to this project:
- Planning of study
- Assisted with bioinformatic analysis
- Primer design, PCR, HRM analysis and sequencing
- Analysis and interpretation of data
- Writing of manuscript

Abstract

Assuming that a significant cause of Parkinson's disease (PD) is genetic, genetic factors have been shown to account for < 10% of all PD cases to date, and it is therefore necessary to identify novel genes. The aim of the present study was to identify PD candidate genes using a bioinformatics approach and to screen them for possible PD-causing mutations. The CAESAR (CAndidatE Search And Rank) program was used in the present study to identify and prioritize PD candidate genes. CAESAR ranks annotated human genes as candidates by using ontologies to semantically map natural language descriptions of the trait under investigation to gene-centric databases. Two of the candidates were selected and screened for mutations in 202 South African PD patients using the High-Resolution Melt (HRM) method. Samples exhibiting altered HRM profiles were sequenced. CAESAR generated a prioritized list of candidates including both known and novel PD genes. The *MAPT* and *SNCAIP* genes were selected for mutation screening from the list of ten highest scoring genes. Two novel missense (A91V and V635I), four synonymous and three intronic sequence variants were identified in *MAPT*. For *SNCAIP*, three novel missense (T383N, R606Q, N906H), one known (E709Q), four synonymous and one intronic sequence variant were found. A bioinformatics approach was used to aid in the identification and selection of PD candidate genes in a group of South African patients. Mutation screening of *MAPT* and *SNCAIP* identified novel sequence variants in both genes and further studies are necessary to determine their possible functional consequences.

Keywords Parkinson's disease; CAESAR; bioinformatic tools; MAPT, SNCAIP

Introduction

Parkinson's disease (PD) (OMIM#168600) is a neurodegenerative disorder which presents with tremor and bradykinesia, in addition to a wide range of non-motor disturbances. To date, six genes (*PARK2, PINK1, DJ-1, SNCA, LRRK2* and *ATP13A2*) have been conclusively implicated in the pathogenesis of PD (Lesage and Brice 2009), although mutations in these genes account for < 10% of all PD cases. Our studies have shown that the main PD causing genes do not play a significant role in the South African population (Bardien et al. 2009; Bardien et al. 2010; Keyser et al. 2010a; Keyser et al. 2010b and unpublished data), and therefore it is important to identify novel PD candidate genes. Identification of novel PD genes may also reveal new molecular pathways that could be involved in neurodegeneration and development of the disorder.

Traditional linkage analysis and genome-wide association studies (GWAS) have been used to identify candidate genes for PD (Maraganore et al. 2005; Fung et al. 2006; Simon-Sanchez et al. 2009; Funayama et al. 2002). A limited number of candidate genes have been found using these methods. Also, some of the findings have not been reproduced in follow-up studies, which may be due to small sample size, inadequate correction for multiple testing and population stratification. Searching through online databases to select candidate genes has become extremely difficult and time consuming because the wealth of molecular information in databases has increased exponentially. Bioinformatic tools such as GeneSeeker (van Driel et al. 2003), TOM (Rossi et al. 2006), Gene Prospector (Yu et al. 2008) and CAESAR (Gaulton et al. 2007) have been developed to assist with this challenging task. These *in-silico* methods are rapid, cost effective and are not prone to many of the biases invariably introduced by researchers.

GeneSeeker searches for relevant genes within a specific chromosomal region as defined by the user (van Driel et al. 2003). The program gathers and combines positional, expression and phenotypic data in an automated way from nine different web-based databases in order to obtain an overview of the candidate genes in the specified chromosomal region. Transcriptomics of OMIM (TOM) requires the user to have knowledge of another disease associated gene as well as the linkage area, or knowledge of two disease loci (Rossi et al. 2006). TOM's algorithm uses information on mapping, expression and function stored in public databases. The Gene Prospector tool selects and prioritizes potential disease-related genes by using HuGe Navigator, a curated published literature database of genetic association studies (Yu et al. 2008). It ranks genes according to the amount of literature available in human genome epidemiology as well as published research on both rat and mouse animal models.

In the present study, CAESAR (CAndidatE Search And Rank) (Gaulton et al. 2007) was used because it was designed specifically for complex diseases such as PD in which the relevant biological processes are not well understood and for which a large number of reasonable candidate genes exist. CAESAR is a computational system that uses ontologies to semantically map trait information to gene and protein-centric information including protein-protein interactions, metabolic pathways and tissue-specific gene expression data from various public data sources such as SwissProt, KEGG and Ensembl. The data mining results are then integrated and a ranked list of candidate genes produced that may be involved with the complex trait. The CAESAR approach mimics the individual steps a researcher would undertake to select candidate genes, but it conducts this faster and in a more quantitative and systematic manner.

In the present study, a bioinformatics strategy was used to identify candidate genes for PD. Two candidates were selected from this strategy and mutation screening was performed to identify possible disease-causing mutations in South African PD patients.

Materials and methods

Study participants

The Committee for Human Research at the University of Stellenbosch, South Africa, approved the study (Protocol number 2002/C059) and all study participants were recruited with informed written consent. The study was conducted on 202 unrelated PD patients recruited from the Movement Disorders clinic at Tygerberg Hospital, Cape Town as well as from the Parkinson's Association of South Africa, and the inclusion criteria included early onset and/or a positive family history of the condition. A movement disorder specialist (JC) examined the patients and they met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD (Gibb and Lees 1988).

The study group had an average age at onset (AAO) of 54 ± 12.34 years and 65% were male. Among the patients, 74 (37%) were ≤50 years old at onset of the disease (age range 17-50 years) and 128 (63%) were >50 years old at onset (age range 51-77). The patients were from all South African sub-population groups, these being distributed as follows: 73.2% Caucasian, 17.3% mixed ancestry (defined in Bardien et al. 2009), 8% Black, and 1.5% Indian. Thirty-six percent of patients had a positive family history and both autosomal recessive and autosomal dominant patterns of inheritance were present. Control samples (62 mixed ancestry, 62 Caucasian and 70 Black) were recruited from unrelated blood donors at the South African Western Province Blood

Transfusion Service blood collection clinics. The controls were 'de-identified' and had not been clinically assessed for signs of PD.

Bioinformatic analysis (CAESAR)

The CAESAR program was obtained from the authors (http://visionalb.bio.unc.edu/) and installed on an Intel Pentium PC running on Ubuntu Linux 9.10 operating system with internet connectivity. Four literature reviews on PD (Bonifati 2007; Klein and Lohmann-Hedrich 2007; Thomas and Beal 2007; Onyango 2008) were downloaded from NCBI's PubMed (http://www.ncbi.nlm.nih.gov/pubmed) and used as the corpus (i.e. input text for CAESAR). The selected publications had to be converted from PDF format to text using the Linux utility program 'pdf2text'. The texts were then edited to remove references and remnants of the image files.

In CAESAR, human gene symbols from the corpus are extracted using a reference list of standard names, symbols, database identifiers and corresponding mouse homologs from Entrez Gene (Maglott et al. 2005) and Ensembl (Birney et al. 2006). The selected genes are then weighted based on the frequency of occurrence of the respective genes and quantified using four catalogued biomedical ontologies namely, gene ontology biological process (GO bp) (Harris et al. 2004), gene ontology molecular function (GO mf) (Harris et al. 2004), mammalian phenotype ontology (MP) (Smith et al. 2005) and eVOC anatomical ontology (Kelso et al. 2003). The quantification is achieved through a similarity search using a vector-space model.

The quantified gene information is then ranked by mapping them onto eight sources of gene-centric information databases resulting in eight scores for each gene. The gene-centric data sources are Mouse Genome Database (MGD), Genetic Association Database (GAD), UniProt Database, Gene Ontology Annotation Database (GOA), Biomolecular Interaction Network Database (BIND), Human Protein Reference Database (HPRD), Kyoto Encyclopaedia of Genes and Genomes Pathway Database (KEGG), and InterPro Protein Domain Database (IPro). Finally, scores generated for each gene through mapping are mathematically integrated to produce a ranked list of candidate genes. Two of the top ten ranked genes from this list were selected as possible candidate genes for mutation screening.

Genetic analysis

Peripheral blood samples were collected for each study participant for genetic analysis and genomic DNA was extracted using established methods. Polymerase chain reaction (PCR) primers for exons of the *MAPT* and *SNCAIP* genes that span known functional domains were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/). Primer sequences are available from the authors on request. PCR reactions were performed in a 2720 Thermal Cycler (Applied Biosystems, USA). The 25µl reactions contained 10ng template DNA, 10 pmoles of each primer, 75µM dNTPs (Promega, USA), 1.5mM $MgCl_2$, 1X NH4 buffer (Bioline, UK), 5% formamide (for the *MAPT* primer sets), 2µM SYTO9 fluorescent dye (Invitrogen, USA) and 0.25 units of BIOTAQ DNA polymerase (Bioline, UK). The PCR cycling parameters were as follows: initial denaturation at 94°C for 5min, followed by 30 cycles of denaturation at 94°C for 30sec, annealing at different temperatures (ranging from 55 to 60°C) according to each primer set for 30sec, extension at 72°C for 45sec, and a final extension step at 72°C for 7min.

The PCR products were then subjected to high-resolution melt (HRM) analysis on a Rotor-Gene 6000 analyzer (Corbett Life Sciences, Australia) in which they underwent a melt from 75ºC to 95ºC rising by 0.1ºC. During 'melting' of the DNA, the release of the fluorescent intercalating dye SYTO9 is measured in order to produce thermal denaturation profiles which are characteristic of a specific PCR product. These profiles are dependent on sequence length, base and GC content of the PCR product (Erali et al. 2008;Wittwer et al. 2003). HRM profiles that differed significantly from that of the wild-type sample were sequenced in order to characterize the genetic variant. Sequencing was performed using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems, USA) and analyzed on a 3130xl Genetic Analyzer (Applied Biosystems). The primers used for sequencing were the same as those used to generate the PCR products for HRM analysis. Ethnically matched control samples were screened in order to determine the frequency of selected sequence variants. For naming of missense variants, protein sequences NP_001116538.2 and NP_005451.2 was used for *MAPT* and *SNCAIP*, respectively. The PolyPhen tool (http://coot.embl.de/PolyPhen/) was used to predict the possible functional consequence of the missense variants. This tool combines information on sequence features, a multiple alignment with homologous proteins and several structural parameters to predict the effect of a variant on the protein function. PolyPhen's output is a prediction that the variant is either 'damaging' i.e. is supposed to affect protein function or 'benign' i.e. most likely lacking any phenotypic effect.

Results

Bioinformatic analysis

CAESAR generated a prioritized list of genes of which the top ten (in order of ranking) were *SNCA*, *PARK2*, *SNCAIP*, *APP*, *MAPT*, *HD*, *BCL2L1*, *SNCB*, *TH* and *UCHL1* (Supplementary Table 1) The *SNCA* and *PARK2* genes have been conclusively implicated in the pathogenesis of PD (Lesage and Brice 2009) which accounts for their ranking as number one and two, respectively. The *UCHL1* gene is currently classified as a PD associated gene with unknown relevance (Lesage and Brice 2009). The *APP* and *HD* genes are associated with Alzheimer's disease and Huntington's disease, respectively (Sisodia and Price 1995; Hoogeveen et al. 1993). After excluding the known PD genes (*SNCA, PARK2* and *UCHL1*) and genes involved in other disorders (*APP* and *HD*), the remaining genes in the list were *SNCAIP, MAPT, BCL2L1*, *SNCB* and *TH*. Of these, the two genes with the highest CAESAR scores were selected for further investigation.

Mutation screening

The *MAPT* (microtubule-associated protein tau) and *SNCAIP* (synuclein alpha interacting protein) genes were selected as plausible candidates for mutation screening in a group of 202 South African PD patients. For *MAPT*, two novel missense variants A91V and V635I were identified in two patients (Table 1). A91V was present in one male patient of mixed ancestry who had an AAO of 32 years. This variant was absent from 124 ethnic-matched control chromosomes. V635I was found in a male Black patient (AAO of 60 years) as well as his three unaffected children. It was also present in 0.7% of control chromosomes. A91V and V635I were predicted by PolyPhen to be benign variants even though they are both located at amino acid positions that are evolutionarily conserved from dog to human (Figure 1).

For *SNCAIP*, we found three novel missense variants T383N, R606Q and N906H as well as one previously-identified sequence variant E709Q (Table 1). T383N was found in one Black patient who had a positive family history and an AAO of 50 years. R606Q was identified in one Caucasian patient who had a late AAO of 70 years and no family history of PD. This variant is located in the ankyrin repeat 5 domain. N906H was found in one Caucasian patient with an AAO of 64 years and no family history of PD. The previously reported E709Q variant (Marx et al. 2003) was identified in one Caucasian patient with an AAO of 64 years and is located in the ankyrin repeat 6 domain. E709Q was found in 1.6% of control chromosomes, indicating that it is probably a non-pathogenic polymorphism.

Table 1 Missense sequence variants identified in *MAPT* and *SNCAIP* in South African PD patients

Gene	Location	Variant	Zygosity	Number of patients/ Ethnicity	AAO (yrs)	Family history	Protein domain	Conserved region	Frequency in ethnic matched control chromosomes	Previously reported/ Reference	PolyPhen prediction
MAPT	Exon 4	A91V (GCG-GTG)	Heterozygous	1/ mixed ancestry	32	-	No	Yes	0/124	Novel	Benign
	Exon 12	V635I (GTC-ATC)	Heterozygous	1/ Black	60	-	Microtubule-binding repeat 2	Yes	1/140 (0.7%)	Novel	Benign
SNCAIP	Exon 5	T383N (ACT-AAT)	Heterozygous	1/ Black	50	+	Between ankyrin repeat 1and 2	No	0/124	Novel	Benign
	Exon 10	R606Q (CGG-CAG)	Heterozygous	1/ Caucasian	70	-	Ankyrin repeat 5	Yes	0/124	Novel	Benign
		E709Q (GAG-CAG)	Heterozygous	1/ Caucasian	64	-	Ankyrin repeat 6	Yes	2/124 (1.6%)	Marx et al., 2003	Benign
		N906H (AAC-CAC)	Heterozygous	1/ Caucasian	64	-	No	Yes	0/124	Novel	Possibly damaging

AAO, age at onset; -, Negative family history of PD; +, Positive family history of PD.

Table 2 Synonymous and intronic sequence variants identified in *MAPT* and *SNCAIP* in South African PD patients

Gene	Location	Variant	Protein domain	Ethnicity of patients with variant	Frequency in patients (n=202)	Previously reported/ Reference
MAPT	Intron 3	IVS3+18 C>T		Caucasian, Mixed ancestry, Indian, Black	49	Novel
	Exon 4	A90A (GCC-GCT)	-	Mixed ancestry	1	Novel
	Intron 4	IVS4+9 A>G		Caucasian, Mixed ancestry	31	Novel
	Exon 11	A562A (GCA-GCG)	-	Caucasian, Mixed ancestry	19	Novel
	Exon 11	N590N (AAT-AAC)	Microtubule-binding repeat 1	Caucasian, Mixed ancestry	19	Novel
	Exon 11	P605P (CCG-CCA)	Microtubule-binding repeat 1	Caucasian, Mixed ancestry, Indian, Black	27	Kim *et al*., 2010
	Intron 11	IVS11+40 C>T		Caucasian	2	Novel
SNCAIP	Exon 5	G390G (GGC-GGA)	Ankyrin repeat 2	Black	1	Novel
	Exon 6	S406S (AGC-AGT)	Ankyrin repeat 2	Mixed ancestry	1	Novel
	Intron 6	IVS6-23 T>C		Caucasian, Black	3	Novel
	Exon 8	L504L (CTG-TTG)	Between ankyrin repeat 4 and coiled-coil domain	Caucasian	2	Novel
	Exon 10	L638L (CTG-CTC)	Between ankyrin repeat 5 and 6	Black	1	Novel

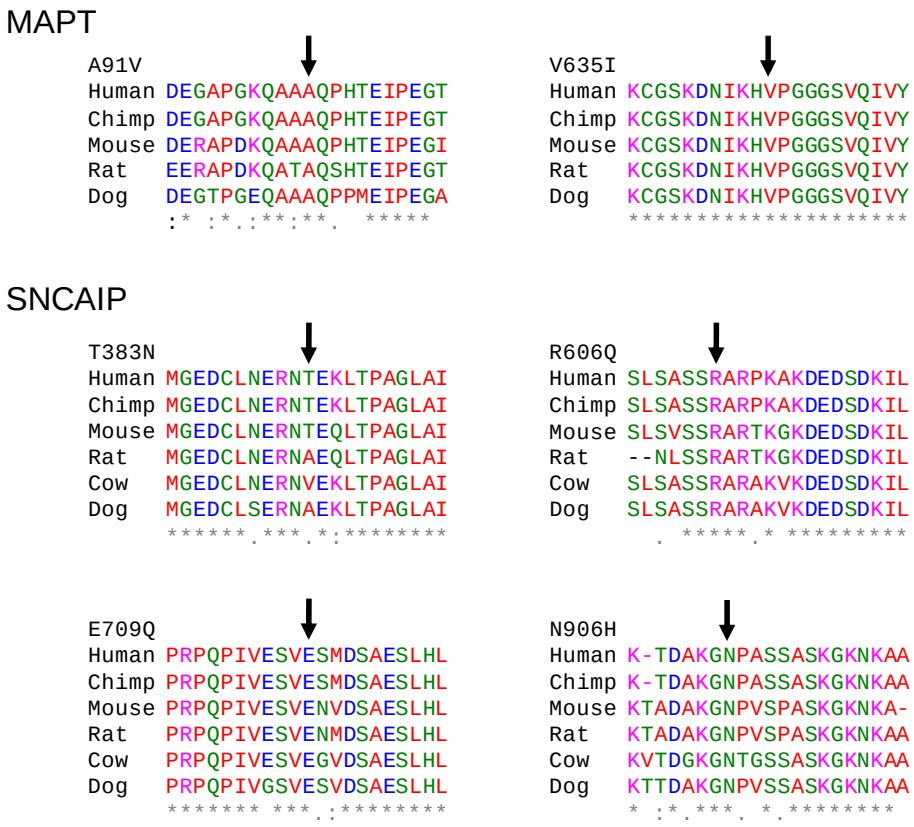


Figure 1 Comparative protein alignments (using ClustalW) of selected regions of MAPT and SNCAIP across different species to determine whether they are evolutionarily-conserved. The protein sequences used were as follows: For MAPT: human (NP_001116538.1), chimp (NP_001009068.1), mouse (NP_001033698.1), rat (ENSRNOP00000006856) and dog (XP_850032.1). For SNCAIP: human (NP_005451.2), chimp (ENSPTRP00000041122), mouse (NP_080684.2), rat (XP_225768.4), cow (XP_591868.3) and dog (XP_538600.2). The arrows indicate the positions of variants. Identical amino acids are indicated by asterisks.

The remaining three *SNCAIP* variants were not found in 124 ethnic-matched control chromosomes and occur within evolutionarily conserved regions of the protein (Figure 1). PolyPhen analysis predicted that they are benign except for N906H which was predicted to be possibly damaging. A number of synonymous and intronic variants were found in both genes (Table 2). The *MAPT* IVS3+18 C>T, IVS4+9A>G, A562A, N590N and P605P are probably non-pathogenic polymorphisms due to the high frequencies observed in our study and the remaining variants are of unknown functional significance but are not likely to be pathogenic.

Discussion

It is important that novel candidate genes for PD are identified. PD genes may exhibit population specificity as has been shown for disease-causing mutations in the *LRRK2* gene which are a relatively common cause of PD in non-Asian populations but to date have not

been found in Asian populations. In the present study, we used a bioinformatics approach to search for novel PD candidates for screening in a South African population. Bioinformatics is a useful tool because it can be used to systematically and methodically search through vast amounts of information stored in various databases and is free of inherent human biases. CAESAR is also much faster, and to have attempted this work manually may have taken many months to accomplish. The *SNCAIP* and *MAPT* were selected and screened in a group of South African PD patients. In both genes, a number of novel variants were found of unknown functional significance.

The *MAPT* gene was selected as a plausible PD candidate as it has been shown to be a susceptibility factor for PD in a number of independent GWAS studies (Edwards et al. 2010; Simon-Sanchez et al. 2009). It is possible that a single gene may harbour both disease-causing and susceptibility alleles for PD, as illustrated by *LRRK2* in which several disease-causing mutations have been found, as well as two susceptibility factors G2385R and R1628P. MAPT, which is highly expressed in neurons, consists of 15 exons which undergo complex and regulated alternative splicing to produce a number of protein isoforms. The longest version, isoform 6, encodes a 776 amino acid protein. MAPT is needed for organizing and maintaining cell structure by modulating microtubules (Weingarten et al. 1975; Hirokawa 1994). Interactions between MAPT and the microtubules occur via the microtubule-binding repeat domains located in the carboxyl-terminus (Lee et al. 1989). Aggregation of MAPT protein results in the formation of filamentous cytoplasmic inclusions which lead to the development of tauopathies including frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17) (Rademakers et al. 2004). The phosphorylation status of this protein is important since hyperphosphorylation of MAPT produces conformational changes which reduce the ability of the protein to interact with microtubules and leads to the formation of the aggregations (Morishima-Kawashima et al. 1995; Rademakers et al. 2004). MAPT has a polymorphic inversion resulting in two main haplotypes: H1 and H2. H1 has been shown to be associated with an increased risk for PD (Farrer et al. 2002; Zabetian et al. 2007; Tobin et al. 2008; Wider et al. 2010); however the functional variant within this haplotype still needs to be identified. It would be interesting to determine in future studies whether the A91V and V635I variants found in the present study have any effect on the normal functioning of MAPT.

SNCAIP was selected for mutation screening in the present study due to the fact that it has been shown to interact with α-synuclein (Engelender et al. 1999) and parkin (Chung et al. 2001). It has been suggested that synphilin-1 assembles α-synuclein and parkin in a multi-protein complex. The *SNCAIP* gene has 11 exons and encodes a 919 amino acid protein that contains several protein-protein interaction domains including ankyrin repeat domains, a coiled-coil domain, and an ATP and GTP binding domain (Engelender et al. 1999;Engelender et al. 2000). Synphilin-1 is ubiquitously expressed throughout the human body including the substantia nigra, with highest levels found in neurons (Engelender et al. 1999). Although its exact function is not known, it is proposed to localize to the presynapse where it could affect dopamine release by binding to synaptic vesicles (Ribeiro et al. 2002). Synphilin-1 is a major component of Lewy bodies, a pathological hallmark of PD, and has been shown to promote the formation of Lewy-body-like ubiquitin-positive cytosolic inclusions when co-expressed with α-synuclein and parkin (Chung et al. 2001).

Only one *SNCAIP* mutation, R621C has been shown to be associated with PD and was found in two unrelated late-onset idiopathic PD German patients (Marx et al. 2003). R621C is also commonly found in controls (Myhre et al. 2008). However, functional studies have established that it causes an increase in cell vulnerability to cellular stress (O'Farrell et al. 2001;Marx et al. 2003) and transgenic mice overexpressing R621C exhibit decreased motor skill learning and motor performance (Nuber et al. 2010). It is possible that the R606Q variant identified in the present study might have a similar effect as the R621C mutation since both occur in the ankyrin repeat 5 domain of the protein.

Future work is planned to determine the biological effect of the various variants identified in the present study. It will be determined whether A91V and V635I in *MAPT* have an effect on mRNA splicing by altering the ratios of isoforms containing either the four repeat or three repeat microtubule binding domains, which has been found to be increased in parkinsonian disorders (Takanashi et al. 2002; Connell et al. 2005). The variants might also affect phosphorylation of the protein or possibly lead to tau aggregation in the cell. For the T383N, R606Q and N906H variants in *SNCAIP*, it will be investigated whether mutant synphilin-1 is able to form cytoplasmic inclusions and is able to be ubiquitinated by parkin. The majority of these variants were predicted to be benign by PolyPhen analysis but this requires verification by wet-lab functional studies. Previously, the E476K variant in *PINK1* was predicted to be

benign but was shown to adversely affect mitochondrial membrane potential in transfected neuroblastoma cells (Abou-Sleiman et al. 2006).

Conclusion

Bioinformatic approaches are useful and constitute an important toolset to identify disease-causing genes especially in complex diseases such as PD as they can systematically search through vast amounts of information and are free of inherent human biases. Using the CAESAR program in the present study, the *MAPT* and *SNCAIP* genes were selected as plausible candidates. Mutation screening identified a number of novel missense sequence variants in both genes in South African PD patients. Functional studies are planned to investigate whether these novel variants affect the normal functioning of these proteins and the possible role they play in PD neurodegeneration.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

We thank the study participants, Ms D Lombard and the South African Western Province Blood Transfusion Service for collection of blood samples. This work was supported by the South African Medical Research Council, the Harry and Doris Crossley Foundation, the University of Stellenbosch, South Africa, and we also acknowledge funding to RJK from the Oppenheimer Memorial Trust.

References

Abou-Sleiman PM, Muqit MM, McDonald NQ, Yang YX, Gandhi S, Healy DG, Harvey K, Harvey RJ, Deas E, Bhatia K, Quinn N, Lees A, Latchman DS, Wood NW (2006) A heterozygous effect for PINK1 mutations in Parkinson's disease? Ann Neurol 60:414-419

Bardien S, Keyser R, Yako Y, Lombard D, Carr J (2009) Molecular analysis of the parkin gene in South African patients diagnosed with Parkinson's disease. Parkinsonism Relat Disord 15:116-121

Bardien S, Marsberg A, Keyser R, Lombard D, Lesage S, Brice A, Carr J (2010) LRRK2 G2019S mutation: frequency and haplotype data in South African Parkinson's disease patients. J Neural Transm 117:847-853

Birney E, Andrews D, Caccamo M, Chen Y, Clarke L, Coates G, Cox T, Cunningham F, Curwen V, Cutts T, Down T, Durbin R, Fernandez-Suarez XM, Flicek P, Graf S, Hammond M, Herrero J,

Howe K, Iyer V, Jekosch K, Kahari A, Kasprzyk A, Keefe D, Kokocinski F, Kulesha E, London D, Longden I, Melsopp C, Meidl P, Overduin B, Parker A, Proctor G, Prlic A, Rae M, Rios D, Redmond S, Schuster M, Sealy I, Searle S, Severin J, Slater G, Smedley D, Smith J, Stabenau A, Stalker J, Trevanion S, Ureta-Vidal A, Vogel J, White S, Woodwark C, Hubbard TJ (2006) Ensembl 2006. Nucleic Acids Res 34:D556-D561

Bonifati V (2007) Genetics of parkinsonism. Parkinsonism Relat Disord 13 Suppl 3:S233-S241

Chung KK, Zhang Y, Lim KL, Tanaka Y, Huang H, Gao J, Ross CA, Dawson VL, Dawson TM. (2001) Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease. Nat Med 7:1144-1150

Connell JW, Rodriguez-Martin T, Gibb GM, Kahn NM, Grierson AJ, Hanger DP, Revesz T, Lantos PL, Anderton BH, Gallo JM (2005) Quantitative analysis of tau isoform transcripts in sporadic tauopathies. Brain Res Mol Brain Res 137:104-109

Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, Wang L, Zuchner S, Konidari I, Wang G, Singer C, Nahab F, Scott B, Stajich JM, Pericak-Vance M, Haines J, Vance JM, Martin ER (2010) Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. Ann Hum Genet 74:97-109

Engelender S, Kaminsky Z, Guo X, Sharp AH, Amaravi RK, Kleiderlein JJ, Margolis RL, Troncoso JC, Lanahan AA, Worley PF, Dawson VL, Dawson TM, Ross CA (1999) Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. Nat Genet 22:110-114

Engelender S, Wanner T, Kleiderlein JJ, Wakabayashi K, Tsuji S, Takahashi H, Ashworth R, Margolis RL, Ross CA (2000) Organization of the human synphilin-1 gene, a candidate for Parkinson's disease. Mamm Genome 11:763-766

Erali M, Voelkerding KV, Wittwer CT (2008) High resolution melting applications for clinical laboratory medicine. Exp Mol Pathol 85:50-58

Farrer M, Skipper L, Berg M, Bisceglio G, Hanson M, Hardy J, Adam A, Gwinn-Hardy K, Aasly J (2002) The tau H1 haplotype is associated with Parkinson's disease in the Norwegian population. Neurosci Lett 322:83-86

Funayama M, Hasegawa K, Kowa H, Saito M, Tsuji S, Obata F (2002) A new locus for Parkinson's disease (PARK8) maps to chromosome 12p11.2-q13.1. Ann Neurol 51:296-301

Fung HC, Scholz S, Matarin M, Simon-Sanchez J, Hernandez D, Britton A, Gibbs JR, Langefeld C, Stiegert ML, Schymick J, Okun MS, Mandel RJ, Fernandez HH, Foote KD, Rodriguez RL, Peckham E, De Vrieze FW, Gwinn-Hardy K, Hardy JA, Singleton A (2006) Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol 5:911-916

Gaulton KJ, Mohlke KL, Vision TJ (2007) A computational system to select candidate genes for complex human traits. Bioinformatics 23:1132-1140

Gibb WR, Lees AJ (1988) A comparison of clinical and pathological features of young- and old-onset Parkinson's disease. Neurology 38:1402-1406

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, Sethuraman A, Theesfeld CL, Botstein D, Dolinski K, Feierbach B, Berardini T, Mundodi S, Rhee SY, Apweiler R, Barrell D, Camon E, Dimmer E, Lee V, Chisholm R, Gaudet P, Kibbe W, Kishore R, Schwarz EM, Sternberg P, Gwinn M, Hannick L, Wortman J, Berriman M, Wood V, de la CN, Tonellato P, Jaiswal P, Seigfried T, White R (2004) The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 32:D258-D261

Hirokawa N (1994) Microtubule organization and dynamics dependent on microtubule-associated proteins. Curr Opin Cell Biol 6:74-81

Hoogeveen AT, Willemsen R, Meyer N, de Rooij KE, Roos RA, van Ommen GJ, Galjaard H. (1993) Characterization and localization of the Huntington disease gene product. Hum Mol Genet 2:2069-2073

Kelso J, Visagie J, Theiler G, Christoffels A, Bardien S, Smedley D, Otgaar D, Greyling G, Jongeneel CV, McCarthy MI, Hide T, Hide W (2003) eVOC: a controlled vocabulary for unifying gene expression data. Genome Res 13:1222-1230

Keyser RJ, Lesage S, Brice A, Carr J, Bardien S (2010a) Assessing the prevalence of PINK1 genetic variants in South African patients diagnosed with early- and late-onset Parkinson's disease. Biochem Biophys Res Commun 398:125-129

Keyser RJ, Lombard D, Veikondis R, Carr J, Bardien S (2010b) Analysis of exon dosage using MLPA in South African Parkinson's disease patients. Neurogenetics 11:305-312

Klein C, Lohmann-Hedrich K (2007) Impact of recent genetic findings in Parkinson's disease. Curr Opin Neurol 20:453-464

Lee G, Neve RL, Kosik KS (1989) The microtubule binding domain of tau protein. Neuron 2:1615-1624

Lesage S, Brice A (2009) Parkinson's disease: from monogenic forms to genetic susceptibility factors. Hum Mol Genet 18:R48-R59

Maglott D, Ostell J, Pruitt KD, Tatusova T (2005) Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res 33:D54-D58

Maraganore DM, de AM, Lesnick TG, Strain KJ, Farrer MJ, Rocca WA, Pant PV, Frazer KA, Cox DR, Ballinger DG (2005) High-resolution whole-genome association study of Parkinson disease. Am J Hum Genet 77:685-693

Marx FP, Holzmann C, Strauss KM, Li L, Eberhardt O, Gerhardt E, Cookson MR, Hernandez D, Farrer MJ, Kachergus J, Engelender S, Ross CA, Berger K, Schols L, Schulz JB, Riess O,

Kruger R (2003) Identification and functional characterization of a novel R621C mutation in the synphilin-1 gene in Parkinson's disease. Hum Mol Genet 12:1223-1231

Morishima-Kawashima M, Hasegawa M, Takio K, Suzuki M, Yoshida H, Watanabe A, Titani K, Ihara Y (1995) Hyperphosphorylation of tau in PHF. Neurobiol Aging 16:365-371.

Myhre R, Klungland H, Farrer MJ, Aasly JO. 2008. Genetic association study of synphilin-1 in idiopathic Parkinson's disease. BMC Med Genet 9:19

Nuber S, Franck T, Wolburg H, Schumann U, Casadei N, Fischer K, Calaminus C, Pichler BJ, Chanarat S, Teismann P, Schulz JB, Luft AR, Tomiuk J, Wilbertz J, Bornemann A, Kruger R, Riess O (2010) Transgenic overexpression of the alpha-synuclein interacting protein synphilin-1 leads to behavioral and neuropathological alterations in mice. Neurogenetics 11:107-120

O'Farrell C, Murphy DD, Petrucelli L, Singleton AB, Hussey J, Farrer M, Hardy J, Dickson DW, Cookson MR (2001) Transfected synphilin-1 forms cytoplasmic inclusions in HEK293 cells. Brain Res Mol Brain Res 97:94-102

Onyango IG (2008) Mitochondrial dysfunction and oxidative stress in Parkinson's disease. Neurochem Res 33:589-597

Rademakers R, Cruts M, van BC (2004) The role of tau (MAPT) in frontotemporal dementia and related tauopathies. Hum Mutat 24:277-295

Ribeiro CS, Carneiro K, Ross CA, Menezes JR, Engelender S (2002) Synphilin-1 is developmentally localized to synaptic terminals, and its association with synaptic vesicles is modulated by alpha-synuclein. J Biol Chem 277:23927-23933

Rossi S, Masotti D, Nardini C, Bonora E, Romeo G, Macii E, Benini L, Volinia S (2006) TOM: a web-based integrated approach for identification of candidate disease genes. Nucleic Acids Res 34:W285-W292

Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, Paisan-Ruiz C, Lichtner P, Scholz SW, Hernandez DG, Kruger R, Federoff M, Klein C, Goate A, Perlmutter J, Bonin M, Nalls MA, Illig T, Gieger C, Houlden H, Steffens M, Okun MS, Racette BA, Cookson MR, Foote KD, Fernandez HH, Traynor BJ, Schreiber S, Arepalli S, Zonozi R, Gwinn K, van der BM, Lopez G, Chanock SJ, Schatzkin A, Park Y, Hollenbeck A, Gao J, Huang X, Wood NW, Lorenz D, Deuschl G, Chen H, Riess O, Hardy JA, Singleton AB, Gasser T (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet 41:1308-1312

Sisodia SS, Price DL (1995) Role of the beta-amyloid protein in Alzheimer's disease. FASEB J 9:366-370

Smith CL, Goldsmith CA, Eppig JT (2005) The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. Genome Biol 6:R7

Takanashi M, Mori H, Arima K, Mizuno Y, Hattori N (2002) Expression patterns of tau mRNA isoforms correlate with susceptible lesions in progressive supranuclear palsy and corticobasal degeneration. Brain Res Mol Brain Res 104:210-219

Thomas B, Beal MF (2007) Parkinson's disease. Hum Mol Genet 16 Spec No. 2:R183-R194

Tobin JE, Latourelle JC, Lew MF, Klein C, Suchowersky O, Shill HA, Golbe LI, Mark MH, Growdon JH, Wooten GF, Racette BA, Perlmutter JS, Watts R, Guttman M, Baker KB, Goldwurm S, Pezzoli G, Singer C, Saint-Hilaire MH, Hendricks AE, Williamson S, Nagle MW, Wilk JB, Massood T, Laramie JM, DeStefano AL, Litvan I, Nicholson G, Corbett A, Isaacson S, Burn DJ, Chinnery PF, Pramstaller PP, Sherman S, Al-hinti J, Drasby E, Nance M, Moller AT, Ostergaard K, Roxburgh R, Snow B, Slevin JT, Cambi F, Gusella JF, Myers RH (2008) Haplotypes and gene expression implicate the MAPT region for Parkinson disease: the GenePD Study. Neurology 71:28-34

van Driel MA, Cuelenaere K, Kemmeren PP, Leunissen JA, Brunner HG (2003) A new web-based data mining tool for the identification of candidate genes for human genetic disorders. Eur J Hum Genet 11:57-63

Weingarten MD, Lockwood AH, Hwo SY, Kirschner MW (1975) A protein factor essential for microtubule assembly. Proc Natl Acad Sci U S A 72:1858-1862

Wider C, Vilarino-Guell C, Jasinska-Myga B, Heckman MG, Soto-Ortolaza AI, Cobb SA, Aasly JO, Gibson JM, Lynch T, Uitti RJ, Wszolek ZK, Farrer MJ, Ross OA (2010) Association of the MAPT locus with Parkinson's disease. Eur J Neurol 17:483-486

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. Clin Chem 49:853-860

Yu W, Wulf A, Liu T, Khoury MJ, Gwinn M (2008) Gene Prospector: an evidence gateway for evaluating potential susceptibility genes and interacting risk factors for human diseases. BMC Bioinformatics 9:528

Zabetian CP, Hutter CM, Factor SA, Nutt JG, Higgins DS, Griffith A, Roberts JW, Leis BC, Kay DM, Yearout D, Montimurro JS, Edwards KL, Samii A, Payami H (2007) Association analysis of MAPT H1 haplotype and subhaplotypes in Parkinson's disease. Ann Neurol 62:137-144

Supplementary Table 1

Table 1 Ranked list of PD candidate genes selected by the CAESAR program

Rank	Gene Symbol	Mathematically integrated score	Publicly available data sources							
			Genetic association data	Gene ontology biological process data	Gene ontology molecular function data	Biomolecular interaction network data	InterPro protein domain data	Mouse genome data	Kyoto encyclopedia of genes and genomes pathway data	UniProt data
1	**SNCA (α-synuclein)**	0.000477719	ND	2.33E-05	3.93E-05	4.39E-05	0.000179919	1.89E-05	0.000139595	3.29E-05
2	**PARK2 (parkin)**	0.000474014	ND	2.68E-05	3.93E-05	0.00017545	2.25E-05	7.03E-05	0.000139595	ND
3	SNCAIP (synphilin-1)	0.000466342	ND	1.61E-05	3.93E-05	0.00017545	6.30E-05	ND	0.000139595	3.29E-05
4	APP (amyloid beta (A4) precursor protein)	0.000456547	ND	4.25E-05	3.93E-05	0.00017545	2.25E-05	4.29E-06	0.000139595	3.29E-05
5	MAPT (microtubule-associated protein tau)	0.000453194	ND	2.55E-05	3.97E-05	0.00017545	1.35E-05	2.66E-05	0.000139595	3.29E-05
6	HD (huntingtin)	0.000449758	ND	6.31E-05	3.93E-05	5.70E-05	6.30E-05	5.49E-05	0.000139595	3.29E-05
7	BCL2L1 (bcl-2-like protein 1)	0.000448355	ND	6.49E-05	2.53E-05	0.00017545	ND	1.03E-05	0.000139595	3.29E-05
8	SNCB (β-synuclein)	0.000427042	ND	ND	1.48E-05	0.00017545	0.000179919	2.40E-05	ND	3.29E-05
9	TH (tyrosine hydroxylase)	0.000426525	1.92E-05	5.95E-05	1.15E-05	0.00017545	1.35E-05	7.71E-06	0.000139595	ND
10	**UCHL1 (ubiquitin carboxyl-terminal esterase L1)**	0.000426485	ND	2.15E-05	3.21E-05	0.00017545	4.50E-05	1.29E-05	0.000139595	ND

Genes selected for mutation screening are shaded. Known PD causing genes are shown in bold. The eight scores generated by CAESAR for each gene using the publically available data sources are shown in the columns. ND, no meaningful data retrieved from that data source and the score generated was too low.

Chapter 7

Additional unpublished results

Additional data obtained that has not been published are presented in this Chapter.

Mutation screening of the *SNCA* gene

Only three missense mutations (A30P, E46K and A53T) have been found in *SNCA* and these occur in exons 2 and 3. These two exons were subjected to mutation screening by HRM analysis in 119 patients that were diagnosed with early- and late-onset PD. No pathogenic mutations or novel sequence variants were identified.

Mutation screening of the *LRRK2* gene

Nine out of the 51 exons of *LRRK2*, namely, exons 21, 23, 24, 25, 27, 31, 32, 35 and 41, in which putative disease-causing mutations have previously been reported were screened by direct sequencing in 195 patients. The patients were selected according to AAO>40 years and autosomal dominant inheritance patterns. Results obtained from this study are shown in Table 7.1. Two known pathogenic mutations were found. One individual had a heterozygous R1441C mutation and another was homozygous for G2019S. Notably, both of these individuals have late AAO of 62 years and 58 years, respectively. This finding brings the total number of G2019S-positive patients in our study group to five. A putative pathogenic mutation (S1228T) was detected in one individual. S1228T has previously been found in two affected siblings and was not present in 1200 controls (Berg *et al.* 2005). PolyPhen analysis predicted that S1228T is a benign variant.

Furthermore, three novel missense sequence variants were found (E899D, R924H, I997N) which occur in evolutionary conserved regions of the gene. PolyPhen analysis predicted the E899D variant to be benign and the R924H variant to be possibly damaging. Both of these variants occur between two functional domains (ANK and LRR). The I997N variant is located in the LRR domain and was predicted by PolyPhen analysis to be probably damaging. A number of synonymous and intronic variants were also identified. Future work will be to assess the

prevalence of the missense variants in ethnic matched control individuals and to determine their pathogenic significance by means of functional studies.

Haplotype studies were performed, according to the methodology in Chapter 4, for the individual with the homozygous G2019S mutation and this revealed that she was homozygous for haplotype-1 (Figure 7.1).

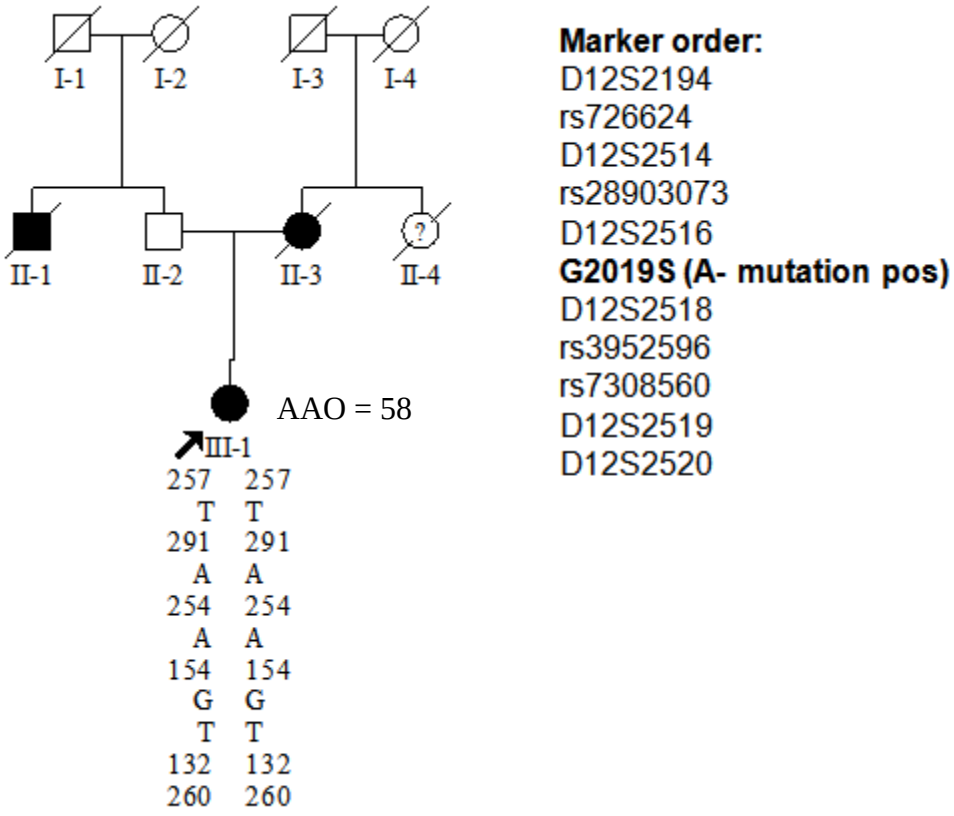


Figure 7.1 Pedigree of the patient with the homozygous G2019S mutation. The disease-associated haplotype is shown. Arrow indicates the proband, *open square* male, *open circle* female, *filled square and circle* PD affected individual, *square and circle with a strike through* deceased individual, AAO: age at onset.

Table 7.1 Pathogenic mutations and novel sequence variants identified in *LRRK2* in South African PD patients

Location		Variant	Zygosity	Number of patients/ethnicity	AAO (yrs)	Family history	Protein domain	Previously reported
Intron 20		IVS20-65 A>T	Heterozygous	2/Black				Novel
	Exon 21	E899D	Heterozygous	1/Black	51	+	between ANK and LRR domain	Novel
	Exon 21	R924H	Heterozygous	1/Afrikaner	44	-	between ANK and LRR domain	Novel
	Exon 23	I997N	Heterozygous	1/Caucasian	65	-	LRR	Novel
Intron 23		IVS23+78 C>T	Heterozygous	1/Caucasian				Novel
Intron 26		IVS26-89 T>A	Heterozygous	3/Black; 2/Mixed ancestry				Novel
Intron 26		IVS26-33 T>C	Heterozygous	2/Black; 1/Mixed ancestry				Novel
Intron 26		IVS26-29 G>A	Heterozygous	2/Mixed ancestry				Novel
	Exon 27	E1224E	Heterozygous	1/Mixed ancestry 1/Black	48 40	-	LRR	Novel
	Exon 27	S1228T	Heterozygous	1/ Caucasian	50	+	LRR	(Berg *et al.* 2005)
Intron 27		IVS27+7 C>T	Heterozygous	1/Caucasian				Novel
	Exon 31	**R1441C**	**Heterozygous**	**1/ Mixed ancestry**	**62**	**-**	**ROC**	**(Mata *et al.* 2005)**
	Exon 32	R1514Q	Heterozygous	4/Caucasian			COR	Known polymorphism (Toft *et al.* 2007)
	Exon 32	P1542S	Heterozygous	3/Afrikaner; 2/Caucasian			COR	Known polymorphism (Di Fonzo *et al.* 2006)
	Exon 35	S1721S	Heterozygous	1/Caucasian 1/Black	56 74	+ -	COR	Novel
Intron 35		IVS35+23 T>A	Heterozygous and homozygous	122/Afrikaner, Caucasian, Mixed ancestry, Black, Indian				Novel
	Exon 41	**G2019S**	**Homozygous**	**1/ Caucasian (Russian)**	**58**	**+**	**MAPKKK**	**(Kachergus *et al.* 2005; Lesage *et al.* 2009)**

Pathogenic mutations are shown in bold. AAO: age at onset; ANK: Ankyrin repeat; COR: C-terminal of ROC; LRR: leucine rich repeat; Roc: Ras of complex protein; MAPKKK: mitogen-activated protein kinase kinase kinase

Reference list

Berg, D., K. Schweitzer, P. Leitner, A. Zimprich, P. Lichtner *et al.* 2005 Type and frequency of mutations in the LRRK2 gene in familial and sporadic Parkinson's disease. *Brain* 128: 3000-3011.

Di Fonzo, A., C. Tassorelli, M. M. De, H. F. Chien, J. Ferreira *et al.* 2006 Comprehensive analysis of the LRRK2 gene in sixty families with Parkinson's disease. *Eur.J.Hum.Genet.* 14: 322-331.

Kachergus, J., I. F. Mata, M. Hulihan, J. P. Taylor, S. Lincoln *et al.* 2005 Identification of a novel LRRK2 mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations. *Am.J.Hum.Genet.* 76: 672-680.

Lesage, S., C. Condroyer, A. Lannuzel, E. Lohmann, A. Troiano *et al.* 2009 Molecular analyses of the LRRK2 gene in European and North African autosomal dominant Parkinson's disease. *J.Med.Genet.* 46: 458-464.

Mata, I. F., J. M. Kachergus, J. P. Taylor, S. Lincoln, J. Aasly *et al.* 2005 Lrrk2 pathogenic substitutions in Parkinson's disease. *Neurogenetics.* 6: 171-177.

Toft, M., I. F. Mata, O. A. Ross, J. Kachergus, M. M. Hulihan *et al.* 2007 Pathogenicity of the Lrrk2 R1514Q substitution in Parkinson's disease. *Mov Disord.* 22: 389-392.

URL reference list

Polymorphism Phenotyping (PolyPhen): [http://genetics.bwh.harvard.edu/pph/] date accessed 11/10/2010

Chapter 8

Identification of a novel functional deletion variant in the 5'-UTR of the DJ-1 gene

Rowena J Keyser[1], Lize van der Merwe[2,3], Mauritz Venter[4], Craig Kinnear[1], Louise Warnich[4], Jonathan Carr[5], and Soraya Bardien[1]

[1]Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Cape Town, 8001, South Africa

[2]Biostatistics Unit, Medical Research Council, Cape Town, 8001, South Africa

[3]Department of Statistics, University of Western Cape, Cape Town, 8001, South Africa

[4]Department of Genetics, University of Stellenbosch, Stellenbosch, 7600, South Africa

[5]Division of Neurology, Faculty of Health Sciences, University of Stellenbosch, Cape Town, 8001, South Africa

Published in: BMC Medical Genetics 2009 Oct 13;10:105.

Impact factor of journal: 2.84

My contribution to this project: Planning of study

Performed all experimental procedures including: generation of constructs, cell culture, transfections and luciferase assay, clonogenic survival assay, RFLP analysis and sequencing

Analysis and interpretation of data

Writing of manuscript

Abstract

Background

DJ-1 forms part of the neuronal cellular defence mechanism against oxidative insults, due to its ability to undergo self-oxidation. Oxidative stress has been implicated in the pathogenesis of central nervous system damage in different neurodegenerative disorders including Alzheimer's disease and Parkinson's disease (PD). Various mutations in the *DJ-1* (*PARK7*) gene have been shown to cause the autosomal recessive form of PD. In the present study South African PD patients were screened for mutations in *DJ-1* and we aimed to investigate the functional significance of a novel 16 bp deletion variant identified in one patient.

Methods

The possible effect of the deletion on promoter activity was investigated using a Dual-Luciferase Reporter assay. The *DJ-1* 5'-UTR region containing the sequence flanking the 16 bp deletion was cloned into a pGL4.10-Basic luciferase-reporter vector and transfected into HEK293 and BE(2)-M17 neuroblastoma cells. Promoter activity under hydrogen peroxide-induced oxidative stress conditions was also investigated. Computational (*in silico*) *cis*-regulatory analysis of *DJ-1* promoter sequence was performed using the transcription factor-binding site database, TRANSFAC via the PATCH™ and rVISTA platforms.

Results

A novel 16 bp deletion variant (g.-6_+10del) was identified in *DJ-1* which spans the transcription start site and is situated 93 bp 3' from a Sp1 site. The deletion caused a reduction in luciferase activity of approximately 47% in HEK293 cells and 60% in BE(2)-M17 cells compared to the wild-type ($P<0.0001$), indicating the importance of the 16 bp sequence in transcription regulation. The activity of both constructs was up-regulated during oxidative stress. Bioinformatic analysis revealed putative binding sites for three transcription factors AhR, ARNT, HIF-1 within the 16 bp sequence. The frequency of the g.-6_+10del variant was determined to be 0.7% in South African PD patients (2 heterozygotes in 148 individuals).

Conclusion

This is the first report of a functional *DJ-1* promoter variant, which has the potential to influence transcript stability or translation efficiency. Further work is necessary to determine the extent to which the g.-6_+10del variant affects the normal function of the *DJ-1* promoter and whether this variant confers a risk for PD.

Background

The *DJ-1* gene (*PARK7*; OMIM 602533) was first described about a decade ago [1] and encodes a 189 amino acid protein which belongs to the DJ-1/Thi/PfpI protein super family [2,3]. It is ubiquitously expressed in a variety of mammalian tissues including the brain, and was initially described in association with oncogenesis and male rat infertility [1,4,5]. Later it was shown to be associated with autosomal recessive early-onset Parkinson's disease (PD) [3,6]. A few PD-causing mutations have been identified including exon deletions, truncations, homozygous and heterozygous point mutations, which predominantly result in loss of function [3,7]. The expression of DJ-1 in the central nervous system (CNS) is not restricted to specific anatomical or functional systems and it is located in neuronal and glial cells within the substantia nigra [8,9].

DJ-1 is proposed to play a role in protecting neurons from oxidative stress and protecting against mitochondrial damage [10,11]. Oxidative stress can be defined as an imbalance between reactive oxygen species (ROS) production and the antioxidant capacity of a cell. Mitochondrial dysfunction leading to increased ROS can cause damage to various cellular components such as unsaturated lipids, proteins, and nucleic acids, and this has been implicated in various neurodegenerative disorders including Alzheimer's disease, amyotrophic lateral sclerosis and PD [12]. Recent studies aimed at elucidating the function of DJ-1, have found that it is able to undergo self oxidation in order to eliminate hydrogen peroxide (H_2O_2) and, in so doing, acts as a scavenger of ROS [11,13]. An increase in susceptibility to oxidative stress of mammalian cells in cases where DJ-1 was knocked down has been reported [10,11]; cells were more susceptible to H_2O_2-induced cell death due to DJ-1 knockdown and over-expression of wild-type DJ-1 rescued the cells [14]. It has been reported that DJ-1 is oxidatively damaged in the brains of PD patients [15,16]. Studies showed that under oxidative stress conditions, DJ-1 undergoes a shift in its isoelectric point which leads to the accumulation of acidic isoforms in PD frontal cortex tissue compared to age-matched controls [15]. Up-regulation of DJ-1 as well as the intracellular redistribution of DJ-1 to the mitochondria under oxidative stress conditions has been reported [10,17]. The mitochondria have a central role in free radical generation and it has been proposed that DJ-1 in the mitochondria may have a role in preventing mitochondrial injury or decreasing mitochondrial ROS production. Redistribution of DJ-1 to the mitochondria might be the neuronal cellular defence mechanism against oxidative insults [9,10,18].

DJ-1 is highly conserved across diverse species [19] and there have been limited reports of sequence variants. An 18 bp insertion/deletion polymorphism (g.168_185del) in *DJ-1's* promoter

region has been reported to not confer a risk for PD [20,21]. However, a homozygous duplication of this 18 bp sequence (g.168_185dup), as well as a homozygous E163K mutation, have been identified in an Italian family with a Parkinsonism-Dementia-Amyotrophic Lateral Sclerosis phenotype [22]. The duplication and the E163K variants were not found in 1,400 and 500 control chromosomes, respectively. It has been reported that the E163K mutation compromises the ability of DJ-1 to protect against oxidative stress induced by H_2O_2 [23]. It is not known whether the phenotype in this family was due to the g.168_185dup mutation or the E163K mutation, or both.

In the present study, we investigated the functional significance of a novel 16 bp deletion which we identified in the promoter region of *DJ-1* in a PD patient by means of luciferase functional expression studies. Unlike the 18 bp insertion/deletion variant, the 16 bp deletion occurs in a transcriptionally important region of the gene as it spans the transcription start site.

Methods

Study participants

The study protocol was approved by the Committee for Human Research at the University of Stellenbosch, South Africa (Protocol number 2002/C059). In the initial part of the study 30 unrelated South African PD patients were recruited with informed written consent from the Movement Disorders clinic at Tygerberg Hospital in South Africa. The patients were examined by a movement disorder specialist (JC) and all met the UK Parkinson's Disease Society Brain Bank Research criteria for diagnosis of PD [24]. Inclusion criteria for recruitment of patients were early onset and/or positive family history of PD. The average age at onset (AAO) of the study group was 43 years and ranged from 17 years to 77 years. The percentage of males in our study group was 66%. Once the 16 bp deletion variant was detected, the frequency of the variant was determined in a total of 148 unrelated South African PD patients (average AAO = 52 years). Blood samples for 62 controls representing the mixed ancestry ethnic group were recruited from unrelated blood donors at the South African Western Province Blood Transfusion Service blood collection clinics. The controls had been 'de-identified' and had not been clinically assessed for signs of PD.

Genetic analysis

The single-strand conformational polymorphism analysis method was used to screen the 30 PD patients for genetic variations in all 7 exons and the 5' UTR of *DJ-1*. Samples exhibiting altered

mobilities on mildly-denaturing polyacrylamide gels (containing 5% glycerol and 15% urea) were sequenced in order to characterize the sequence variation. Direct sequencing was performed using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems, Foster City, USA) and analysis was conducted on a 3130 xl Genetic Analyser (Applied Biosystems).

Construction of luciferase vectors

Functional analysis of a novel 16 bp deletion identified in the 5'-UTR was performed using a Dual-Luciferase Reporter Assay System (http://www.promega.com/pGL4). In this assay, the activities of the firefly (*Photinus pyralis*; the experimental reporter gene) and the *Renilla* (*Renilla reniformis* or sea pansy; the internal control gene) luciferases are measured sequentially from a single sample. The firefly luciferase gene is present on the promoterless pGL4.10 [*luc*2] vector and the *Renilla* luciferase gene is on the phRL-SV40 vector. The pGL4 luciferase reporter vector was used in this study because it has been codon optimized for more efficient expression in mammalian cells. The reporter gene and the vector backbone have been engineered to reduce the number of consensus transcription factor binding sites, which leads to reduced background luminescence and the risk of anomalous transcription. In this experiment the pGL4.10 [*luc*2] and phRL-SV40 vectors are used to co-transfect mammalian cells and due to their distinct evolutionary origins the activity of two luciferases can be distinguished since they have different enzyme structures and substrate requirements.

The promoter region of *DJ-1* containing sequence spanning the 16 bp deletion was PCR amplified from genomic DNA from the individual harbouring the deletion variant using primers containing restriction sites for cloning. Primers were designed to PCR amplify a fragment from position 1 to 2119 bp (GenBank accession number AB045294). It was decided to include the entire 5'UTR sequence in the luciferase assay to ensure that all sites important for transcription initiation and regulation, especially under hydrogen peroxide-induced oxidative stress conditions, were present. A similar sized construct had previously been used successfully to assess promoter function of *DJ-1* in a study done by Taira et al., 2001 [19].

Two constructs were generated; wild-type [pDJ-1(wt)Luc] and a 16 bp deletion variant [pDJ-1(del)Luc]. Primer sequences used were: forward 5'-ATC GTA TCG CTC GAG GGA TCC TTC TAA GCT CAT TC-3' and reverse 5'-CAG AGC TCT TTT GGA AGC AAG CTT CGATACGAT-3'. PCR reactions were performed in a 2720 Thermal Cycler (Applied Biosystems). The 50 µl reactions contained 200 ng template DNA (wild-type or deletion), 20

pmoles of each primer, 75 μM dNTPs (Promega), 1.5 mM $MgCl_2$, 1× NH_4 buffer (Bioline), 5% DMSO and 0.5 units of BIOTAQ DNA polymerase (Bioline). The PCR cycling parameters entailed an initial denaturation step at 95°C for 7 min, followed by 30 cycles of denaturation at 95°C for 30 sec, annealing at 50°C for 30 sec and extension at 72°C for 2 min, and a final extension step at 72°C for 10 min. The 2,151-kb sized PCR products (either wild-type sequence or 16 bp deletion sequence), as well as pGL4.10 [*luc*2] were digested with *XhoI* and *HindIII* (Fermentas) and purified (Wizard SV Gel and PCR Clean-up system, Promega). Vector arms were dephosphorylated (CIP, Promega) and ligated overnight to the digested PCR fragments. The pDJ-1(wt)Luc and pDJ-1(del)Luc constructs were subcloned into E. coli DH5α cells and single colonies were miniprepped (GeneJET plasmid miniprep kit, Promega). Insertion of the *DJ-1* promoter sequence in-frame with the firefly luciferase gene in the pGL4.10 [*luc*2] vector was verified by direct sequencing using the BigDye Terminator Sequence Ready Reaction kit version 3.1 (Applied Biosystems) and analyzed on a 3130 xl Genetic Analyzer (Applied Biosystems).

PCR-restriction fragment length polymorphism (RFLP) analysis

PCR-restriction fragment length polymorphism (RFLP) analysis was used to screen 148 PD patients and 62 control samples for the presence of the 16 bp deletion variant. The control samples were of South African mixed ancestry descent which can be defined as an admixture of indigenous African populations and immigrants from mainly Western Europe, Madagascar, Malaysia and India. The following PCR primers were designed which span the 16 bp deletion: forward 5'-ACC CAG GGC TGT CCA GCT A-3' and reverse 5'-GTC CAG CAC AGG GAC ACC-3' and produced a PCR product of 321 bp and 305 bp for the wild type and deletion alleles, respectively. A total of 8 μl of the PCR product was digested overnight at 37°C with 5 units of *KpnI* (Promega) in a final volume of 20 μl. Thereafter, the digested products were electrophoresed on 12% polyacrylamide gels and the bands visualized by silver staining. Following the *KpnI* digest, the samples could be genotyped since the wild type allele produced two fragments of 129 bp and 192 bp whereas the 16 bp deletion allele produced two fragments of 129 bp and 176 bp.

Cell culture

Human embryonic kidney cells (HEK293) (Highveld Biological, Pty, Ltd, South Africa) and human dopaminergic neuroblastoma BE(2)-M17 cells (American Type Culture Collection, USA) were cultured separately under sterile conditions in Dulbecco-modified Eagle medium (LONZA

BioWhittaker®) supplemented with 10% fetal calf serum, penicillin (100 units/ml), and streptomycin (100 µg/ml) in 5% CO_2 humidified atmosphere at 37°C.

Plasmid transfection and luciferase assay

Cultured HEK293 and M17 cells were plated 24 hrs prior to transfection into 6-well culture dishes at 80-90% confluence and maintained at an atmosphere of 5% CO_2 at 37°C. For transfections, 1 µg of DNA was added with 3 µl Fugene (Roche Biochemicals) and serum-free media to a total volume of 100 µl (DNA/Fugene/media complex). The complex was then added to the cells in the culture dishes in a drop-wise manner. The pDJ-1(wt)Luc and pDJ-1(del)Luc constructs were co-transfected with phRL-SV40, an internal control vector containing the *Renilla* luciferase gene (Promega), in a ratio of 50:1 (pGL4 versus phRL-SV40), in order to control for transfection efficiency. Twenty four hours after transfection, the cells were exposed to H_2O_2 (0-75 µM), in order to produce ROS formation. Forty eight hours after transfection, the cells were gently rinsed with phosphate buffered saline and harvested with Passive Lysis Buffer (Promega). The Dual-luciferase® Reporter Assay System (Promega) was used to measure luciferase activity. Twenty microliters of cell lysate were added to 100 µl of luciferase assay reagent II (Promega) and the firefly luminescence was read using a Modulus 96 Luminometer. Next, 100 µl of Stop&Glo® reagent (Promega) was added to the lysates and renilla luminescence was read. Luminescence values of firefly were normalized with renilla for each construct within an experiment. The luciferase assay was repeated in four independent experiments for each cell type and the luminescence readings were read in triplicate.

Clonogenic survival assay

A clonogenic survival assay was conducted in order to determine the appropriate concentration of H_2O_2 to use in the luciferase assay. The cultured HEK293 and M17 cells were plated in appropriate dilutions (1000-2000 cells/well) into 6-well plates, and incubated for 48 hrs. The cells were then exposed to varying concentrations of H_2O_2 (0, 25, 50 and 75 µM) and incubated for approximately one week. Thereafter, the cells were fixed and stained with 0.2% crystal violet (Merck) overnight for visual colony count. The IC50 values (concentration of H_2O_2 required to inhibit colony formation by 50%) were determined for both cell lines.

Bioinformatics

Computational (*in silico*) *cis*-regulatory analysis of *DJ-1* promoter sequence was performed using the transcription factor-binding site database, TRANSFAC via the PATCH™ platform (http://www.gene-regulation.com). For PATCH™ analysis, the primary parameter settings were

set as: Minimum Length of Sites at 4 and Lower Score Boundary at 85%. The multi-species conserved sequence analyses were performed using the rVISTA (regulatory VISTA) platform (http://genome.lbl.gov/vista/index.shtml) at default cut-off criteria: ≥ 70% identity over a contiguous paired sequence length ≥ 100 bp. rVISTA combines searching the major transcription factor binding site database TRANSFAC with a comparative sequence analysis. *DJ-1* promoter sequences (2 kb upstream and 5'-UTR) from human (Ensembl Gene ID: ENSG00000116288; *H. sapiens*), chimpanzee (ENSPTRG00000000102, *P. troglodytes*), orangutan (ENSPPYG00000001925, *P. pygmaeus*), macaque (ENSMMUG00000019671, *M. mulatta*), tree shrew (ENSTBEG00000009474, *T. belangeri*), horse (ENSECAG00000017251, *E. caballus*), dog (ENSCAFG00000019674, *C. familiaris*), cow (ENSBTAG00000020518, *B. taurus*), mouse (ENSMUSG00000028964, *M. musculus*) and rat (ENSRNOG00000018289, *R. norvegicus*) were retrieved from the Ensembl Genome Browser (archive version at http://jul2008.archive.ensembl.org/index.html).

Statistical analysis

The open-source programming environment R (freely available from http://www.r-project.org) was used for the statistical analysis. A linear mixed-effects model was created to analyse the data. All the relative light units (RLU) values were first log-transformed towards symmetry, and then modelled as a function of the factors including cell line (HEK293 or M17), Type (WT or deletion) and H_2O_2 concentration (0, 25, 50 or 75 μM) as fixed effects. We also included interactions between Type and cell line, as well as between Type and H_2O_2 concentration, both of which were independently significant after adjusting for all the other effects. As the individual experiments yielded RLU values that were dissimilar, each experiment were included in the model as a random effect. All the results reported here come from the same model, and were therefore adjusted for all the other factors in the model, as well as the random effect of individual experiments.

Results

During mutation screening of South African PD patients a novel 16 bp deletion in the promoter region of *DJ-1* was identified. The frequency of this deletion (g.-6_+10del) was estimated to be 0.7% (95% confidence interval 0.2% to 2.4%), based on 2 heterozygotes in 148 South African PD patients screened. Both patients were sporadic individuals of South African mixed ancestry. One of the patients presented with PD at age 38 years. He had displayed early onset of prominent autonomic dysfunction and hallucinations (with psychosis) and had died at the age of 47 years.

The other patient had an age at onset of 56 years and also has a history of psychotic episodes. Furthermore, the g.-6_+10del variant was observed in a homozygous state in one of 62 mixed ancestry control individuals screened (1.6%). This individual was aged 34 years and had not been clinically assessed for signs of PD.

The 16 bp deletion spans the transcription start site of *DJ-1* which had previously been identified by Taira et al. [19] using the method of 'CAP site hunting'. This variant is situated 93 bp downstream of a Sp1 site, which has been shown to be necessary for *DJ-1* transcription [19], and 157 bp upstream from the known g.168_185del polymorphism (Figure 1). We hypothesized that this deletion would disrupt transcription of DJ-1 and this hypothesis was tested using a luciferase reporter assay. DJ-1 has previously been shown to be expressed in the HEK293 cell line [19] and these cells were therefore used in the initial experiments.

To determine the *in vitro* effect of the deletion variant on *DJ-1* promoter activity, two different constructs (wild-type and deletion variant) were generated. The promoter region of *DJ-1* containing the sequence spanning the 16 bp deletion, as well as transcriptionally important sites such as the transcription start site and Sp1 site, was linked to the firefly luciferase gene and the luciferase activity was measured in RLU. Both constructs contained the 18 bp of the polymorphic variant (g.168_185del). The results obtained are therefore likely to be due only to the presence or absence of the 16 bp sequence.

The linear mixed-effects model for log(RLU) values, showed a significant interaction between construct type and H_2O_2 concentration ($p = 0.0010$) and also between construct type and cell line ($p = 0.0119$). All three factors: cell line, construct type and H_2O_2 concentration, had highly significant effects ($p < 0.0001$) on log(RLU) after adjusting for the interactions and experiments. All the P-values presented in the present study come from this model and they are therefore adjusted for the differences between the individual experiments as well as all the other factors in the model [the cell line (HEK293 or M17), construct type (WT or deletion) and H_2O_2 concentration (0, 25, 50 or 75 μM) as well as the interactions between Type and cell line, and between Type and H_2O_2 concentration]. We compared the geometric means (the mean of the logs, which is then anti-logged) of the observed RLU values obtained for each combination of the M17 cells, the construct type and the H_2O_2 concentration to the predicted means generated by the model (results not shown). These results indicated that the model fitted the data well and we therefore have confidence in our results.

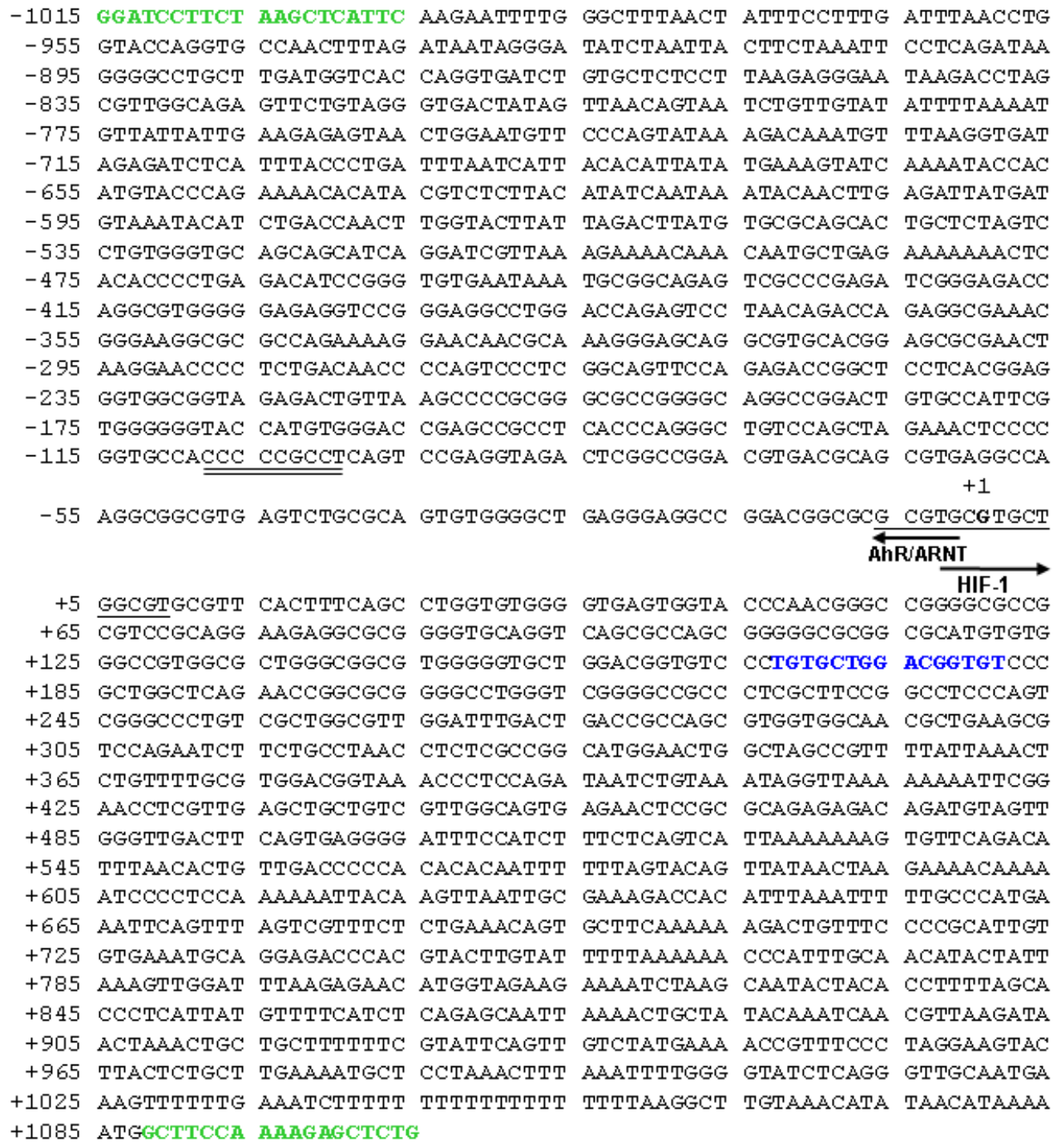


Figure 1 Nucleotide sequence of the promoter region of the human *DJ-1* gene (GenBank Accession number AB045294) showing the positions of significant sites. The +1 indicates the transcription start site. The position of the 16 bp deletion sequence is indicated by a single underline and the Sp1 site by a double underline. The position of the 18 bp insertion/deletion polymorphism (g.168_185del) is indicated in blue font and the positions of the primers used to generate the luciferase constructs are shown in bold green font. The positions of putative *cis*-motifs (AhR/ARNT and HIF-1) present in the 16 bp deletion sequence are indicated by arrows.

The wild-type construct produced consistently higher luciferase activity compared to the activity of a promoterless vector, which indicated that the insert contained a functional promoter (results not shown). Figure 2 shows the distribution of the normalized RLU values obtained for the wild-type and deletion variant in both cell lines. There was a significant difference between the wild-type and the deletion variant in the HEK293 cell line, after adjusting for experiments and all other factors. The deletion caused a reduction in luciferase activity of approximately 47% compared to the wild-type, indicating that the 16 bp sequence is probably important for the transcriptional regulation of *DJ-1* (Figure 2A). Conducting the assay in M17 neuroblastoma cells (Figure 2B) produced similar results with an even larger reduction of 60% in luciferase activity observed for the deletion compared to the wild-type. From the box plots of the data, it was noted that for both cell lines, approximately 75% of the RLU values of the deletion variant were less than the lower 25% quartile of the RLU values for the wild-type (Figure 2), which illustrates the significant difference in the RLU values between the two constructs.

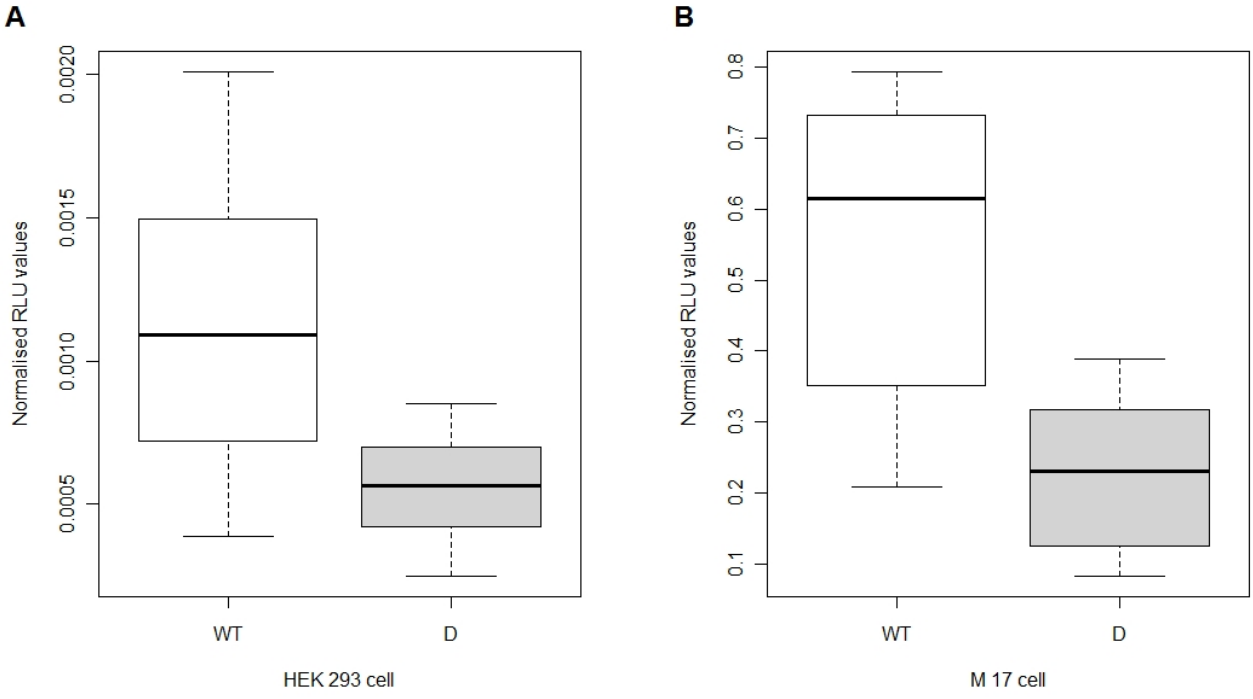


Figure 2 The *DJ-1* 16 bp deletion variant exhibited significantly reduced transcription levels compared to the wild-type ($P < 0.0001$) in two different cell lines. Box plots of dual-luciferase assays of the wild-type and deletion variant using (A) HEK293 cells and (B) neuroblastoma M17 cells. They include median (horizontal line within box), inter-quartile interval i.e. 25^{th} to 75^{th} percentile (box) and the range of variation (whiskers). The 2 kb promoter fragments (containing either wild-type sequence or the deletion) were linked to the luciferase gene and the constructs were transfected into two mammalian cell lines. For each experiment, firefly luciferase activity was divided by renilla activity to normalize for transfection efficiency. Data are representative of four transfection experiments (triplicate points obtained over four independent experiments). WT: wild-type; D: deletion variant; RLU: relative light units.

H_2O_2 treatment of the M17 neuroblastoma cells

The effect of the deletion variant on *DJ-1* promoter activity under H_2O_2-induced oxidative stress conditions was investigated using M17 neuroblastoma cells. DJ-1 is a scavenger of ROS and is proposed to play a role in protecting neurons from oxidative stress. A clonogenic survival assay was first performed to assess the concentration of H_2O_2 to use for these experiments. The IC50 values (concentration of H_2O_2 required to inhibit colony formation by 50%) were shown to be 50 μM for both cell lines (results not shown). The H_2O_2 concentration of 75 μM had the highest cytotoxic effect on both cell lines.

Under H_2O_2-induced oxidative stress conditions, the deletion variant again exhibited significantly reduced promoter activity compared to the wild-type for the untreated as well as at all three concentrations of H_2O_2 (Figure 3). The promoter activity of both the wild-type and deletion was moderately upregulated during these oxidative stress conditions with increasing concentrations of H_2O_2. However, this trend was not observed at the highest H_2O_2 concentration of 75 μM, possibly because of increased cell death at this dosage. Although both the wild-type and deletion exhibited increased promoter activity with increasing concentrations of H_2O_2, the promoter activity for the deletion variant was always lower than that of the wild-type. In a post-hoc analysis with H_2O_2 as dichotomous [treated (all concentrations > 0) versus untreated] we found a highly significant difference ($p < 0.0001$) in promoter activity between untreated cells and H_2O_2 treated cells (after adjusting for all other factors in model, namely the construct type, the cell line and the H_2O_2 concentrations (0, 25, 50 or 75 μM) and the interactions between type and each of the other two factors).

Computational cis-regulatory analysis of DJ-1

Interestingly, the 16 bp deletion sequence contains three perfect 5'-GCGT-3' and one imperfect repeat (5'-GCTG-3'). The functionality of this particular motif is currently not known. Analysis of the 16 bp deletion sequence (using the TRANSFAC database via PATCH and rVISTA) revealed the over-representation of three putative *cis*-motifs that correlate to the transcription factors Aryl hydrocarbon receptor (AhR; TRANSFAC Acc. Nr. T00018), Ahr nuclear translocator (ARNT; T01797) and Hypoxia induced factor 1 (HIF-1; T01609) within the 16 bp sequence (Figure 1). The core binding sequence for AhR and ARNT is 5'-CACGC-3' and for HIF-1 is 5'-RCGTG-3'. These transcriptional regulators control a variety of developmental and physiological events including metabolism of toxins and responses to hypoxia [25,26], and share regulatory cross-talk with many other transcription factors.

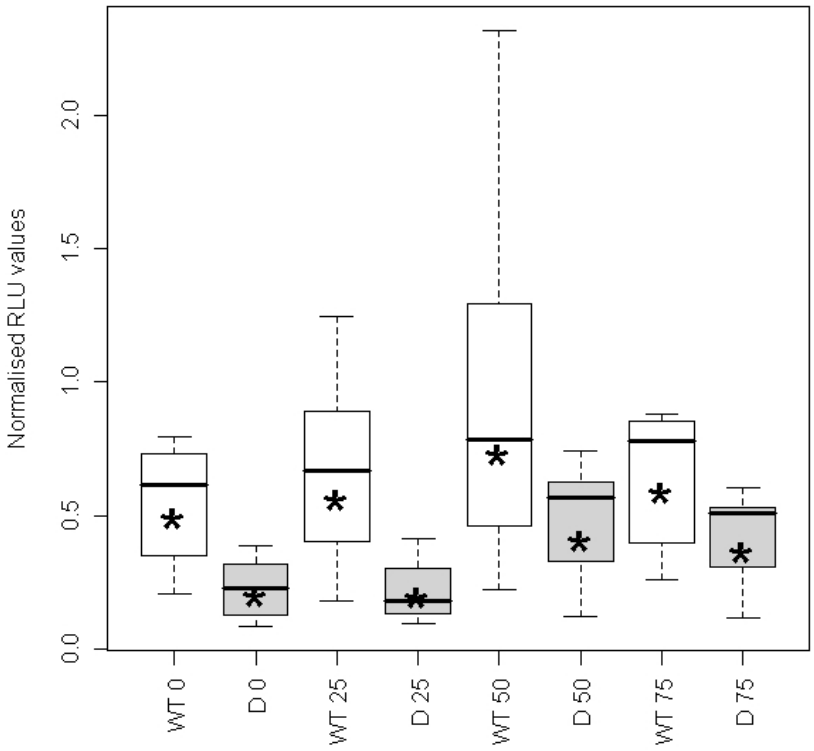


Figure 3 H_2O_2 dose-dependent up-regulation of *DJ-1* promoter activity. Box plots of the dual-luciferase assay of the 16 bp *DJ-1* deletion variant using the neuroblastoma M17 cell line and exposure to different concentrations of H_2O_2 (0, 25, 50 and 75 μM). *: geometric means; WT: wild-type; D: deletion variant; 0, 25, 50, 75: H_2O_2 concentrations.

When the 16 bp sequence is deleted a transcription factor site is created. *In silico* promoter analysis revealed that the sequence in the (-) strand (GCGCGTTC) of the 16 bp deleted region conforms to a novel *cis*-motif that putatively correlated to *cis/trans* interaction with the E2F-class of transcription factors.

Multi-species comparative analysis using rVISTA, [27], showed regions of ≥ 70% conserved identity (over contiguous sequence length: 100 bp) of the human *DJ-1* promoter region to chimpanzee (1754 bp), orangutan (1347 bp), macaque (1239 bp), tree shrew (233 bp) and horse (248 bp). All the other species investigated including rat and mouse showed ≤ 70% conservation (Figure 4).

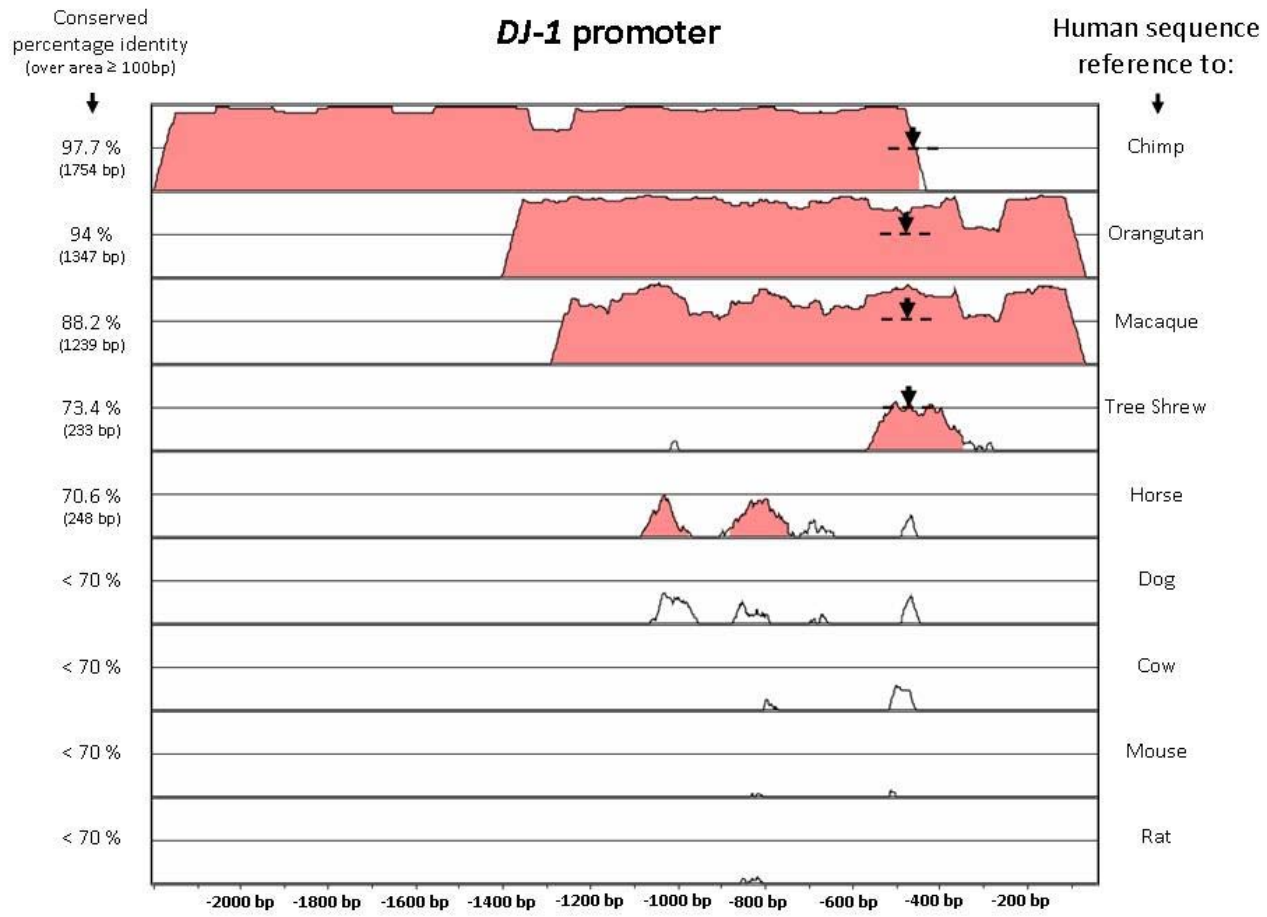


Figure 4 Comparative multi-species analysis using rVISTA of the *DJ-1* promoter region of approximately 2000 bp + 5'-UTR. Conserved percentage identity is indicated with cut-off criteria: ≥ 70% identity over ≥ 100 bp are indicated in orange and the white regions represent either no alignment or ≤ 70% sequence identity over 100 bp contiguous sequence length. The 16 bp deletion sequence is represented by an arrow on a dashed line and is found in a region sharing highly conserved identity between species.

Discussion

In the present study we report a novel 16 bp deletion (g.-6_+10del) in the promoter region of *DJ-1*. The deletion was shown to span the transcription start site, according to Taira et al. 2001 [19] and is situated 93 bp downstream of a Sp1 transcription regulatory sequence which is essential for *DJ-1* promoter activity. In addition, the g.-6_+10del variant is located 157 bp upstream from a known 18 bp duplication mutation (g.168_185dup). The duplication mutation has previously been reported to co-segregate, together with a homozygous E163K mutation, with early-onset Parkinsonism-Dementia-Amyotrophic Lateral Sclerosis in an Italian family [22]. Due to its location within the *DJ-1* promoter, we hypothesised that the 16 bp deletion variant might influence the level of DJ-1 expression, the transcript stability or the translation efficiency.

Functional expression studies using the dual-luciferase reporter system found that the g.-6_+10del variant significantly reduced *DJ-1* promoter activity in both HEK293 and neuroblastoma M17 cells, with the M17 cells exhibiting a higher reduction in activity. This indicates the possible importance of the 16 bp sequence in transcriptional regulation of *DJ-1*. Recent studies aiming to elucidate the function of DJ-1, have identified it as a scavenger of ROS due to its ability to undergo self oxidation in order to eliminate H_2O_2 [11]. It was therefore decided to assess the effect of the deletion variant on *DJ-1* promoter activity under H_2O_2-induced oxidative stress conditions using the neuroblastoma M17 cell line. In the present study it was shown that cells placed under oxidative stress conditions showed a dose-dependent moderate up-regulation in both wild-type and deletion variant's promoter activities with the deletion variant retaining its lower RLU values compared to the wild-type.

Furthermore, bioinformatic analysis identified binding sites for the transcription factors; Aryl hydrocarbon receptor (AhR), Ahr nuclear translocator (ARNT) and Hypoxia induced factor 1 (HIF-1) within the 16 bp deletion sequence. AhR and ARNT are known to dimerize to form an active transcription factor complex that binds defined DNA sequences, the xenobiotic-responsive element (XRE), with high affinity causing an increase in transcription of AhR-regulated genes [26,28,29]. The AhR/ARNT complex up-regulates cytochrome P450 enzymes that play diverse roles in metabolism of endogenous substances, environmental chemicals and various drugs [29]. AhR functions as the prime transcription factor and ARNT as a DNA binding partner. Transcriptional up-regulation during hypoxia is mediated principally by HIF-1 (a dimer of HIF-1α and ARNT). The anticancer property of curcumin has been shown to be due to inactivation of HIF-1 by degradation of ARNT via oxidation and ubiquitination processes [30]. It was also shown that curcumin induces proteasomal degradations of both AhR and ARNT and this is mediated by oxidative stress [31]. It is possible that the AhR/ARNT or HIF-1 complexes may interact with DJ-1's promoter thereby regulating its transcription and therefore indirectly influencing the transcription of various genes associated with oxidative stress, apoptosis and neurotoxicity that are regulated by DJ-1 [32].

Further analysis found that when the 16 bp sequence is deleted, a binding site for the E2F-class of transcription factors is created. E2F has been shown to induce transcription of pro-apoptotic proteins and repression of E2F-responsive genes is required for neuronal survival [33]. Recent studies found that the pRb/E2F cell-cycle pathway is activated in dopaminergic neurons in PD patients, as well as in a PD mouse model [34].

DJ-1 is highly conserved at both the nucleotide and amino acid level, possibly due its important role in protection against oxidative stress. Mouse and human DJ-1 is 83% and 90% identical at the DNA and protein level, respectively [19]. Similarly, the conservation between porcine and human DJ-1 is 96% at the protein level [35]. There have been no reports however on the level of conservation of DJ-1's 5'-UTR sequence. We showed through multi-species comparative analysis that the promoter region of *DJ-1* is highly conserved, as expected, between man and other primates as well as to a few other species. More importantly, the 16 bp deletion sequence is present in a highly conserved promoter region (Figure 4) indicating the possible significance of this sequence in higher order mammals and thus an evolutionarily retained functionality.

It would be interesting to determine the prevalence of the g.-6_+10del variant in other populations and to establish whether it plays a role in susceptibility to neurodegenerative diseases such as PD. Further studies are necessary to determine whether endogenous levels of DJ-1 are affected by the 16 bp deletion and whether an alternative transcriptional start site is present. In addition, introducing point mutations into the predicted transcription factor binding sites within the 16 bp sequence to determine the effect on DJ-1's transcription and the cell's response to oxidative stress would be of interest.

Conclusion

We report a novel sequence variant in the highly conserved DJ-1 gene. Functional expression studies found that this variant significantly reduced *DJ-1* promoter activity in two separate mammalian cell lines, which indicates the possible importance of the 16 bp sequence in transcriptional regulation of *DJ-1*. In addition, the activity of three transcription factors with recognition sites within the deletion sequence might be influenced by the g.-6_+10del variant.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RJK performed the laboratory work and drafted the manuscript. LvdM performed the statistical analysis. MV conducted the bioinformatic analysis. LW and JC critically reviewed the manuscript. CK assisted with laboratory work and critically reviewed the manuscript. SB participated in the conception and design, and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the study participants and Ms D Lombard for collection of blood samples. Ms L Korkie, Mr R Vogt, Dr. M Zaahl and Dr. J Riedemann are thanked for technical assistance with tissue culture, luciferase assays and clonogenic survival assays. This work was supported by the South African Medical Research Council, the Harry and Doris Crossley Foundation and the University of Stellenbosch, South Africa.

References

1. Nagakubo D, Taira T, Kitaura H, Ikeda M, Tamai K, Iguchi-Ariga SM, Ariga H: **DJ-1, a novel oncogene which transforms mouse NIH3T3 cells in cooperation with ras**. *Biochem Biophys Res Commun* 1997, **231(2)**:509-513.

2. Huai Q, Sun Y, Wang H, Chin LS, Li L, Robinson H, Ke H: **Crystal structure of DJ-1/RS and implication on familial Parkinson's disease.** *FEBS Lett*. 2003, **549(1-3)**:171-175.

3. Bonifati V, Rizzu P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, Dekker MCJ, Squitieri F, Ibanez P, Joosse M, et al.: **Mutations in the DJ-1 gene associated with autosomal recessive early-onset Parkinsonism**. *Science* 2003, **299(5604):**256-259.

4. Kim RH, Peters M, Jang Y, Shi W, Pintilie M, Fletcher GC, DeLuca C, Liepa J, Zhou L, Snow B, Binari RC, Manoukian AS, Bray MR, Liu FF, Tsao MS, Mak TW: **DJ-1, a novel regulator of the tumor suppressor PTEN.** *Cancer Cell.* 2005, **7(3):**263-273.

5. Wagenfeld A, Yeung CH, Strupat K, Cooper TG: **Shedding of a rat epididymal sperm protein associated with infertility induced by ornidazole and alpha-chlorohydrin**. *Biol Reprod* 1998, **58(5):**1257-1265.

6. Abou-Sleiman PM, Healy DG, Wood N: **Causes of Parkinson's disease: genetics of DJ-1**. *Cell Tissue Res* 2004, **318**:185-188.

7. Bonifati V, Oostra BA, Heutink P: **Linking DJ-1 to neurodegeneration offers novel insights for understanding the pathogenesis of Parkinson's disease**. *J Mol Med* 2004, **82**:163-174.

8. Bader V, Ran Zhu X, Lübbert H, Stichel CC: **Expression of DJ-1 in the adult mouse CNS**. *Brain Res* 2005, **1041(1)**:102-111.

9. Zhang L, Shimoji M, Thomas B, Moore DJ, Yu SW, Marupudi NI, Torp R, Torgner IA, Ottersen OP, Dawson TM, et al.: **Mitochondrial localization of the Parkinson's disease related protein DJ-1: implications for pathogenesis**. *Hum Mol Genet* 2005, **14**:2063-2073.

10. Lev N, Ickowicz D, Melamed E, Offen D: **Oxidative insults induce DJ-1 upregulation and redistribution: Implications for neuroprotection**. *Neurotoxicology* 2008, **29(3)**:397-405

11. Taira T, Saito Y, Niki T, Iguchi-Ariga SM, Takahashi K, Ariga H: **DJ-1 has a role in antioxidative stress to prevent cell death.** *EMBO Rep* 2004, **5(2):**213-218.

12. Lin MT, Beal MF: **Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases**. *Nature* 2006, **443(7113):**787-795.

13. Kinumi T, Kimata J, Taira T, Ariga H, Niki E: **Cysteine-106 of DJ-1 is the most sensitive cysteine residue to hydrogen peroxide-mediated oxidation in vivo in human umbilical vein endothelial cells**. *Biochem Biophys Res Commun* 2004, **317(3):**722-728.

14. Yokota T, Sugawara K, Ito K, Takahashi R, Ariga H, Mizusawa H: **Down regulation of DJ-1 enhances cell death by oxidative stress, ER stress, and proteasome inhibition**. *Biochem Biophys Res Commun* 2003, **312(4):**1342-1348.

15. Choi J, Sullards MC, Olzmann JA, Rees HD, Weintraub ST, Bostwick DE, Gearing M, Levey AI, Chin LS, Li L, et al.: **Oxidative damage of DJ-1 is linked to sporadic Parkinson and Alzheimer diseases**. *J Biol Chem* 2006, **281(16):**10816-10824.

16. Bandopadhyay R, Kingsbury AE, Cookson MR, Reid AR, Evans IM, Hope AD, Pittman AM, Lashley T, Canet-Aviles R, Miller DW, et al.: **The expression of DJ-1 (PARK7) in normal human CNS and idiopathic Parkinson's disease**. *Brain* 2004, **127**:420-430.

17. Lev N, Ickowicz D, Barhum Y, Lev S, Melamed E, Offen D: **DJ-1 protects against dopamine toxicity**. *J Neural Transm*. 2009,**116(2)**:151-160

18. Canet-Avilés RM, Wilson MA, Miller DW, Ahmad R, McLendon C, Bandyopadhyay S, Baptista MJ, Ringe D, Petsko GA, Cookson MR: **The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfinic acid-driven mitochondrial localization.** *Proc Natl Acad Sci U S A*. 2004, **101(24)**:9103-9108.

19. Taira T, Takahashi K, Kitagawa R, Iguchi-Ariga SM, Ariga H: **Molecular cloning of human and mouse DJ-1 genes and identification of Sp1-dependent activation of the human DJ-1 promoter.** *Gene* 2001, **263(1-2):**285-292.

20. Eerola J, Hernandez D, Launes J, Hellstrom O, Hague S, Gulick C, Johnson J, Peuralinna T, Hardy J, Tienari PJ, et al.: **Assessment of a DJ-1 (PARK7) polymorphism in Finnish PD**. *Neurology* 2003, **61**:1000-1002.

21. Morris CM, O'Brien KK, Gibson AM, Hardy JA, Singleton AB: **Polymorphism in the human DJ-1 gene is not associated with sporadic dementia with Lewy bodies or Parkinson's disease.** *Neurosci Lett* 2003, **352(2):**151-153.

22. Annesi G, Savettieri G, Pugliese P, D'Amelio M, Tarantino P, Ragonese P, La Bella V, Piccoli T, Civitelli D, Annesi F, et al.: **DJ-1 mutations and parkinsonism-dementia-amyotrophic lateral sclerosis complex.** *Ann Neurol* 2005, **58**:803-807.

23. Ramsey CP, Giasson BI: **The E163K DJ-1 mutant shows specific antioxidant deficiency.** *Brain Res* 2008, (1239) 1-11

24. Gibb WR, Lees AJ: **A comparison of clinical and pathological features of young-and old-onset Parkinson's disease.** *Neurology*. 1988, **38(9)**:1402-1406.

25. Wilson CL, Safe S: **Mechanisms of ligand-induced aryl hydrocarbon receptor-mediated biochemical and toxic responses**. *Toxicol Pathol*. 1998, **26(5)**:657-671.

26. Pocar P, Fischer B, Klonisch T, Hombach-Klonisch S: **Molecular interactions of the aryl hydrocarbon receptor and its biological and toxicological relevance for reproduction.** *Reproduction* 2005, **129**:379-389.

27. Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM: **rVista for comparative sequence-based discovery of functional transcription factor binding sites**. *Genome Res.* 2002, **12(5):**832-839.

28. Denison MS, Fisher JM, Whitlock JP Jr: **The DNA recognition site for the dioxin-Ah receptor complex. Nucleotide sequence and functional analysis**. *J Biol Chem*. 1988, **263(33):**17221-17224.

29. Rivera SP, Wang F, Saarikoski ST, Taylor RT, Chapman B, Zhang R, Hankinson O: **A novel promoter element containing multiple overlapping xenobiotic and hypoxia response elements mediates induction of cytochrome P4502S1 by both dioxin and hypoxia**. *J Biol Chem*. 2007, **282(15):**10881-10893.

30. Choi H, Chun YS, Kim SW, Kim MS, Park JW: **Curcumin inhibits hypoxia-inducible factor-1 by degrading aryl hydrocarbon receptor nuclear translocator: a mechanism of tumor growth inhibition.** *Mol Pharmacol* 2006, **70(5):**1664-1671.

31. Choi H, Chun YS, Shin YJ, Ye SK, Kim MS, Park JW: **Curcumin attenuates cytochrome P450 induction in response to 2,3,7,8-tetrachlorodibenzo-p-dioxin by ROS-dependently degrading AhR and ARNT**. *Cancer Sci* 2008, **99(12):**2518-2524.

32. da Costa CA**: DJ-1: a newcomer in Parkinson's disease pathology.** *Curr Mol Med* 2007, 7(7):650-657.

33. Liu DX, Greene LA: **Regulation of neuronal survival and death by E2F-dependent gene repression and derepression.** *Neuron*. 2001, **32(3)**:425-438.

34. Höglinger GU, Breunig JJ, Depboylu C, Rouaux C, Michel PP, Alvarez-Fischer D, Boutillier AL, Degregori J, Oertel WH, Rakic P, et al.: **The pRb/E2F cell-cycle pathway mediates cell death in Parkinson's disease**. *Proc Natl Acad Sci* 2007, **104(9):**3585-3590.

35. Larsen K, Madsen LB, Høj A, Bendixen C: **Porcine DJ-1: cloning of PARK7 cDNA, sequence comparison, expression analysis and chromosomal localization**. *Cytogenet Genome Res* 2007, **116**(1-2):93-09.

Chapter 9

Conclusion

PD had been considered the archetypal non-genetic disorder until the discovery 13 years ago of the first PD-causing mutation in the *SNCA* gene. Since then at least six genes, *PARK2, PINK1, DJ-1, SNCA, LRRK2* and *ATP13A2,* have been conclusively implicated in PD pathogenesis. However, these genes have been shown to account for < 10% of all PD cases which highlights the need to identify additional PD candidate genes. Identification of these genes and the pathways that they are involved in may ultimately result in the development of new therapeutic modalities.

The present study is the first to investigate the genetic basis of PD in South Africa and it was hypothesized that, due to the unique ancestry of this country's populations, these patients may harbor novel mutations. The mixed ancestry sub-population, in particular, has a complex ancestry, being an admixture of at least five different population groups (African, European, Madagascan, Malaysian and Indian). The Afrikaner population is also unique to South Africa. They are known to have undergone a genetic bottleneck at the end of the 19th century, and founder effects for a number of disorders including variegate porphyria and familial hypercholesterolemia have been identified in this group (Jenkins 1990). It could be speculated that a founder effect for PD may also exist in this group. In addition, the South African Black PD population is of interest since Black individuals with PD have been reported to present with atypical clinical features, and a late-onset akinetic-rigid variant (Chaudhuri *et al.* 2000). PD is also thought to be uncommon in the Black population worldwide.

During our investigation of the genetic etiology of PD in South Africa, we found the following disease-causing mutations: *parkin* (exonic duplications and deletions in five probands; 40bp homozygous deletion in one proband); *PINK1* (homozygous Y258X in one proband); *LRRK2* (G2019S in five probands and R1441C in one proband) and *SNCA* (whole-gene triplication in one proband). These findings have important implications for these individuals and their family members. Identification of pre-symptomatic mutation carriers in these families provides a unique opportunity for recruitment of these individuals for neuroprotective clinical trials and longitudinal studies in order to identify biomarkers of neurodegeneration.

These mutations were mainly found in Caucasian individuals as indicated in Figure 9.1. The parkin gene is predominantly involved in early-onset forms of PD and the higher prevalence of parkin mutations found in the present study might be attributed to the fact that the study group consisted mostly of individuals with early-onset PD. Higher mutation frequencies in the known PD genes have been reported for other populations and we concluded that these genes are not a

common cause of the disease in the South African population. Due to this finding, a bioinformatic approach was employed in order to identify novel PD candidate genes for mutation screening. The *MAPT* and *SNCAIP* genes were selected by this approach and a number of novel sequence variants of unknown pathogenic significance were found in these genes. The known and putative mutations identified in the present study are shown in Table 9.1.

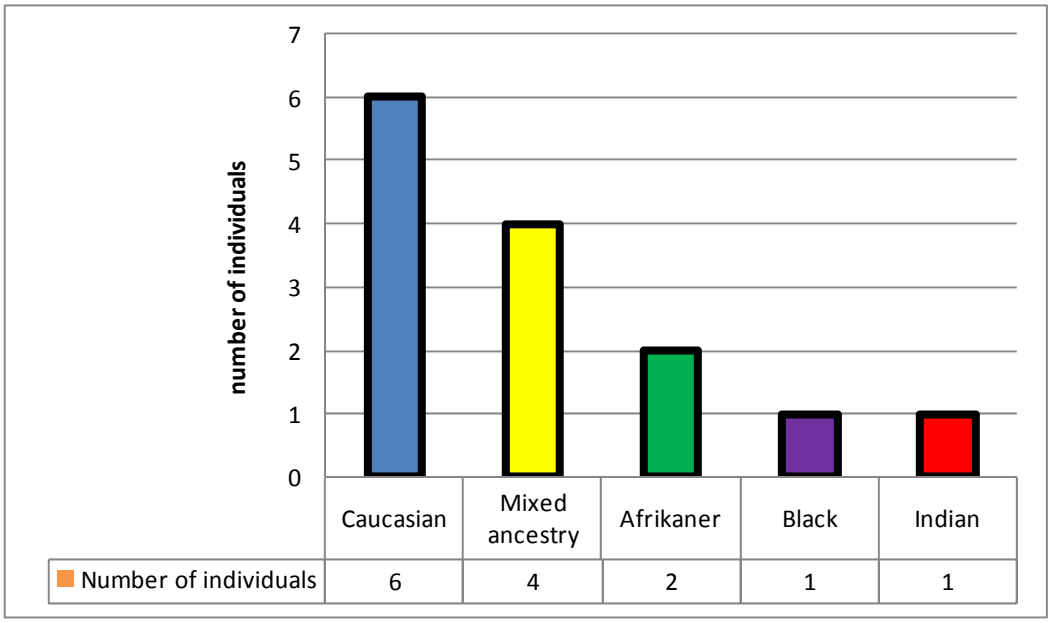


Figure 9.1 Number of individuals from the different South African sub-population groups in which pathogenic mutations were found.

Table 9.1 Known and putative mutations identified in the present study

Gene	Variant [a]	Number of patients/ Ethnicity [b]	AAO (yrs)	Family history	Frequency in control chromosomes	Previously reported/ reference
Autosomal recessive genes						
Parkin	**Duplication exon 2 + deletion exon 9**	1/ Black (37.12)	45	+	ND	(Madegowda *et al.* 2005; Simon-Sanchez *et al.* 2008)
	Duplication exon 2	1/ Afrikaner (55.54)	56	-	ND	(Simon-Sanchez *et al.* 2008)
	Duplication exons 2+3	1/ Mixed ancestry (68.22)	50	-	ND	(Periquet *et al.* 2003)
	Deletion exons 3+4 (homozygous)	1/ Mixed ancestry (53.44)	27	-	ND	(Hattori *et al.* 1998))
	Deletion exon 4 (homozygous)	1/ Afrikaner (56.45)	27	+	ND	(Kitada *et al.* 1998)
	40bp deletion in exon 3 (T113fsX163) (homozygous)	1/ Caucasian (77.60)	25	+	ND	(Farrer *et al.* 2001)
	H200Q	1/ Caucasian (65.79)	55	+	0/106	novel
	D280N	1/ Indian (68.10)	37	+	0/118	(Lucking *et al.* 2000; Oliveira *et al.* 2003)
	E310D	1/ Afrikaner (51.70)	42	-	0/110	novel

	R402C	1/ Afrikaner (42.06)	37	-	0/100	(Kay *et al.* 2007)
PINK1	**Y258X (homozygous)**	1/ Indian (68.10)	37	+	0/130	(Tan *et al.* 2006)
	P305A	1/ Black (42.03)	30	+	2/108 (1.9%)	novel
	E476K	1/ Mixed ancestry (63.69)	48	-	2/144 (1.4%)	(Valente *et al.* 2004)
DJ-1	g.-6_+10del	1/ Mixed ancestry (50.31)	38	-	2/148 (0.7%)	novel
Autosomal dominant genes						
LRRK2	**R1441C**	1/ Mixed ancestry (85.57)	62	-	ND	(Mata *et al.* 2005)
	G2019S	1/ Mixed ancestry (60.47)	47	+	0/158	(Kachergus *et al.* 2005; Lesage *et al.* 2009)
		1/ Caucasian (Russian-Ashkenazi Jewish) (68.06)	42	+		
		1/ Caucasian (Polish Jewish) (81.64)	70	-		
		1/ Caucasian (Jewish) (82.47)	63	-		
	G2019S (homozygous)	1/ Caucasian (Russian) (84.25)	58	+		
	E899D	1/ Black (43.59)	51	+	ND	novel
	R924H	1/ Afrikaner (38.36)	44	-	ND	novel
	I997N	1/ Caucasian (67.66)	65	-	ND	novel
	S1228T	1/ Caucasian (67.31)	50	+	ND	(Berg *et al.* 2005)
SNCA	**Triplication of whole gene**	1/ Caucasian (French-Italian) (42.35)	46	+	ND	(Singleton *et al.* 2003; Farrer *et al.* 2004))
Bioinformatically selected genes						
MAPT	A91V	1/ mixed ancestry (51.30)	32	-	0/124	novel
	V635I	1/ Black (60.43)	60	-	1/140 (0.7%)	novel
SNCAIP	T383N	1/ Black (52.23)	50	+	0/124	novel
	R606Q	1/ Caucasian (67.92)	70	-	0/124	novel
	E709Q	1/ Caucasian (83.15)	64	-	2/124 (1.6%)	(Marx *et al.* 2003)
	N906H	1/ Caucasian (82.92)	64	-	0/124	novel

[a] all variants are heterozygous unless stated otherwise.
[b] patient sample numbers are shown in brackets.
AAO: age at onset; ND: not determined.
Mutations with known pathogenic significance are shown in bold and shaded boxes indicate novel sequence variants.

9.1 Summary of main research findings

9.1.1 HRM screening of the *parkin* gene

More PD-causing mutations have been found in the *parkin* gene than in any other gene (Parkinson's disease Mutation Database, http://grenada.lumc.nl/LOVD2/TPI/home.php). For this reason mutation screening of all 12 exons of *parkin* was conducted in South African PD patients. The HRM technique was used and we showed that this method was effective in detecting sequence variants and that it was less time- and labor-intensive than the single strand conformation polymorphism (SSCP) technique. In our study, four missense sequence variants were identified (H200Q, D280N, E310D and R402C) but currently it is not known whether these are rare non-pathogenic polymorphisms. Two genomic rearrangement mutations (a homozygous deletion of exon 4 and a homozygous deletion of exons 3 and 4) were found. In addition, *parkin* was found to be highly polymorphic in the South African population. Our study revealed that the S167N, M192L and R334C variants, which had previously been reported to be mutations, were in fact polymorphisms due to their high prevalence in control chromosomes.

The low frequency of *parkin* mutations found were unexpected given that our patients had predominantly early-onset PD. *Parkin* mutations were also not found in one patient with juvenile onset PD, with an AAO of 17 years. From our investigation it became clear that point mutations or homozygous exon deletions in *parkin* are not a major cause of PD in the South African population.

9.1.2 Analysis of the *LRRK2* gene

The initial part of this study on the *LRRK2* gene focused on the G2019S mutation since this is the most commonly occurring mutation worldwide for PD (Wider *et al.* 2010). The HRM method was used to screen 205 patients and proved efficient in detecting the mutation. The frequency of G2019S was found to be low in this population compared to findings from other countries. It was detected at a frequency of 2.8% in patients with familial PD and 1.5% in apparently sporadic patients, and was absent in Afrikaner and Black patients. These low frequencies may indicate that genetic screening for this mutation is possibly less important in the South African population. Two of the G2019S-positive patients have late AAO PD with no apparent family history which concurs with other studies, and indicates that the familial and idiopathic late-onset forms of PD may share a common genetic basis.

The G2019S mutation has previously been shown to be present on only three haplotypes, referred to as haplotypes 1, 2 and 3 (Lesage *et al.* 2010). Haplotype studies revealed that the five South African G2019S-positive patients (four are Caucasian and one is of mixed ancestry) have the same disease-associated haplotype (haplotype 1) which has been found in a number of families worldwide of diverse origins including Ashkenazi Jews, North African Arabs, Caucasian Europeans and North Americans (Lesage *et al.* 2010). The one individual who had a homozygous G2019S was found to be homozygous for haplotype-1.

The second part of this study on *LRRK2* involved direct sequencing of nine exons in 195 patients, which had been selected according to AAO >40 years and autosomal dominant inheritance patterns. Using this approach, another one of the known pathogenic mutations, R1441C, was identified in only one patient. In addition, a previously identified putative pathogenic mutation, S1228T (Berg *et al.* 2005; Gandhi *et al.* 2009), was found in a patient. PolyPhen analysis however predicts this variant to be benign. Also, three novel missense sequence variants were found (E899D, R924H, I997N) which occur in evolutionary conserved regions of the gene. PolyPhen analysis predicts the E899D variant to be benign and the R924H variant to be possibly damaging. Both these variants occur in between the ANK (ankyrin) and LRR (leucine rich repeat) domains. The I997N variant, located in the LRR domain, is predicted by PolyPhen analysis to be probably damaging. A number of synonymous and intronic variants were also identified. Future work will be to assess the prevalence of the missense variants in ethnic matched control individuals and to determine their significance by means of functional studies.

9.1.3 HRM screening of the *PINK1* gene

After *parkin*, mutations in *PINK1* have been shown to be the second most common cause of PD (Klein & Lohmann-Hedrich 2007). For this reason all 8 exons of the *PINK1* gene were screened for pathogenic mutations. Only one mutation, homozygous Y258X, was found in three affected members of an Indian family. This mutation causes a significant truncation of the protein, which results in loss of the functional serine/threonine kinase domain.

Two missense variants (P305A and E476K) were also identified in *PINK1* as well as 13 polymorphisms. The P305A variant was present in a Black Xhosa individual who had typical clinical features of PD, which is in contrast to the atypical features that have been reported for

Black PD patients. Upon further investigation of the frequency of the P305A and E476K variants in ethnic matched control chromosomes, their location in evolutionary conserved regions and PolyPhen analysis, the two variants were found to be most likely non-pathogenic polymorphisms.

PINK1 was found to be polymorphic in the South African population, and the IVS1-7A>G variant was found at a frequency of 30%. Our preliminary findings suggest that IVS1-7A>G might be associated with a decreased AAO in our study participants ($P < 0.001$; t-test). This variant has been shown to decrease the AAO in patients exposed to various environmental risk factors (Godeiro, Jr. *et al.* 2010); however, their finding could not be verified in the present study due to the lack of comprehensive data on environmental exposures for all our study participants.

The very low frequency of mutations detected in the present study correlates with recent studies in which no mutations were detected (Bras *et al.* 2008; Godeiro-Junior *et al.* 2009; Zhang *et al.* 2010) or only a few heterozygous variants of unknown pathogenic significance were found (Mellick *et al.* 2009; Biswas *et al.* 2010). Earlier studies however reported higher frequencies which might indicate that the contribution of this gene varies according to ethnicity or that the frequency of mutations is not as common as was previously thought.

9.1.4 HRM screening of the *SNCA* gene

SNCA has been found to be a relatively rare cause of PD worldwide (Wider *et al.* 2010). To date, only three missense mutations (A30P, E46K and A53T) have been identified. As these mutations are found in exons 2 and 3, these two exons were subjected to mutation screening in the present study. No pathogenic missense mutations or novel sequence variants were identified.

9.1.5 Analysis of exon dosage/genomic rearrangements

Exon dosage/genomic rearrangements are a common type of mutation found in PD patients (Hedrich *et al.* 2001; Hedrich *et al.* 2002). This type of mutation can not be detected by traditional mutation screening methods such as DNA sequencing. The MLPA method was used in the present study and was found to be effective in detecting a variety of genomic rearrangements as well as a specific point mutation (Y258X in *PINK1*). These findings revealed the importance of including exon dosage studies as part of a PD mutation screening strategy.

The *parkin* gene presented with the highest proportion of genomic rearrangements compared to the other known PD genes (Table 9.1). As our study group was mostly early- onset PD patients, it was proposed that the higher frequency of *parkin* mutations detected was due to the fact that this gene is predominantly involved in early-onset forms of the disease. The frequency of *parkin* mutations (6.8%) detected in the present study was however much lower than that reported in many previous studies in which up to 50% of patients were found to have exonic rearrangements (Hattori *et al.* 1998; Hedrich *et al.* 2004). Notably, no parkin exonic rearrangements were found in the patient with an AAO of 17 years.

Our study reported a family with a *SNCA* triplication mutation, which is the 4th family worldwide to have this specific mutation. *SNCA* triplication mutations are thought to lead to a more severe form of early-onset PD with autonomic dysfunction and subsequent dementia (Singleton *et al.* 2003; Farrer *et al.* 2004). Our study identified mutation-positive presymptomatic members of the family who can now be offered the option of genetic testing with appropriate genetic counseling.

Another important finding from this study was that a common polymorphism in the South African population (M192L in *parkin*; frequency of 6.7% in controls) disrupts the ligation of the probes in kit P051 resulting in false positives (deletions) in exon 5. This highlights the importance of verifying all MLPA positive results with an independent method.

No genomic rearrangements were found in four of the major PD genes (*LRRK2, PINK1, DJ-1,* and *ATP13A2*) or in the *UCHL1, LPA*, *TNFRSF9, CAV2, CAV1, GCH1* genes in our study participants.

9.1.6 HRM screening of the *MAPT* and *SNCAIP* genes

Since we have shown that the known PD genes are not a common cause of the disease in the South African population we speculated that novel genes might be involved. A bioinformatics approach was used to select novel PD candidate genes for further study. The CAESAR bioinformatic program (Gaulton *et al.* 2007) was used and it produced a ranked list of hundreds of genes containing known PD-causing genes as well as novel candidates. The *MAPT* and *SNCAIP* genes, which encode tau and synphilin-1, respectively, were selected from the list of ten highest scoring genes for mutation screening based on current literature showing potential involvement in PD pathogenesis; GWAS have recently revealed *MAPT* to be a risk factor for idiopathic PD (Simon-Sanchez *et al.* 2009; Edwards *et al.* 2010) and synphilin-1 has been shown

to interact with α-synuclein, parkin and LRRK2 (Engelender *et al.* 1999; Chung *et al.* 2001). Mutation screening identified novel sequence variants in both genes that may have potential functional implications. Further functional studies are needed to determine the possible pathogenic significance of these sequence variants.

In the present study, the *ATP13A2* gene was not included in our mutation screening experiments since the phenotype of *ATP13A2*-linked individuals is juvenile-onset atypical parkinsonism with pyramidal tract dysfunction, supranuclear gaze paresis, dementia, and cognitive dysfunction (Di Fonzo *et al.* 2007) and our study participants did not exhibit these features.

9.1.7 Functional study of the *DJ-1* 5'UTR variant

For all novel sequence variants identified, functional studies are needed to assess their functional significance. For variants found in the 5'UTR regions of genes, luciferase reporter assays can be used to determine if they have an affect on the transcription or translation efficiency. In the present study, SSCP analysis of the *DJ-1* gene in 30 study participants had revealed a novel 16bp deletion variant (g.-6_+10del) in the promoter region and the functional significance of this variant was investigated by means of a Dual-luciferase reporter assay. The variant was found to span the transcription start site and was situated 93bp downstream of a Sp1 transcription regulatory sequence which is essential for *DJ-1* promoter activity (Taira *et al.* 2001). It was hypothesized that the variant might influence the level of *DJ-1* expression, the transcript stability or the translation efficiency. The variant was found to reduce luciferase activity in both HEK293 and neuroblastoma M17 cells, with the M17 cells exhibiting a higher reduction in activity. Induced oxidative stress conditions led to a dose-dependent moderate up-regulation in the promoter activity with the variant retaining its lower activity compared to the wild-type *DJ-1*. It was therefore proposed that the 16bp sequence might be important in transcriptional regulation of the *DJ-1* gene.

In addition, bioinformatic analysis found putative binding sites for three transcription factors (AhR, ARNT and HIF-1) within the 16bp deletion sequence and we suggested that the binding of these factors might be influenced by the variant. These findings are important since AhR and ARNT control a variety of developmental and physiological events including metabolism of toxins and responses to hypoxia (Wilson & Safe 1998; Pocar *et al.* 2005). AhR and ARNT are known to dimerize to form an active transcription factor complex that binds to Xenobiotic Response Elements (XRE) in promoter regions and alters expression of genes involved in

metabolism of endogenous substances, environmental chemicals and various drugs, such as the genes belonging to the Cytochrome P450 family (Rivera *et al.* 2007). Further studies are needed to verify that *DJ-1*'s expression is regulated by the AhR, ARNT and HIF-1 transcription factors. Furthermore, the possible importance of the 16bp sequence in higher order mammals, and therefore an evolutionary retained functionality, was suggested due to our finding that the deletion sequence is present in a highly conserved region between man and other primates.

9.2 Limitations of the study

A limitation of the experimental strategy used in the present study is not using direct sequencing to screen for pathogenic mutations in all study participants. However, due to cost-constraints in a low-resource setting like ours, sequencing was not feasible and therefore HRM was used instead.

9.2.1 Limitations of HRM

HRM has limitations which could results in both false positive and false negatives results. For this technique, the template DNA has to be of high purity and the PCR products should not contain contaminants such as primer dimers. Poor quality template DNA can result in spurious melt curves which are difficult to analyze. Primer dimers produce an additional melt curve which interferes with curve normalization and analysis of results. The efficacy of HRM also depends on the instruments and the type of intercalating dye used (Wittwer 2009). Different dyes are available (LCGreen® Plus, SYTO®9, EvaGreen® and SYBR Green I) and the type of analysis conducted determines which one should be used. PCR products should typically not exceed 250 base pairs and should not contain more than one melting domain as this will lead to poor quality melt curves. In our experience, HRM has been found to be less effective in detecting small insertions and deletions, and homozygous substitutions. Furthermore, samples with similar melt curves do not necessarily indicate the presence of the same variant and the results always need to be verified by sequencing.

9.2.2 Limitations of MLPA

A limitation of the MLPA technique is that it is very sensitive to contaminants such as high salt concentrations, which results in aberrant peak profiles. The initial template DNA has to be at the same concentration for all samples, must be of good quality and the same DNA isolation method needs to be used across all samples screened. The MLPA method can also detect point mutations which occur at the ligation site of two adjacent probes. Therefore, polymorphisms that occur at

this position, e.g. M192L in *parkin*, as shown in the present study (Chapter 3) will produce false positive results. Due to this problem all MLPA positive results need to be verified by sequencing and an alternative exon dosage detection method, thereby introducing additional reagent and labor costs.

9.2.3 Limitations of the study group and controls

A limitation of the study group was the lack of information regarding family history of the disorder for all the study participants. The inheritance pattern of PD was therefore not known for a large proportion of the study group which made it difficult to subdivide the patients into autosomal recessive and dominant categories for mutation screening purposes. Another limitation was the fact that the study group consisted predominantly of families with one or two affected members. For larger families with more affected individuals it would have been possible to determine whether sequence variants cosegregated with the disease in a family, in order to establish the pathogenicity of novel variants. Also, it would have been possible to conduct haplotype analysis studies. Furthermore, a major short-coming was the limited number of Black PD patients. Despite intensive efforts, no additional patients could be recruited due to a variety of reasons including the lack of contact details, death of patients and lack of willingness of family members to participate in the research. Ascertainment problems are also a factor as Black PD patients who live in rural areas often do not present to clinicians or neurologists. A limitation of the control individuals was that they had not been clinically assessed for signs of PD.

9.3 Future work

The present study identified a number of novel sequence variants in our study participants (Table 9.1) which should be investigated using functional studies in order to determine their possible involvement in PD pathogenesis. For these studies, mammalian expression constructs encoding wild-type and specific sequence variants for the *parkin, PINK1, LRRK2, MAPT,* and *SNCAIP* genes could be generated for the following functional experiments.

parkin

In vitro ubiquitylation assays and immunobloting are techniques which could be used in order to assess whether the parkin sequence variants negatively affect the ubiquitylation function of the protein. In addition, the possibility of the sequence variants to cause spontaneous aggregation of

the protein, as has been shown for some *parkin* mutations (Gu *et al.* 2003) could be assessed by means of immunohistochemistry and pulse-chase experiments.

LRRK2

LRRK2 mutations have been shown to result in reduced GTPase activity as well as an increase in kinase activity and autophosphorylation (Aasly *et al.* 2010). In order to investigate the enzymatic properties of the novel sequence variants identified in *LRRK2*, *in vitro* kinase and GTP-binding assays could be conducted. *LRRK2* mutations have also been associated with greater rate of apoptosis under oxidative stress conditions (Tan *et al.* 2007). Whether the novel variants identified in the study leads to the same consequences could be determined by performing an apoptosis assay and confocal microscopy.

PINK1

The fluorescence-activated cell sorting mitochondrial membrane potential assay could be used to determine whether the *PINK1* sequence variants leads to increased susceptibility of cells to cellular stress as has been shown for other mutations in *PINK1* (Abou-Sleiman *et al.* 2006). Also, the effect of the P305A variant on mitochondrial morphology could be assessed by fluorescence microscopy.

MAPT

Immunofluorescence confocal microscopy could be employed in order to determine whether the *MAPT* variants result in abnormalities such as failure of microtubules to assemble, reduced length of microtubules, and whether the microtubules occur as fragments in the cytoplasm. Western blotting analysis could be performed to assess the biochemical profiles of the mutant proteins (Sahara *et al.* 2000). Quantitative PCR could be used in order to assess the ratio of tau containing either four or three repeats of the microtubule binding domains as this has been shown by other studies to be increased in PD patients (Tobin *et al.* 2008).

SNCAIP

In order to determine the effect of *SNCAIP* variants, cells could be treated with specific toxins such as proteasome inhibitors, in order to assess the formation of inclusion bodies. It has been suggested that mutations in *SNCAIP* that reduce the ability of cells to form intracellular inclusions may sensitize neurons to cellular stress (Marx *et al.* 2003). In addition, the ability of a mutant synphilin-1 protein to interact with α-synuclein, parkin and LRRK2 could be assessed by means of a co-immunoprecipitation assay.

9.3.1 Bioinformatic analysis to predict functionality of novel variants

Bioinformatic analysis in collaboration with researchers at the South African National Bioinformatics Institute (SANBI) could be conducted to investigate whether the sequence variants found in the present study result in significant protein changes (such as alterations in structure, charge, absence/presence of disulphide bridges and position of hydrogen-bonds in side chains), which could negatively affect the molecular pathways in which the proteins are involved.

SIFT (Sorting Intolerant From Tolerant, http://sift.jcvi.org/), PolyPhen (Polymorphism Phenotyping, http://genetics.bwh.harvard.edu/pph/), SNPs3D (Single Nucleotide Polymorphisms 3D, http://www.snps3d.org/) and SNPeffect (Single Nucleotide Polymorphism effect, http://snpeffect.vib.be/) are some of the computational tools that can be used for this analysis.

9.3.2 Mutation screening of PD candidate genes

The study participants that had not been screened for mutations in the *parkin*, *PINK1*, *LRRK2* and *SNCA* genes in the present study should be investigated (this work is currently ongoing in our laboratory). Furthermore, other PD candidate genes identified by the CAESAR program, namely *BCL2L1* (bcl-2-like protein 1), *SNCB* (β-synuclein) and *TH* (tyrosine hydroxylase), should also be screened for mutations in the South African PD population. Other suitable candidates are the genes involved in spinocerebellar ataxia types 2, 3, 6 and 17 that have been found to be implicated in PD patients (Klein *et al.* 2009). Another exciting prospect would be to investigate genes involved in maintenance and function of mitochondria, and this would include both mitochondrial-encoded and nuclear-encoded genes.

9.4 PD research in sub-Saharan Africa

There is a paucity of data on PD in African populations. A literature search conducted by Okubadejo and colleagues using search terms ‘Parkinson’s disease and Africa’ and related terms revealed that there were publications from 13 countries over the period 1944 to 2004 (Okubadejo *et al.* 2006). The bulk of this literature was based on clinical descriptions of patients and genetic studies had only been conducted on populations from the Northern African countries, Tunisia and Algeria. The lack of genetic studies in sub-Saharan African countries was thought to be due to the scarcity of resources including finances, manpower and equipment.

A literature search conducted a few years later in 2007, this time focusing specifically on genetic studies using search terms 'Parkinson's disease and genetics and Africa' and related terms found a very limited number (12) of publications (Okubadejo *et al.* 2008). Eleven of the 12 studies were on Northern African populations (Morocco, Algeria and Tunisia) and only one study was on a sub-Saharan African country. This study examined the autosomal dominant and recessive inheritance patterns of seven Zambian families with familial PD (Atadzhanov *et al.* 2005).

An interesting finding from the North African populations is that the LRRK2 G2019S mutation appears to be a common cause of PD (41%) in these populations compared to Europe and North America (Lesage *et al.* 2006). The only molecular genetic studies conducted on a sub-Saharan population are the work reported in this dissertation. Further genetic studies into the diverse populations of sub-Saharan Africa are warranted as they may reveal insights into novel mechanisms, causes and patho-physiology of PD.

9.5 Next-generation sequencing technologies

Exome sequencing is a new technology that is currently being employed to identify disease-causing genes and rare variants. This approach was first used to successfully identify the genetic cause of Miller syndrome, a rare Mendelian disorder (Ng *et al.* 2010). It involves sequencing the complete exome (all protein-coding sequences in the genome) as well as splice donor and acceptor sites in disease affected individuals. Bioinformatics approaches are used to exclude known and likely benign variants by screening dbSNP or HapMap databases and mutation databases. Also the possible effect of the variants on the normal function of the protein is investigated (Ng *et al.* 2010; Kobelka 2010). This results in a list of possible pathogenic variants based on gene function, evolutionary conservation and likely mutational impact. Exome sequencing requires only 5% as much sequencing as whole-genome sequencing and is therefore more cost effective (Ng *et al.* 2010). The probability of finding the disease causing gene and variant is higher when sequencing the exome than when sequencing the genome. A limitation of this method is that it is not able to detect cytogenetic abnormalities, genomic copy number variants, exon rearrangements, non-coding sequence variants, and inherited epigenetic modifications (Kobelka 2010).

Whole-genome sequencing is currently still too expensive to be used as a routine mutation screening technique in most laboratories. However, next-generation sequencing technologies which involves various platforms including that of Illumina Genome Analyzer and Applied

Biosystem's SOLiD has revolutionized the field. It is anticipated that this will eventually lead to a dramatic reduction of costs to only ~$1,000 per genome (Service 2006). Next-generation sequencing technologies comprise different approaches for template preparation, sequencing and imaging, genome alignment and assembly methods.

According to a recent survey conducted in 2010, the Illumina platform is the most popular amongst laboratories worldwide (http://www.genomeweb.com/sequencing/survey-illumina-solid-and-454-gain-ground-research-labs-most-users-mull-addition). It produces 18 gigabases of sequence data per run. The average read length is about 72 base pairs for single reads and 80 base pairs for paired reads. The SOLiD platform produces 30 gigabases of sequence data per run. The average read length is 50 base pairs for single reads and 25, 35, or 50 base pairs for paired reads (Hong & Oh 2010). Problems that users of the Illumina and SOLiD platforms have reported include sample and complex library preparations, software and mechanical complications. According to the survey, most of the users of next-generation sequencing technologies are universities or non-profit research institutes based in North America and Europe, and the top three applications include mRNA sequencing, whole-genome de novo sequencing and small RNA sequencing.

These exciting new technologies could be used to make important discoveries about genetic causes of PD in South African populations. Given that Black African populations have been under-studied, whole- genome or exome sequencing data from these populations would be of interest for comparison with that of Caucasian and Asian genomes. Ultimately, genome sequencing of South African PD patients would facilitate the 'holy grail' of personalized medicine (Hong & Oh 2010) but would also be beneficial for the generation of new hypotheses and development of novel therapeutics for PD.

9.6 Concluding remarks

Investigating the genetic basis of predominantly early-onset PD in the South African population has revealed that the *parkin*, *PINK1*, *LRRK2* and *SNCA* genes are not major causes of PD in this particular population. Pathogenic mutations were found in only 14 patients and of these, the majority (six) are Caucasian. It is important that the PD-causing mutations in diverse populations worldwide are identified, and it is speculated that these mutations may occur in as yet unidentified PD genes (Okubadejo *et al.* 2008; Bardien *et al.* 2009). In this regard, bioinformatic *in-silico* approaches are an important tool for the identification of disease-causing genes,

especially in complex diseases such as PD, as they can be used to systematically and methodically search through vast amounts of information and are cost effective. Also bioinformatic approaches are not prone to many of the inherent biases invariably introduced by researchers.

There is currently very limited data available on the genetic basis of PD in sub-Saharan Africa. The usefulness for these populations of the therapeutic modalities or predictive genetic tests, developed by the first world countries, will only become apparent once we know the genetic basis of PD in Africa's populations (Okubadejo *et al.* 2008). This knowledge may also have an impact on our understanding of disease pathogenesis and may provide important insights into the interactions between genetic and lifestyle factors.

From the functions of the known PD-causing genes it has become clear that there is considerable cross-talk between the different pathways proposed to be involved in the pathogenesis. These processes include abnormal handling of misfolded proteins by the ubiquitin proteasome and autophagy-lysosomal systems (*parkin*, *ATP13A2*, *SNCA*), mitochondrial dysfunction (*PINK1*, *DJ-1*, *parkin*), and aberrant kinase signaling (*PINK1*, *LRRK2*). It has also recently been shown that some of the apparently unrelated neurodegenerative disorders may share overlapping genetic causes. Recent GWAS studies have found that in the European population, multiple polymorphisms within and near *MAPT* are associated with idiopathic late onset PD (Simon-Sanchez *et al.* 2009). *MAPT* has also been implicated in Alzheimer's disease in which β-amyloid aggregates to form plaques and the neurotoxicity is thought to be mediated, in part by tau (encoded by *MAPT*) which can also aggregate to form neurofibrillary tangles (Shulman & De Jager 2009). In addition, patients with pathogenic triplet expansions in the genes involved in spinocerebellar ataxia types 2, 3, 6 and 17 have been found in patients that present with clinical features of PD (Klein *et al.* 2009). Taken together, these findings imply that understanding the biological pathways or mechanisms in PD will shed light on the disease mechanisms underlying other neurodegenerative disorders.

Recently, mitochondrial dysfunction has been conclusively implicated as a major role player in PD (Zheng *et al.* 2010). Zheng and colleagues (2010) performed gene expression profiling on 322 brain and 88 blood samples, and found that the expression of genes involved in mitochondrial electron transport, mitochondrial biogenesis, glucose utilization, and glucose sensing are reduced in PD patients. Specifically, genes that are expressed in response to the transcriptional coactivator PGC-1α (peroxisome proliferator-activated receptor-gamma

coactivator -1, alpha) were under-expressed. PGC-1α over-expression resulted in increased transcription of subunits of the mitochondrial respiratory chain and importantly, could block dopaminergic neuron loss in cell culture models of PD. PGC-1α is necessary for induction of numerous ROS-detoxifying enzymes thereby serving as a 'master regulator' of ROS metabolism in the cell (St-Pierre *et al.* 2006). The findings by Zheng and colleagues are a major breakthrough in PD research as it reveals that PGC-1α-activating drugs may be potential therapeutic targets for early intervention. It is still unclear though whether the PGC-1α- pathway is specifically suppressed in PD or if the suppression is a consequence of more widespread mitochondrial damage.

An improved understanding of the complex mechanisms and role players involved in neurodegeneration in PD, such as PGC-1α and other agents that block mitochondrial dysfunction including co-enzyme Q and creatine (Thomas & Beal 2010), brings us one step closer to more effective therapy and possibly a cure, which will offer a beacon of hope for the thousands of patients suffering from this debilitating disorder.

9.7 Reference list

Aasly, J. O., C. Vilarino-Guell, J. C. Dachsel, P. J. Webber, A. B. West *et al.* 2010 Novel pathogenic LRRK2 p.Asn1437His substitution in familial Parkinson's disease. *Mov Disord.* **25**: 2156-2163.

Abou-Sleiman, P. M., M. M. Muqit, N. Q. McDonald, Y. X. Yang, S. Gandhi *et al.* 2006 A heterozygous effect for PINK1 mutations in Parkinson's disease? *Ann.Neurol.* **60**: 414-419.

Atadzhanov, M., A. Zumla, and P. Mwaba, 2005 Study of familial Parkinson's disease in Russia, Uzbekistan, and Zambia. *Postgrad.Med.J.* **81**: 117-121.

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.Relat Disord.* **15**: 116-121.

Berg, D., K. Schweitzer, P. Leitner, A. Zimprich, P. Lichtner *et al.* 2005 Type and frequency of mutations in the LRRK2 gene in familial and sporadic Parkinson's disease*. *Brain* **128**: 3000-3011.

Biswas, A., T. Sadhukhan, S. Majumder, A. K. Misra, S. K. Das *et al.* 2010 Evaluation of PINK1 variants in Indian Parkinson's disease patients. *Parkinsonism.Relat Disord.* **16**: 167-171.

Bras, J., R. Guerreiro, M. Ribeiro, A. Morgadinho, C. Januario *et al.* 2008 Analysis of Parkinson disease patients from Portugal for mutations in SNCA, PRKN, PINK1 and LRRK2. *BMC.Neurol.* **8**: 1.

Chaudhuri, K. R., M. T. Hu, and D. J. Brooks, 2000 Atypical parkinsonism in Afro-Caribbean and Indian origin immigrants to the UK. *Mov Disord.* **15**: 18-23.

Chung, K. K., Y. Zhang, K. L. Lim, Y. Tanaka, H. Huang *et al.* 2001 Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy-body formation in Parkinson disease. *Nat.Med.* **7**: 1144-1150.

Di Fonzo, A., H. F. Chien, M. Socal, S. Giraudo, C. Tassorelli *et al.* 2007 ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease. *Neurology* **68**: 1557-1562.

Edwards, T. L., W. K. Scott, C. Almonte, A. Burt, E. H. Powell *et al.* 2010 Genome-wide association study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. *Ann.Hum.Genet.* **74**: 97-109.

Engelender, S., Z. Kaminsky, X. Guo, A. H. Sharp, R. K. Amaravi *et al.* 1999 Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. *Nat.Genet.* **22**: 110-114.

Farrer, M., P. Chan, R. Chen, L. Tan, S. Lincoln *et al.* 2001 Lewy bodies and parkinsonism in families with parkin mutations. *Ann.Neurol.* **50**: 293-300.

Farrer, M., J. Kachergus, L. Forno, S. Lincoln, D. S. Wang *et al.* 2004 Comparison of kindreds with parkinsonism and alpha-synuclein genomic multiplications. *Ann.Neurol.* **55**: 174-179.

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. *J.Neurosci.Res.* **87**: 1283-1295.

Gaulton, K. J., K. L. Mohlke, and T. J. Vision, 2007 A computational system to select candidate genes for complex human traits. *Bioinformatics.* **23**: 1132-1140.

Godeiro, C., Jr., P. M. Aguiar, A. C. Felicio, O. G. Barsottini, S. M. Silva *et al.* 2010 PINK1 polymorphism IVS1-7 A-->G, exposure to environmental risk factors and anticipation of disease onset in Brazilian patients with early-onset Parkinson's Disease. *Neurosci.Lett.* **469**: 155-158.

Godeiro-Junior, C., P. M. de Carvalho-Aguiar, A. C. Felicio, O. G. Barsottini, S. M. Silva *et al.* 2009 PINK1 mutations in a Brazilian cohort of early-onset Parkinson's disease patients. *Mov Disord.* **24**: 1693-1696.

Gu, W. J., O. Corti, F. Araujo, C. Hampe, S. Jacquier *et al.* 2003 The C289G and C418R missense mutations cause rapid sequestration of human Parkin into insoluble aggregates. *Neurobiol.Dis.* **14**: 357-364.

Hattori, N., T. Kitada, H. Matsumine, S. Asakawa, Y. Yamamura *et al.* 1998 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. *Ann.Neurol.* **44**: 935-941.

Hedrich, K., C. Eskelson, B. Wilmot, K. Marder, J. Harris *et al.* 2004 Distribution, type, and origin of Parkin mutations: review and case studies. *Mov Disord.* **19**: 1146-1157.

Hedrich, K., M. Kann, A. J. Lanthaler, A. Dalski, C. Eskelson *et al.* 2001 The importance of gene dosage studies: mutational analysis of the parkin gene in early-onset parkinsonism. *Hum.Mol.Genet.* **10**: 1649-1656.

Hedrich, K., K. Marder, J. Harris, M. Kann, T. Lynch *et al.* 2002 Evaluation of 50 probands with early-onset Parkinson's disease for Parkin mutations. *Neurology* **58**: 1239-1246.

Hong, K. W., and B. Oh, 2010 Overview of personalized medicine in the disease genomic era. *BMB.Rep.* **43**: 643-648.

Jenkins, T., 1990 Medical genetics in South Africa. *J.Med.Genet.* **27**: 760-779.

Kachergus, J., I. F. Mata, M. Hulihan, J. P. Taylor, S. Lincoln *et al.* 2005 Identification of a novel LRRK2 mutation linked to autosomal dominant parkinsonism: evidence of a common founder across European populations. *Am.J.Hum.Genet.* **76**: 672-680.

Kay, D. M., D. Moran, L. Moses, P. Poorkaj, C. P. Zabetian *et al.* 2007 Heterozygous parkin point mutations are as common in control subjects as in Parkinson's patients. *Ann.Neurol.* **61**: 47-54.

Kitada, T., S. Asakawa, N. Hattori, H. Matsumine, Y. Yamamura *et al.* 1998 Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**: 605-608.

Klein, C., and K. Lohmann-Hedrich, 2007 Impact of recent genetic findings in Parkinson's disease. *Curr.Opin.Neurol.* **20**: 453-464.

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. *Mov Disord.* **24**: 2042-2058.

Kobelka, C. E., 2010 Exome sequencing: expanding the genetic testing toolbox. *Clin.Genet.* **78**: 132-134.

Lesage, S., C. Condroyer, A. Lannuzel, E. Lohmann, A. Troiano *et al.* 2009 Molecular analyses of the LRRK2 gene in European and North African autosomal dominant Parkinson's disease. *J.Med.Genet.* **46**: 458-464.

Lesage, S., A. Durr, M. Tazir, E. Lohmann, A. L. Leutenegger *et al.* 2006 LRRK2 G2019S as a cause of Parkinson's disease in North African Arabs. *N.Engl.J.Med.* **354**: 422-423.

Lesage, S., E. Patin, C. Condroyer, A. L. Leutenegger, E. Lohmann *et al.* 2010 Parkinson's disease-related LRRK2 G2019S mutation results from independent mutational events in humans. *Hum.Mol.Genet.* **19**: 1998-2004.

Lucking, C. B., A. Durr, V. Bonifati, J. Vaughan, M. G. De *et al.* 2000 Association between early-onset Parkinson's disease and mutations in the parkin gene. *N.Engl.J.Med.* **342**: 1560-1567.

Madegowda, R. H., A. Kishore, and A. Anand, 2005 Mutational screening of the parkin gene among South Indians with early onset Parkinson's disease. *J.Neurol.Neurosurg.Psychiatry* **76**: 1588-1590.

Marx, F. P., C. Holzmann, K. M. Strauss, L. Li, O. Eberhardt *et al.* 2003 Identification and functional characterization of a novel R621C mutation in the synphilin-1 gene in Parkinson's disease. *Hum.Mol.Genet.* **12**: 1223-1231.

Mata, I. F., J. M. Kachergus, J. P. Taylor, S. Lincoln, J. Aasly *et al.* 2005 Lrrk2 pathogenic substitutions in Parkinson's disease. *Neurogenetics.* **6**: 171-177.

Mellick, G. D., G. A. Siebert, M. Funayama, D. D. Buchanan, Y. Li *et al.* 2009 Screening PARK genes for mutations in early-onset Parkinson's disease patients from Queensland, Australia. *Parkinsonism.Relat Disord.* **15**: 105-109.

Ng, S. B., K. J. Buckingham, C. Lee, A. W. Bigham, H. K. Tabor *et al.* 2010 Exome sequencing identifies the cause of a mendelian disorder. *Nat.Genet.* **42**: 30-35.

Okubadejo, N., A. Britton, C. Crews, R. Akinyemi, J. Hardy *et al.* 2008 Analysis of Nigerians with apparently sporadic Parkinson disease for mutations in LRRK2, PRKN and ATXN3. *PLoS.One.* **3**: e3421.

Okubadejo, N. U., J. H. Bower, W. A. Rocca, and D. M. Maraganore, 2006 Parkinson's disease in Africa: A systematic review of epidemiologic and genetic studies. *Mov Disord.* **21**: 2150-2156.

Oliveira, S. A., W. K. Scott, E. R. Martin, M. A. Nance, R. L. Watts *et al.* 2003 Parkin mutations and susceptibility alleles in late-onset Parkinson's disease. *Ann.Neurol.* **53**: 624-629.

Periquet, M., M. Latouche, E. Lohmann, N. Rawal, M. G. De *et al.* 2003 Parkin mutations are frequent in patients with isolated early-onset parkinsonism. *Brain* **126**: 1271-1278.

Pocar, P., B. Fischer, T. Klonisch, and S. Hombach-Klonisch, 2005 Molecular interactions of the aryl hydrocarbon receptor and its biological and toxicological relevance for reproduction. *Reproduction.* **129**: 379-389.

Rivera, S. P., F. Wang, S. T. Saarikoski, R. T. Taylor, B. Chapman *et al.* 2007 A novel promoter element containing multiple overlapping xenobiotic and hypoxia response elements mediates induction of cytochrome P4502S1 by both dioxin and hypoxia. *J.Biol.Chem.* **282**: 10881-10893.

Sahara, N., T. Tomiyama, and H. Mori, 2000 Missense point mutations of tau to segregate with FTDP-17 exhibit site-specific effects on microtubule structure in COS cells: a novel action of R406W mutation. *J.Neurosci.Res.* **60**: 380-387.

Service, R. F., 2006 Gene sequencing. The race for the $1000 genome. *Science* **311**: 1544-1546.

Shulman, J. M., and P. L. De Jager, 2009 Evidence for a common pathway linking neurodegenerative diseases. *Nat.Genet.* **41**: 1261-1262.

Simon-Sanchez, J., S. Scholz, M. M. Matarin, H. C. Fung, D. Hernandez *et al.* 2008 Genomewide SNP assay reveals mutations underlying Parkinson disease. *Hum.Mutat.* **29**: 315-322.

Simon-Sanchez, J., C. Schulte, J. M. Bras, M. Sharma, J. R. Gibbs *et al.* 2009 Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nat.Genet.* **41**: 1308-1312.

Singleton, A. B., M. Farrer, J. Johnson, A. Singleton, S. Hague *et al.* 2003 alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**: 841.

St-Pierre, J., S. Drori, M. Uldry, J. M. Silvaggi, J. Rhee *et al.* 2006 Suppression of reactive oxygen species and neurodegeneration by the PGC-1 transcriptional coactivators. *Cell* **127**: 397-408.

Taira, T., K. Takahashi, R. Kitagawa, S. M. Iguchi-Ariga, and H. Ariga, 2001 Molecular cloning of human and mouse DJ-1 genes and identification of Sp1-dependent activation of the human DJ-1 promoter. *Gene* **263**: 285-292.

Tan, E. K., K. Yew, E. Chua, K. Puvan, H. Shen *et al.* 2006 PINK1 mutations in sporadic early-onset Parkinson's disease. *Mov Disord.* **21**: 789-793.

Tan, E. K., Y. Zhao, L. Skipper, M. G. Tan, F. A. Di *et al.* 2007 The LRRK2 Gly2385Arg variant is associated with Parkinson's disease: genetic and functional evidence. *Hum.Genet.* **120**: 857-863.

Thomas, B., and M. F. Beal, 2010 Mitochondrial therapies for Parkinson's disease. *Mov Disord.* **25 Suppl 1**: S155-S160.

Tobin, J. E., J. C. Latourelle, M. F. Lew, C. Klein, O. Suchowersky *et al.* 2008 Haplotypes and gene expression implicate the MAPT region for Parkinson disease: the GenePD Study. *Neurology* **71**: 28-34.

Valente, E. M., S. Salvi, T. Ialongo, R. Marongiu, A. E. Elia *et al.* 2004 PINK1 mutations are associated with sporadic early-onset parkinsonism. *Ann.Neurol.* **56**: 336-341.

Wider, C., T. Foroud, and Z. K. Wszolek, 2010 Clinical implications of gene discovery in Parkinson's disease and parkinsonism. *Mov Disord.* **25 Suppl 1**: S15-S20.

Wilson, C. L., and S. Safe, 1998 Mechanisms of ligand-induced aryl hydrocarbon receptor-mediated biochemical and toxic responses. *Toxicol.Pathol.* **26**: 657-671.

Wittwer, C. T., 2009 High-resolution DNA melting analysis: advancements and limitations. *Hum.Mutat.* **30**: 857-859.

Zhang, B. R., Z. X. Hu, X. Z. Yin, M. Cai, G. H. Zhao *et al.* 2010 Mutation analysis of parkin and PINK1 genes in early-onset Parkinson's disease in China. *Neurosci.Lett.* **477**: 19-22.

Zheng, B., Z. Liao, J. J. Locascio, K. A. Lesniak, S. S. Roderick *et al.* 2010 PGC-1{alpha}, A Potential Therapeutic Target for Early Intervention in Parkinson's Disease. *Sci.Transl.Med.* **2**: 52ra73.

URL reference list

GenomeWeb: [http://www.genomeweb.com/sequencing/survey-illumina-solid-and-454-gain-ground-research-labs-most-users-mull-addition] date accessed: 18/01/2011

Single Nucleotide Polymorphism effect (SNPeffect): [http://snpeffect.vib.be/] date accessed: 11/10/2010

Single Nucleotide Polymorphisms 3D (SNPs3D): [http://www.snps3d.org/] date accessed: 11/10/2010

Sorting Intolerant From Tolerant (SIFT): [http://sift.jcvi.org/] date accessed: 11/10/2010

Parkinson’s disease Mutation Database: [http://grenada.lumc.nl/LOVD2/TPI/home.php] date accessed: 11/10/2010

Polymorphism Phenotyping (PolyPhen): [http://genetics.bwh.harvard.edu/pph/] date accessed: 11/10/2010

Appendix

List of all SNPs and pathogenic variants found in

250 South African patients screened

Table 1 List of all SNPs and pathogenic variants found in 250 South African patients screened

Nr.	Lab Nr.	AAO	PARK2 (parkin)	PARK7 (DJ-1)	PARK8 (LRRK2)	PARK1/4 (SNCA)	PARK 6 (PINK1)	SNCAIP	MAPT
1	37.12	45	ex2 het DUP & ex9 het DEL						
2	37.45	38	het V380L, het D394N				het N521T		het P605P
3	38.37	40	hom 5'UTR -227 A>G						
4	38.36	44	het 5'UTR -258 T>G		het R924H				
5	38.23	40							
6	42.19	38	hom 5'UTR -258 T>G						
7	42.26	39					het N521T		
8	42.03	30	het C238C				het P305A, het A340T, het N521T		
9	42.04	64	het V380L				het N521T		
10	42.05	41							
11	42.06	37	het R402C,het 5'UTR -258 T>G, het V380L						het IVS4+9A>G, het A562A, het N590N, hom IVS11+40C>T
12	42.30	39							
13	42.35	40	het 5'UTR -258 T>G, het V380L			Triplication		ND	ND
14	42.56	43	het 5'UTR -258 T>G, het R402R						
15	43.59	51			het E899D				
16	56.45	27	ex4 hom DEL						het IVS4+9A>G
17	50.31	38	het 5'UTR -258 T>G; het Q34R' het IVS2 +10C>T; het IVS2 +25T>C, het S167N	g.-6_+10 het DEL					
18	51.30	32	het V380L						het A91V
19	50.86	41	het 5'UTR -258 T>G, het V380L						
20	50.03	17					hom IVS1-7A>G		
21	51.70	42	het E310D, het 5'UTR -227 A>G				hom IVS1-7A>G		
22	52.23	50						het T383N, het G390G	
23	53.44	27	ex3 & 4 hom DEL, het 5'UTR -227 A>G; het P34P; het IVS2 +25T>C				hom IVS1-7A>G		
24	54.73	69	het IVS7 -35 G>A, het IVS7 -68C>G						
25	54.74	63						het L638L	
26	55.41	77	het M192L, het R402R						
27	55.45	52	het IVS8 -21_-17del, het R334C				het 3'UTR+37A>T		
28	55.50	55							
29	54.89	26						het IVS6-23 T>C	
30	55.52	42	het L261L						
31	55.54	56	ex2 het DUP	ND			hom IVS1-7A>G, hom 3'UTR+37A>T		
32	55.65	40	het IVS2 +25T>C, hom IVS7 -35 G>A	ND	het IVS26-89T>A; het E1224E				
33	55.77	47		ND	het IVS26-89T>A				
34	55.81	72		ND	het P1542S				
35	59.91	29		ND					het A562A, het N590N
36	60.38	47	het 5'UTR -227 A>G; het S167N, het V380L	ND			hom IVS1-7A>G, het N521T, hom 3'UTR+37A>T		
37	60.39	55	het P34P; het IVS2 +25T>C; hom IVS3 -20 T>C, het V380L	ND	het IVS26-89T>A; het IVS26-33T>C				
38	60.43	60	het IVS7 -35 G>A, het IVS8 +43A>G, het IVS8 +48C>T, het D394N, het 3'UTR *103C>T	ND	het IVS20-65A>T				het V635I
39	60.45	71		ND					

Nr.	Lab Nr.	AAO	PARK2 (parkin)	PARK7 (DJ-1)	PARK8 (LRRK2)	PARK1/4 (SNCA)	PARK 6 (PINK1)	SNCAIP	MAPT
40	60.47	47		ND	het G2019S				
41	60.48	60		ND					
42	60.92	52		ND					
43	61.37	44		ND					
44	61.79	74	hom M192L, het C238C, het 3'UTR repeat expansion: WT 6CAs, 6CGs; variant 5CAs, 8CGs	ND	het S1721S				
45	61.80	50	het 3'UTR *16G>A	ND	het IVS26-89T>A			het S406S	
46	61.81	55	het M192L	ND	het IVS26-33T>C				
47	63.40	48		ND					het IVS3+18C>T
48	63.42	74		ND	het IVS26-89T>A				
49	63.43	35		ND			het A340T		
50	63.52	71		ND					
51	63.68	60	het D394N	ND			hom IVS1-7A>G		
52	63.69	48		ND	het E1224E		het E476K		
53	63.73	59	het P34P; het IVS2 +25T>C	ND					
54	64.85	38	hom IVS2 +25T>C	ND					
55	64.88	56		ND					
56	64.89	58	hom IVS3 -20 T>C, hom V380L	ND					hom IVS4+9A>G,hom A562A, hom N590N
57	65.25	59		ND					
58	65.26	50	het IVS2 +35G>A, , het D394N	ND					
59	65.65	69		ND					
60	65.68	50	het R334C	ND					
61	65.79	55	het H200Q, het D394N	ND					
62	67.31	50		ND	het S1228T				
63	67.32	67	hom IVS8 +48C>T	ND					
64	67.34	63		ND					
65	67.52	57		ND					
66	67.61	70		ND					
67	67.63	65		ND					hom IVS3+18C>T,hom IVS4+9A>G, hom A562A, hom N590N
68	67.64	49	het IVS8 +48C>T	ND					
69	67.65	61		ND					
70	67.66	65	het 3'UTR *94A>G	ND	het I997N				
71	67.67	64	het V380L	ND					het IVS11+40C>T
72	67.68	68		ND					
73	67.69	52		ND			het A340T		
74	67.70	58		ND					
75	67.71	71		ND					hom IVS3+18C>T,hom A562A, hom N590N
76	67.72	65		ND			hom IVS1-7A>G		
77	67.80	57	het IVS2 +35G>A; hom IVS3 -20 T>C	ND					
78	67.82	54		ND	het P1542S				
79	67.92	70		ND				het R606Q	
80	67.93	70		ND			het 3'UTR+37A>T		
81	67.97	68	het L63L	ND			hom L63L		
82	67.98	52		ND					
83	67.99	58		ND					
84	68.00	65		ND					
85	68.06	42		ND	het G2019S				
86	68.07	50		ND	het Q2089R			het L504L	
87	68.08	62		ND					

Nr.	Lab Nr.	AAO	PARK2 (parkin)	PARK7 (DJ-1)	PARK8 (LRRK2)	PARK1/4 (SNCA)	PARK 6 (PINK1)	SNCAIP	MAPT
88	68.10	37	het D280N	ND			hom Y258X		hom P605P
89	68.16	47		ND				het IVS6-23 T>C	
90	68.17	50		ND					
91	68.22	50	ex2 & 3 het DUP	ND			het IVS1-7A>G,het 3'UTR+37A>T		
92	68.27	55		ND					
93	68.37	60		ND					
94	69.33	74		ND			het A340T		
95	69.24	55		ND					
96	69.54	51		ND			hom IVS1-7A>G		
97	70.01	68		ND			het A340T, hom N521T		
98	70.03	55		ND					
99	70.05	45		ND	het R1514Q				
100	70.32	40		ND					
101	70.77	74		ND					
102	76.92	76		ND				ND	ND
103	72.68	41	het IVS2 +25T>C	ND			het A340T		
104	73.65	75		ND					
105	77.60	25	c.337-376 hom DEL	ND					
106	78.67	49		ND					
107	78.74	56	het G430D	ND	het IVS20-65A>T				
108	78.76	27	PARK2 ex4 het DEL, het G430D, het IVS8 +48C>T	ND					
109	78.82	55		ND					
110	78.84	50		ND					
111	78.95	58		ND					
112	78.97	53	het IVS2 +35G>A	ND				het L504L	
113	79.04	34		ND			het IVS1-7A>G,het 3'UTR+37A>T		
114	79.16			ND					
115	79.17	38	het M192L	ND					
116	79.29	40		ND					
117	79.35	48		ND					
118	79.91	46		ND					
119	79.99	39	het IVS8 +48C>T; het V380L	ND	het IVS27+7C>T		hom IVS1-7A>G		hom P605P
120	80.59	33	het IVS2 +25T>C; IVS2 +62G>A	ND		ND			
121	81.03	48	PARK2 ex2-6 het DUP + ex5 het DUP, het IVS8 +48C>T; het D394N	ND		ND			
122	81.04	54		ND		ND			
123	81.05	62	het S167N	ND		ND			
124	81.06	42	het R334C, het S167N	ND		ND			
125	81.23	55		ND		ND			
126	81.27	49	PARK2 ex4 het DEL, het IVS8 +43A>G; het 3'UTR *16G>A	ND		ND			
127	81.28	51		ND		ND			
128	81.58	38		ND		ND			
129	81.60	49	het IVS8 +43A>G	ND	het IVS26-29G>A	ND			
130	81.61	42		ND		ND			
131	81.62	53		ND		ND	het IVS1-7A>G		
132	81.63	65		ND		ND		het IVS6-23 T>C	
133	81.64	70	het P437L	ND	het G2019S	ND			
134	81.65	62		ND		ND	het A340T		
135	81.66	70		ND	het R1514Q	ND			

Nr.	Lab Nr.	AAO	PARK2 (parkin)	PARK7 (DJ-1)	PARK8 (LRRK2)	PARK1/4 (SNCA)	PARK 6 (PINK1)	SNCAIP	MAPT
136	81.67	63	het IVS8 +48C>T	ND		ND	het A340T		
137	81.68	54	het IVS1 +42C>T; het IVS2 +20del(C); het IVS2 +25T>C ; het M192L	ND	het IVS26-33T>C	ND	ND		
138	81.69	53		ND		ND	ND		
139	81.70	55	het M192L; het 3'UTR *16G>A	ND		ND	ND		
140	81.74	48	het D394N	ND		ND	het Q115L, het IVS1-65C>G,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
141	81.81	40		ND		ND	het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
142	81.90	50		ND		ND	het L63L,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
143	82.18	37		ND		ND	het IVS6+43C>T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		het IVS3+18C>T,het A562A, het N590N
144	82.32	59		ND		ND	ND		
145	82.33	70		ND		ND	ND		hom IVS3+18C>T,hom IVS4+9A>G, hom A562A, hom N590N
146	82.34	75		ND		ND	ND		
147	82.35	62		ND		ND	ND		
148	82.36	55		ND		ND	ND		
149	82.38	68	het D394N	ND		ND	ND		het A562A, het N590N
150	82.39	50		ND		ND	het A340T, hom N521T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
151	82.44	56		ND		ND	ND		
152	82.45	80	het D394N	ND		ND	ND		
153	82.46	42		ND		ND	het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
154	82.47	63	het A82E, het S167N; het D394N	ND	het G2019S	ND	ND		
155	82.48	68		ND		ND	ND		
156	82.49	53		ND		ND	ND		
157	82.50	38	A397A	ND		ND	het L63L,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		het A90A
158	82.53	65		ND		ND	ND		
159	82.85	77	het V380L	ND		ND	ND		
160	82.68	64	het IVS2 +25T>C; IVS2 +35T>C	ND		ND	ND		hom IVS4+9A>G
161	82.74	55		ND		ND	ND		
162	82.75	56		ND		ND	ND		
163	82.82	51		ND		ND	ND		
164	82.89	56		ND		ND	ND		
165	82.90	55	het IVS8 +48C>T	ND		ND	ND		
166	82.91	53		ND		ND	ND		
167	82.92	64	het A82E	ND		ND	ND	het N906H	
168	82.93	60		ND	het R1514Q	ND	ND		
169	82.94	66		ND		ND	ND		
170	82.95	56		ND	het S1721S	ND	ND		
171	82.96	60	het IVS8 +48C>T	ND		ND	ND		
172	82.97	60		ND	het IVS23+78C>T	ND	ND		
173	82.99	46	het S167N	ND		ND	het IVS6+43C>T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
174	83.00	62	hom IVS2 +25T>C	ND		ND	ND		
175	83.01	55	het 5'UTR -89C>T	ND		ND	ND		
176	83.02	72		ND		ND	ND		
177	83.03	66		ND		ND	ND		
178	83.04	65		ND		ND	ND		
179	83.05	75		ND		ND	ND		
180	83.06	52		ND		ND	ND		

Nr.	Lab Nr.	AAO	PARK2 (parkin)	PARK7 (DJ-1)	PARK8 (LRRK2)	PARK1/4 (SNCA)	PARK 6 (PINK1)	SNCAIP	MAPT
181	83.07	67	het IVS2 +25T>C; het P37P	ND		ND	ND		
182	83.08	68		ND		ND	ND		
183	83.09	75		ND		ND	ND		
184	83.10	72		ND		ND	ND		
185	83.11	34		ND		ND	het L63L,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
186	83.12	60		ND		ND	ND		
187	83.13	51		ND		ND	ND		
188	83.14	65		ND		ND	ND		
189	83.15	64		ND		ND	ND	het E709Q	
190	83.16	47		ND	het R1514Q	ND	hom IVS6+43C>T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T		
191	83.17	62		ND	het P1542S	ND	ND		
192	83.18	56		ND	het P1542S	ND	ND		
193	83.19	62		ND		ND	ND		
194	83.38	55	het IVS2 +25T>C	ND	het P1542S	ND	ND		
195	83.39	63		ND		ND	ND		
196	83.40	53		ND		ND	ND		
197	83.41	63	het V380L	ND		ND	ND		
198	83.42	68		ND		ND	ND		
199	83.43	73		ND		ND	ND		
200	83.44	52		ND		ND	ND		het IVS3+18C>T
201	83.45	52		ND		ND	ND		
202	83.46	56		ND		ND	ND		
203	83.47	57		ND		ND	ND		
204	83.48	76		ND		ND	ND		
205	83.49	58		ND		ND	ND	ND	ND
206	84.25	58	het S167N	ND	hom G2019S	ND	ND	ND	ND
207	83.88	52		ND		ND	ND	ND	ND
208	83.91	60		ND		ND	ND	ND	ND
209	83.93	45		ND		ND	het N521T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
210	84.02	75	het D394N	ND		ND	ND	ND	ND
211	84.04	57	het D394N	ND		ND	ND	ND	ND
212	84.14	56		ND		ND	ND	ND	ND
213	84.15	73		ND		ND	ND	ND	ND
214	84.18	59		ND		ND	ND	ND	ND
215	84.21	60		ND		ND	ND	ND	ND
216	84.30	34		ND		ND	het 5'UTR-35C>T, het L63L,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
217	84.47	61	PARK2 ex3 & 4 het DEL	ND		ND	ND	ND	ND
218	84.52	57		ND		ND	het IVS4+56G>T,het N521T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
219	84.66	57	het IVS8 +43A>G; het S167N	ND		ND	ND	ND	ND
220	84.67	55		ND		ND	ND	ND	ND
221	84.68	48	het IVS2 +35G>A	ND		ND	het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
222	85.40	46		ND		ND	het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
223	85.56	58		ND	het IVS26-29G>A	ND	ND	ND	ND
224	85.57	62		ND	het R1441C	ND	ND	ND	ND
225	85.58	45		ND		ND	het IVS6+43C>T,het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
226	86.48	35	het R334C	ND		ND	het L63L, het IVS4+72G>C, 3'UTR+43G>A, het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND

Nr.	Lab Nr.	AAO	PARK2 (parkin)	PARK7 (DJ-1)	PARK8 (LRRK2)	PARK1/4 (SNCA)	PARK 6 (PINK1)	SNCAIP	MAPT
227	86.86	54	het S167N, het M192L	ND		ND	ND	ND	ND
228	88.28	40		ND		ND	het A340T, het N521T, 3'UTR+43G>A, het IVS1-7A>G, het IVS7-8T>G, het 3'UTR+40A>T	ND	ND
229	88.29	57		ND		ND	ND	ND	ND
230	88.40	70	ND	ND		ND	ND	ND	ND
231	88.52	65	ND	ND		ND	ND	ND	ND
232	88.53	66	ND	ND		ND	ND	ND	ND
233	88.74	37	ND	ND		ND	ND	ND	ND
234	88.98	58	ND	ND		ND	ND	ND	ND
235	88.99	23	ND	ND		ND	ND	ND	ND
236	89.01	66	ND	ND		ND	ND	ND	ND
237	89.02	56	ND	ND		ND	ND	ND	ND
238	89.03	57	ND	ND		ND	ND	ND	ND
239	90.03	57	ND	ND		ND	ND	ND	ND
240	90.86	70	ND	ND	ND	ND	ND	ND	ND
241	90.87	37	ND	ND	ND	ND	ND	ND	ND
242	90.88	80	ND	ND	ND	ND	ND	ND	ND
243	90.91	59	ND	ND	ND	ND	ND	ND	ND
244	90.92	75	ND	ND	ND	ND	ND	ND	ND
245	90.94	40	ND	ND	ND	ND	ND	ND	ND
246	90.95	27	ND	ND	ND	ND	ND	ND	ND
247	91.84	43	ND	ND	ND	ND	ND	ND	ND
248	91.85	65	ND	ND	ND	ND	ND	ND	ND
249	91.86	59	ND	ND	ND	ND	ND	ND	ND
250	91.87	44	ND	ND	ND	ND	ND	ND	ND

AAO: age at onset, DUP: duplication, DEL: deletion, het: heterozygous, hom: homozygous, ND: not determined
Proven mutations are shown in red
Possible mutations are shown in blue
Polymorphisms are shown in black