Pathophysiology, pleotrophy and paradigm shifts: genetic lessons from Parkinson’s disease

O.A. Ross and M.J. Farrer

Department of Neuroscience, Mayo Clinic, Jacksonville, Fl 32224, U.S.A.

Abstract

PD (Parkinson’s disease) is an aetiologically heterogeneous disorder characterized by a clinical phenotype consisting of resting tremor, rigidity and bradykinesia. Motor symptoms are associated with a progressive loss of dopaminergic neurons, with Lewy body inclusions within surviving neurons. Although heritability studies have shown evidence of familial aggregation, twin studies have provided limited support for a genetic aetiology. Nevertheless, classical linkage methods have nominated 11 regions of the genome and pathogenic mutations have been identified in several genes, including α-synuclein, parkin, ubiquitin C-terminal hydrolase L1, oncogene DJ-1, PTEN-induced protein kinase 1 and microtubule-associated protein tau. Most recently, heterozygous mutations in LRRK2 (leucine-rich repeat kinase 2) were found to cause late-onset, autosomal-dominant PD. Despite their consistent clinical phenotype, family members with LRRK2 mutations can have variable α-synuclein and tau pathologies. Lrrk2 is a member of the Roc (Ras of complex proteins) family, with Ras GTPase and MAPKKK (mitogen-activated protein kinase kinase kinase) catalytic domains. Thus its discovery highlights vesicle dynamics and secondary-messenger signalling in disease pathophysiology. To diagnose a disease accurately and effectively treat it, requires an understanding of its molecular pathogenesis. Herein, we provide an overview of the genetics of PD, how these discoveries are revolutionizing long-held beliefs and more importantly how this knowledge may be translated into patient therapy.

James Parkinson (1755–1824) first described a clinical syndrome characterized by resting tremor and bradykinesia in his 1817 publication ‘An Essay on the Shaking Palsy’. The renowned French neurologist Jean Martin Charcot (1825–1893) defined the syndrome, noting the presentation of rigidity and in tribute named the disorder Parkinson’s disease (PD). PD is the second most prevalent neurodegenerative disorder after AD (Alzheimer’s disease), with approx. 1% of the population older than 50 years being affected. Predominant motor symptoms are the result of a profound deficiency of dopamine neurons in the basal ganglia with degeneration of the substantia nigra pars compacta. The disease is typically asymmetric at onset and gradually progressive. In most patients, proteinaceous Lewy bodies and Lewy neurites are found within remaining neurons post mortem.

While a family history of Parkinsonism is an important contributor to risk [1,2], most cross-sectional twin studies have not supported a genetic aetiology [3]. However, smaller longitudinal studies do not concur, with disease concordance rates in monogygotic and dizygotic twins of 75 and 22% respectively [2]. Reduced/age-associated disease penetrance within mutation carriers and a lack of sufficient statistical power may explain the failure to reach a consensus [4]. Despite this background, in the last decade molecular genetic research in familial Parkinsonism has linked 11 regions of the genome to disease and identified pathogenic mutations in seven genes [SNCA, PRKN, UCH-L1, MAPT, DJ-1, PINK1 and LRRK2 (leucine-rich repeat kinase 2)].

The first gene implicated in dominantly inherited Parkinsonism was α-synuclein (SNCA); missense mutations leading to A30P, E46K and A53T amino acid substitutions cause neurodegeneration in a small number of families [5–7]. α-Synuclein is a major component of Lewy bodies, and common genetic variability within the promoter, leading to enhanced expression, has also been implicated as a risk factor for idiopathic PD [8,9]. More recently, large SNCA genomic triplication and duplication mutations have been identified (PARK1, MIM 168601; and PARK4, MIM 605543) [10–12]. Remarkably, a direct relationship between wild-type SNCA dosage, expression and the age of disease onset, progression and phenotypic severity has been discovered. Indeed, overexpression of the wild-type gene may lead to presentations of PD, Parkinsonism with dementia (PDD) and dementia with Lewy bodies (DLB), consistent with a diagnosis of diffuse Lewy body disease on autopsy (DLBD). This continuum supports the recent staging of Lewy body pathology proposed by Braak et al. [13].

In the UCH-L1 gene, a causal missense mutation (I93M) and common genetic variability (S18Y) has been implicated in dominantly inherited and idiopathic PD (PARK5, MIM 191342) [14,15]. Uch-I is a component of Lewy bodies, and
before α-synuclein reagents, anti-Uch-l1 antibodies were the pathological standard in staining Lewy bodies in PD [16]. Mutations in the MAPT gene, encoding tau, have also been associated with atypical Parkinsonism, with a primary phenotype of FTDP-17 (frontotemporal dementia with Parkinsonism linked to chromosome 17) [17,18]. Nevertheless, common variability in MAPT has also been associated with both idiopathic PD and PSP (progressive supranuclear palsy) [19–21]. FTDP-17 and PSP are characterized by neurofibrillary tangles of the tau protein, rather than Lewy bodies post mortem.

In the case of autosomal recessive forms of early-onset Parkinsonism (<50 years at disease onset), three genes have been identified [22–24]. Mutations in PRKN, DJ-1 and PINK1 leading to loss of function were originally described in consanguineous families with recessively inherited juvenile and early-onset disease (PARK2, MIM 600116 and MIM 602544; PARK7, MIM 606324; and PARK6, MIM 605909). Most pathogenic mutations and rearrangements have been identified in the PRKN locus, encoding parkin on chromosome 6q34 [22]. In referral-based series, the frequency of PRKN mutations has been estimated at 49% in families consistent with recessive inheritance, with onset before the age of 45 years [25]. In idiopathic early-onset Parkinsonism, the frequency of PRKN mutations is estimated at approx. 18%, whereas in community-based, idiopathic and typically late-onset PD the frequency is approx. 3%, similar to that observed in controls [26]. Approximately six patients have come to autopsy and, whereas complete loss of function of the gene causes nigral degeneration and neuronal loss without co-existing pathology, tauopathy and synucleinopathy have been described for compound heterozygous mutations [27–30]. Mutations in the DJ-1 and PINK1 genes were first identified in consanguineous families, although carriers have been identified in several families from different populations [23,24,31,32]. It remains unclear whether haplo-insufficiency for these genes is a risk factor for familial or sporadic PD, as sufficiently powered, epidemiological studies are lacking. The carrier frequency for PINK1 and DJ-1 mutations has yet to be formally estimated and affected family members have yet to come to autopsy [33].

Mutations in LRRK2, encoding the novel Lrrk2 protein, now explain the largest number of patients with familial Parkinsonism (PARK8; OMIM*609007). The PARK8 locus was originally mapped in a large Japanese family, the Sagamihara kindred, presenting with autosomal-dominant Parkinsonism [34]. Affected members of this family had typical levodopa-responsive Parkinsonism with variable age at onset, but mainly in the sixth decade, and pathologically exhibited nigral degeneration without distinctive inclusions [34]. Our group replicated this finding in a number of families of European ancestry, and refined the disease-linked chromosomal region before describing five disease-segregating mutants in the Lrrk2 protein (L1114L, I1122V, I2020T, R1441C and Y1699C) in six unrelated families (Figure 1) [35,36]. Interestingly, two mutations affecting the same codon were reported in a group of Basque families (R1441G) and in a British family (Y1699C) [37]. Linkage of disease in the Basque families to chromosome 12q12 and the PARK8 locus was subsequently shown [37a]. The identification of a common Lrrk2 G2019S substitution, reported at a relatively high frequency in European population has confirmed the importance of this gene in neurodegeneration [38–41]. The frequency, origin and founder effect of some of these pathogenic amino acid substitutions has been reported for different populations [37,40].

The Lrrk2 protein appears to be expressed in most regions of the brain, and cellular-based studies have shown that the wild-type protein appears to be cytoplasmic (Figure 2). Lrrk2, a 286 kDa protein, is a member of the recently defined ROCO [Roc (Ras of complex proteins)/COR (C-terminal of Roc)] family [42]. These proteins are multidomain proteins and have been found in species ranging from mammals to metazoans with a variety of functions including tumour suppression [42]. In human, mouse and rat, members of the ROCO protein family have five conserved domains [42]. Lrrk2 has a large N-terminus ending with leucine-rich repeats and there is a WD40 repeat domain at the carboxylic terminus. The Roc domain contains a GTPase like domain with homology to all four members of the GTPase superfamily. All ROCO proteins contain a novel COR domain, which is approx. 300–400 amino acids long, adjacent to the C-terminus of the Roc domain. In addition to the two domains that give the name to this group of proteins, a third kinase domain is always present in this family, suggesting that the three domains are somehow functionally dependent.

The kinase domain belongs to the MAPKKK (mitogen-activated protein kinase kinase kinase) subfamily of mixed-lineage protein kinases, with catalytic activity for both serine/threonine and tyrosine residues. The active site of all
Figure 2 | Cellular localization of Lrrk2–EGFP tagged fusion protein.
HEK-293T cells, grown on coverslips, were transfected with full-length Lrrk2–EGFP tagged wild-type protein. Cells were fixed with paraformaldehyde after 36 h, mounted and analysed using confocal microscopy. Wild-type Lrrk2–EGFP showed a cytoplasmic distribution throughout the cell; image courtesy of J.P. Taylor (Department of Neuroscience, Mayo Clinic, Jacksonville, FL, U.S.A.).

MAPKKK of this class is located in a cleft between an N- and a C-terminal lobe and is often covered by an ‘activation loop’ in its inactive form (Figure 3). The activation loop undergoes crucial conformational change to provide a docking platform and access to protein substrate, and to orientate key catalytic amino acids within the cleft of the kinase [43]. In different kinases, the activation loop starts and ends with the conserved residues Asp–Phe–Gly (DFG) and Ala–Pro–Glu (APE) respectively [44]. The common pathogenic Lrrk2 G2019S substitution changes the highly conserved glycine (G) within the ‘hinge’, at the start of this sequence [40]. In other kinases, oncogenic mutations in residues within the activation loop of the kinase domain have an activating effect [45]. It is postulated that G2019S may have an activating effect on the kinase activity of Lrrk2 [40]. A mutation causing ‘gain of function’ would be compatible with the dominant mode of disease transmission observed in our families [40]. It may also alter substrate specificity. Although most other Lrrk2 mutations are relatively rare, G2019S is the most common pathogenic substitution observed to date [38–41]. The clinical phenotype of LRRK2-associated disease is indistinguishable from asymmetric, levodopa responsive, late-onset PD [46]. In referral series, this variant may explain up to 5% of familial PD and perhaps as much as 0.5–1% of community-based, apparent sporadic PD [38–41]. To date, three of these ‘sporadic’ cases of PD with G2019S substitutions have come to autopsy with typical Lewy body disease and nigral neuronal loss [41]. Notably, one German family has a pathogenic I2020T substitution, located adjacent to G2019, suggesting that tight regulation of kinase activity is critically important [35].

Figure 3 | Molecular modelling of Lrrk2 Roc and MAPKKK domains
Left panel: the Roc domain between amino acids 1334 and 1456 is shown with the GTP-binding domain (magenta) complexed with GTP analogue GppNHp (green) and highlights switch I (yellow) and switch II regions (red, including the site of the R1398H polymorphism). Specific downstream effector molecules are proposed to bind to the variable RabSF region (turquoise) in addition to detecting the nucleotide-binding state. The Lrrk R1441 substitution (turquoise residue) changes a highly conserved residue on this surface. Right panel: MAPKKK domain between amino acids 1800 and 2138. The active site of the mixed lineage kinase (yellow) is centrally located alongside the ATP-binding region (turquoise). The activation segment (magenta) lies between conserved residues Asp–Phe–Gly (DFG) and Ala–Pro–Glu (APE) respectively [18]. The Lrrk2 G2019S and I2020T substitutions change highly conserved glycine (G) and isoleucine (I) (coloured residues indicated by an arrow) at the start of the activation segment. Figures courtesy of J. Johnson (Department of Neuroscience, Mayo Clinic, Jacksonville, FL, U.S.A.).
Another cluster of pathogenic amino acid substitutions, R1441C/G is located in the Roc domain of the protein [35,37,47]. This motif belongs to the Ras GTPase superfamily, a large class of small proteins that regulate a wide variety of cellular processes, such as signalling, differentiation and growth, through binding and hydrolysis of GTP [42]. Our alignment and molecular modelling studies suggest that the protein region containing the R1441, similar to a RabSF (Rab subfamily) motif, may play a role in protein interactions and cellular localization (Figure 3). RabSF regions are located on the surface of Rab GTPases and are thought to be involved in molecular interactions and protein binding [48,49]. In addition, this motif has recently been implicated in the membrane targeting of some Rab GTPases to specific organelles [48]. Parkinsonism in Family D, which contains affected individuals harbouring an R1441C Lrrk2 variant, has been comprehensively described. Affected members present with a phenotype characterized by late-onset, levodopa-responsive Parkinsonism, for the most part clinically indistinguishable from classic, idiopathic PD [50]. At autopsy, one family member had only nigral neuronal loss and gliosis without co-existing pathology, two family members had α-synuclein pathology consistent with Lewy body PD and diffuse Lewy body disease, whereas one had tau pathology, characterized by neurofibrillary tangles, coiled bodies and tufted astrocytes, the histological hallmarks of PSP. Interestingly, this patient presented with features typical of idiopathic PD but, five years after onset, at the age of 83, developed ocular gaze palsy [50].

The COR domain of the protein harbours the majority of coding variability including nine amino acid substitutions of which some might be expected to have an effect on protein structure (R1514Q, P1542S, V1598E, R1725STOP; [50a]) and is also the location of the pathogenic Y1699C substitution. Although most patients with Lrrk2-associated disease present with typical PD, substitutions within this domain may lead to rather atypical clinical and pathological presentations. Family A, a large German–Canadian kindred, harbours an Lrrk2 Y1699C substitution and has been thoroughly described [51]. Whereas the majority of affected family members presented with parkinsonism, some developed co-existing amyotrophy while others presented with dementia. Two family members have come to autopsy with neuropathological examination and both have neuronal loss and gliosis in the substantia nigra, but Lewy body pathology was not observed. Surprisingly, ubiquitin positive cytoplasmic and nuclear inclusions reminiscent of Marcinkos bodies were observed, with modest involvement of spinal motor neurons [51].

Lrrk2 substitutions are the most common, genetic cause of familial Parkinsonism and idiopathic, late-onset PD identified to date [52]. Our group’s in silico and sequencing studies in families with dominantly inherited Parkinsonism have identified 39 variants within the exonic, open reading frame of the gene [50a]. Of these 39 changes, 27 result in non-synonymous amino acid substitutions. The pleomorphic pathology observed within and across families with Lrrk2 substitutions is perhaps the most surprising finding and highlights a greater role for Lrrk2 substitutions, causal mutations and perhaps common polymorphisms, in neurodegenerative disease. Hence, we postulate that the MAPKKK and Rab GTPase activities of Lrrk2 must lie within a pathway upstream of other proteins implicated in the pathogenesis of Parkinsonism and dopaminergic neuronal survival. It is worthy to note that a number of mixed-lineage MAPKKK inhibitors have been in clinical trials for oncogenic disease and as a neuroprotectant therapy in PD [53]. Thus therapeutic targets aimed at Lrrk2-associated disease, and the underlying molecular pathway, have the potential to revolutionize diagnosis and treatment for a substantial proportion of patients.

Note added in proof (received 15 June 2005)
Since going to press, affected individuals in the original PARK8 family, the Sagamihara kindred, have been found to have an I2020T Lrrk2 substitution [54].

We thank all those who have contributed to our research. This work would not have been possible without the gift of patient participation, and the considerable time and effort of the many physicians involved. This work is support by a Morris K. Udall Parkinson’s Disease Research Center of Excellence (NINDS P01 #NS40256) and the National Institute of Aging (P01 #AG17216). More information regarding studies, resources and subject participation can be found at http://www.mayo.edu/fpd.

References

©2005 Biochemical Society