The genetics of Parkinson’s disease: review of current and emerging candidates

Caroline Ran
Andrea Carmine Belin
Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden

Abstract: Parkinson’s disease (PD) is the second most common neurodegenerative disorder, affecting more than 1% of the population over the age of 65 worldwide. Certain rare forms of the disease are monogenetic, and there is increasing evidence that multiple genetic risk factors are also important for common forms of PD. We have summarized the results from candidate gene and genome-wide association findings in sporadic PD as well as linkage and next-generation sequencing studies of familial PD. To date, 19 genetic loci, PARK1–19, have been reported for rare forms of PD, including autosomal-dominant and autosomal-recessive PD. At 14 of these loci, genes have been identified carrying mutations that are linked to affected family members. These genes have also been shown to constitute candidate genes for idiopathic forms of PD, since they may also carry other mutations that merely increase risk. Multiple genetic factors combine in different ways to increase or decrease risk, and several of these risk factors need to be identified in order to begin unraveling the causative pathways leading to the different forms of PD. In this review, we present current and emerging PD candidate genes to help explain the pathways leading to neurodegeneration.

Keywords: polymorphism, linkage, synuclein, mitochondria, toxins, GWAS

Introduction
Parkinson’s disease (PD) is a common neurodegenerative disorder in individuals over the age of 65, affecting approximately 1%–2% with no ethnic preference. It is characterized primarily by the degeneration of dopamine (DA) neurons in substantia nigra (SN), but there is also evidence of degeneration in other neurons as well.1 Another typical PD characteristic is the presence of proteinaceous intracellular inclusions, called Lewy bodies (LB), in the brain stem and elsewhere.2 Although there is a relatively large body of knowledge on the pathology and pathophysiology of PD, there is little understanding of its etiology. The main principles for treating PD today – replacing DA, improving the effect of DA, or inhibiting the breakdown of DA – do not cure the disease; they only reduce the symptoms caused by the loss of DA neurons. In order to create a curative treatment for PD, the underlying causes must be understood. While less common forms of PD are monogenic, representing 5%–10% of all PD cases, there is increasing evidence that multiple genetic factors are also important for sporadic PD. Several genetic approaches have been used to screen for genetic risk factors, linkage studies, candidate gene studies, genome-wide association studies (GWASs), and next-generation sequencing. To date, the PARK1–15 and PARK17–19 loci have been identified as risk factors for rare familial forms of PD, while the PARK16 locus was identified as a risk factor in sporadic PD. At 14 of these loci, genes referred to as...
PARK genes have been identified which carry mutations that are linked to affected family members in several PD families. These genes also constitute candidate genes for sporadic PD, since those same genes may carry other mutations that increase risk or confer protection. Positive as well as negative genetic association studies in PD have been reported for genetic variants in a plethora of candidate genes, but results from replication studies have sometimes been ambiguous. GWASs have resulted in identifying new candidate gene loci as well as the replication of findings from linkage studies. In this review, we summarize the latest results from candidate gene and genome-wide association findings in sporadic PD, as well as linkage and next-generation sequencing studies of familial PD to help explain the pathways leading to neurodegeneration in PD.

**Linkage studies**

**SNCA/PARK1/PARK4**

The first linkage study on PD was conducted by Polymeropoulos et al, who mapped a PD locus to chromosome 4q21, known as PARK1. A nonsynonymous mutation (A53T) was later identified in the α-synuclein gene (SNCA) in Greek PD families with an autosomal dominant pattern of inheritance. Although this mutation only constitutes a rare cause of PD in three unrelated families of Greek origin, it highlighted the genetic variability in PD. The discovery of the point mutation A53T in SNCA was subsequently followed by the identification of A30P and E46K in unrelated German and Spanish PD families, with the latter also showing LB dementia. More recently, H50Q in exon 4 of SNCA was identified in one individual of British origin affected by PD. SNCA encodes the α-synuclein protein, for which the normal physiological function has not yet been determined, although α-synuclein has a potential role in synaptic plasticity and vesicle dynamics. It has also been shown in a number of studies that α-synuclein is abundantly expressed in the brain and that it constitutes a major component of LBs and Lewy neurites. The role of LBs as pathogenic, passive bystanders or possibly protective entities in PD has not been elucidated. The gene dose of SNCA appears to be critical, since SNCA duplication and triplication have also been found to cause PD. A triplication of the SNCA locus (incorrectly designated PARK4) results in a doubling in α-synuclein protein levels. Moreover, a repetitive sequence in the promoter region and single nucleotide polymorphisms (SNPs) in intronic sequences have also been found to be associated with PD, suggesting that regulation of gene transcription activity and messenger (m)RNA stability might play a role in α-synuclein-associated PD. SNCA is one of the few genes reported to be associated with PD that has been repeatedly replicated in GWASs.

**LRRK2/PARK8**

Leucine-rich repeat kinase 2 (LRRK2), localized at the PARK8 locus, was suggested as a candidate gene for PD in 2004. Many LRRK2 mutations have now been linked to PD. The G2019S mutation in LRRK2 is the largest single known cause of PD today. It contributes to both sporadic and familial disease in populations of different ethnic origins. We identified G2019S in Swedish sporadic PD patients and analyzed LRRK2 gene expression in the brain. We found that this gene has a relationship to the DA system, evidenced by its high expression in striatum (but not in DA neurons) in rodents and human postmortem tissue. Regarding other reported mutations in LRRK2, numerous replication studies have been performed on sporadic cohorts, showing variable results. However, the LRRK2 locus has been confirmed to be associated with PD in several GWASs. Also a recent meta-analysis based on 61 independent studies in Asian populations further confirmed LRRK2 mutations as risk factors for PD. Despite extensive research, the exact function of LRRK2 remains unknown. LRRK2 has been shown to bind to a large number of proteins via two protein-protein interaction domains, a leucine-rich repeat domain and a carboxy terminal WD40 domain. LRRK2 also harbors two enzymatic domains, one kinase and one GTPase domain, and is probably involved in quite diverse cellular functions such as mitochondrial function, apoptosis, membrane trafficking, and microtubule polymerization that may contribute to PD pathogenicity. Interestingly, LRRK2 is potentially interacting with many other PD candidate genes described in this review; ie, Ras-related protein Rab7-L1, vacuolar protein sorting 35 homologue (VPS35), SNCA, V-akt murine thymoma viral oncogene homologue 1 (AKT1), and microtubule-associated protein Tau (MAPT).

**Glucocerebrosidase/GBA**

Mutations in the glucocerebrosidase gene (GBA), originally identified as the cause of Gaucher’s disease, have been associated with both familial and sporadic PD. GBA is located on chromosome 1q21 and encodes a lysosomal protein, glucocerebrosidase (GCase), which cleaves the beta-glycosidic linkage of glucosylceramide, an intermediate formed during glycolipid metabolism. Genetic variations in GBA are known to alter GCase activity.
might lead to lysosomal dysfunction or interfere with the binding of α-synuclein to its putative receptor at the lysosomal membrane. This could result in reduced α-synuclein degradation, leading to cell toxicity and potentially PD when affecting DA neurons, for example.

The most common GBA mutations found in PD are L444P and N370S. Parkinson patients carrying heterozygous mutations in GBA are characterized by relatively early disease onset (≤50 years of age), but are otherwise clinically similar to patients with sporadic PD and respond well to L-DOPA (L-3,4-dihydroxyphenylalanine). Known GBA mutations affect α-synuclein protein levels both in cells and in transgenic mice. This is confirmed by the finding that induced pluripotent stem cell-derived DA neurons from Gaucher’s disease patients carrying N370S show increased levels of α-synuclein compared to DA neurons derived from healthy individuals. They also show lower GCase protein levels and enzymatic activity. Mazzulli et al demonstrated that elevated α-synuclein levels caused by low GCase leads to further reduction of lysosomal GCase activity and increase of the GCase retained in the endoplasmic reticulum in murine cortical neurons, thus suggesting that improper GCase handling results in increased α-synuclein accumulation.

**VPS35/PARK17**

Recently, VPS35 was identified through exome sequencing, a next-generation sequencing technique, as a risk gene for PD in families with several affected members originating from Austria and Switzerland. VPS35 is a component of the retromer complex involved in the transport of specific membrane proteins, in particular acid hydrolase receptors, from endosomes to the trans-Golgi network to avoid their degradation in lysosomes; dysfunction of this pathway could potentially lead to PD pathogenesis. The first description of this form of autosomal dominant PD came from a Swiss family. Patients were described to have an initially asymmetric Parkinsonism with tremor similar to classic sporadic PD. Sequencing of the Swiss pedigree, as well as an independent Austrian PD pedigree, revealed that all affected family members carried the D620N mutation in the VPS35 gene. It is noteworthy that seven mutation carriers were detected in middle-aged neurologically healthy individuals, which indicates that the penetrance is high but might not be complete. The D620N mutation has also been reported in PD in Germany.

In silico analysis of the substitution of aspartic acid at position 620 to the basic amino acid asparagine predicted the mutation to be deleterious. The substitution is likely to affect protein conformation at the C-terminus, since the asparagine residue forms fewer chemical bonds, thus making the protein more flexible. Taken together, the D620N substitution has also been identified and could potentially influence the risk of developing PD; several of these variants were identified only in one individual, while others were found in patient populations and in neurologically healthy controls. These data were confirmed in a large multicenter replication study performed on almost 9,000 PD patients, in which D620N was found in seven patients while the other mutations were found in both patients and control subjects. Because of the rarity of these genetic variants, further studies in large cohorts are required to determine the contribution of VPS35 mutations to develop PD.

**EIF4G1/PARK18**

Analysis of a French pedigree with PD revealed significant linkage to chromosome 3q26-q28. One mutation, R1205H, in the eukaryotic translation initiation factor 4-gamma 1 (EIF4G1) segregated with disease. EIF4G1 is the scaffold protein of the RNA translation initiation complex. The R1205H mutation is located close to the putative binding site of EIF3E, another component of the complex; when analyzed in silico, the mutation was predicted to have a deleterious effect on protein function. Indeed, the R1205H substitution decreased the binding between EIF4G1 and EIF3E in transfected cell lines. In addition, cells with mutated EIF4G1 had a more pronounced loss of mitochondrial membrane potential when exposed to stress. Mitochondrial dysfunction has also been implicated as an important risk factor in the pathogenesis of PD, since it can lead to apoptosis of midbrain DA neurons.

Replication studies have so far failed to detect additional EIF4G1 R1205H mutation carriers among PD cases or detected the mutation exclusively in control subjects; however, it is possible that the replication studies were underpowered to detect this rare mutation (0.2% of the PD patients in the initial study). It should also be noted that one of the replication studies was performed on a cohort with a different ethnic background, while haplotype analysis in the discovery study indicated that mutation carriers were likely to share an ancestral founder. It is also worth noting that the healthy control individuals carrying R1205H came from a general population cohort and might be susceptible to the development of PD later in life. Nevertheless, it cannot be excluded that R1205H is not itself disease causing and...
that the real disease-causing variant at the large multigene 3q26-q28 locus remains to be discovered.

Numerous mutations in EIF4G1 have been discovered in the extensive genetic analyses described above, but the evidence of pathogenicity is ambiguous. Many mutations were found in too few patients, segregation analysis was not practicable, or variants were also reported in control subjects; however, missense mutation A502V is a promising candidate. In conformity with R1205H, A502V was predicted to be deleterious to protein function and is expected to alter the binding of the EIF4E component of the translation initiation complex.57

DNAJC6/PARK19

Recently, pathogenic DnaJ (Hsp40) homologue subfamily C, member 6 (DNAJC6) mutations have been reported in two independent consanguineous families with juvenile Parkinsonism using next-generation sequencing.53,54 Also, member 13 of the DnaJ (Hsp40) homologue subfamily C (DNAJC13) has been discovered to harbor mutations potentially causing PD.55 DNAJC6 on chromosome 1p31.3 encodes auxilin, which is a neuron-specific co-chaperone protein. Together with the molecular chaperone, Hsp70c, auxilin is responsible for the removal and recycling of clathrin from coated endocytosis vesicles.56 DNAJC6’s malfunction leads to impaired synaptic vesicle recycling and disturbed clathrin-mediated endocytosis.53 Dopamine receptors are internalized by clathrin-mediated endocytosis to be degraded by the endosomal system; thus, mutations in DNAJC6 can lead to an abnormal DA receptor metabolism, which can be an important factor for developing PD.57

The first pathogenic mutation in DNAJC6 was found in two Palestinian brothers homozygous for the potentially deleterious mutation c.801-2 A>G,53 while unaffected family members were heterozygous for the genetic variation. In a second study, another DNAJC6 mutation, c.182A>T, was found to segregate with Parkinsonism in a Turkish family.54 Patients in both families have rare forms of juvenile-onset Parkinsonism. Replication studies have so far been negative;58,59 however, this does not exclude DNAJC6 as a candidate gene for PD. More research is needed to understand the possible implication of this gene in PD, and should tentatively be focused on PD patients with juvenile or early onset disease.

Summary of linkage studies

Linkage studies first revealed the existence of a genetic component in PD etiology. After decades of genetic research, the disease is now viewed as a complex genetic disorder. Mutations in genes first identified as disease causing for familial PD have also been discovered in sporadic disease, and SNPs in those genes have also been found to associate with sporadic disease. Herein, we have summarized findings on a few important genes identified through linkage studies (Table 1): SNCA, which was the first gene discovered to harbor PD-associated mutations, as well as LRRK2 and GBA, which together constitute the most common known causes for sporadic PD today. Not all genes found in familial studies are relevant in sporadic PD, nor have they been replicated. VPS35, EIF4G1, and DNAJC6 are the most recent candidate genes, none of which have yet been established as risk factors for common PD. Even so, these genes are of great interest, since they offer indications of molecular pathways that may be disrupted in PD pathology. VPS35, for instance, is a member of the retromer complex, which is involved in the transport of proteins between the Golgi network and endosomes. Alterations in this pathway could potentially interfere with normal lysosome function leading to PD pathogenesis, thus shedding more light on the hypothesis suggesting a role for defective lysosomes in PD as indicated by an involvement of GBA, for example. Concerning VPS35 in particular, the PD-associated mutation has been suggested to cause an endosome phenotype where endosomes are relocalized to a perinuclear position.60 Somehow this impairs transport of the lysosomal protease cathepsin D to the lysosomal compartment, which might affect lysosome function. Furthermore, this ties VPS35 to another risk gene for PD, SNCA, since α-synuclein is a known cathepsin D substrate.61

Candidate gene studies

Candidate genes with several different functional properties have been proposed to be involved in the pathogenesis of PD. In this section, we describe candidate genes involved in four areas that have all been suggested to be important in the...
development of PD: abnormal protein aggregation, detoxifying enzymes, mitochondrial function, and oxidative stress. We also describe a recent candidate gene, integrin alpha 8 (ITGA8), with no known link to PD pathophysiology.

Abnormal protein aggregation
In many neurological disorders, there is evidence of abnormal protein aggregation in postmortem brain tissue of patients. In PD, these aggregations are observed as LB, which are large proteinaceous inclusions in the cytoplasm of neurons, usually seen in the DA neurons of SN. The main component of these inclusions is α-synuclein, but other proteins are also present. Attempts to deduce the composition of these protein aggregations have been difficult, since they display unspecific binding to antibodies.

MAPT/Tau is also a potential component of LBs. Although not regarded as part of the pathological hallmarks of PD, neurofibrillar tangles, which are large intracellular aggregations of hyperphosphorylated tau, have also been noted in postmortem brains of PD patients. Interestingly, tau pathology is more pronounced in PD patients suffering from cognitive decline.62,63

The interest in MAPT increased in the late 1990s after the discovery that MAPT was associated with two disorders related to PD, frontotemporal dementia with Parkinsonism and progressive supranuclear palsy.64–66 The first reported association between MAPT and late onset PD was made by Pastor et al in 2000.67 The H1 haplotype described in supranuclear palsy by Baker et al was found to be overrepresented in PD patients compared to control subjects.65,67 Since these publications, MAPT association with PD has been confirmed both in association studies and in GWAS.13–18,20–22 Exactly how tau influences the pathological changes in the Parkinsonian brain is unclear. Tau binds to microtubules and promotes their stabilization. When phosphorylated, tau is detached from microtubules, with subsequent disruption of microtubule organization. Most pathological genetic variations seen in tau affect the capacity of tau to bind microtubules and to aggregate.68 Interestingly, expression of the tau 4R isoform, which has four microtubule binding repeats and is a product of alternative splicing of exon 10, is increased in PD patients. This is likely to affect the capacity of binding to microtubules, which is critical for normal neuronal function.

Detoxifying enzymes
A few genes associated with PD encode detoxifying enzymes. SNPs and mutations in genes that affect enzymatic activity and or expression levels are prone to affect the risk of developing PD in concert with exposure to environmental risk factors. This combination of effects might mask genetic correlation, making these genes more difficult to study.

Pesticide exposure is one environmental factor that has been suggested to increase the risk to develop PD. The arylesterase paraoxonase 1 (PON1) hydrolyzes organophosphates (such as pesticides) and harbors the L54M mutation. L54M impairs enzyme activity, leads to decreased PON1 expression levels, and has been reported to be associated with PD.69 The minor allele of a PON1 promoter polymorphism, rs854571, has been reported to be more common among controls than PD cases, thus suggesting a protective effect.70 rs854571 is in strong linkage disequilibrium (LD) with another PON1 promoter polymorphism, rs854572, reported to increase PON1 gene expression.71

Mitochondria
Several genes involved in mitochondrial function have been suggested as candidate genes for PD. One such gene is DNA polymerase gamma (POLG1), which codes for the catalytic subunit of the heterotrimeric mitochondrial DNA polymerase and is involved in replication and repair of mitochondrial DNA.72 In addition to its 5′ to 3′ polymerase activity, POLG1 also has a 3′ to 5′ exonuclease activity that is important in the repair process. POLG1 has an N-terminal polyglutamine tract that is encoded by a CAG sequence in exon 2. Variations in repeat length, possibly affecting POLG1 function, have been reported to be associated with PD by several research groups.73

Oxidative stress
Oxidative stress has also been proposed as a mechanism underlying the development of PD. There are currently two hypotheses; the first postulates that the DA neurons are more susceptible to oxidative stress, while the second postulates that increased oxidative stress has global effects in the PD brain. The latter explanation is in line with the notion that none of the many genes discovered to date which have been postulated to cause or increase risk for PD are specific for the DA system and that PD pathology often includes other parts of the nervous system. Dysfunction of mitochondria and impaired protein degradation, for example, are two cellular events that occur during PD which could increase cellular stress and oxidative stress in particular.74

AKT1 on chromosome 14q32.32 encodes a serine/threonine protein kinase that is a potent promoter of cell survival which could possibly protect the PD brain from oxidative stress or, when impaired, increase levels of stress.
In 2008, a protective AKT1 haplotype was found to associate with PD in Greece,73 but this genetic association could not be replicated by us in Swedish PD patients.76 While GWASs have not confirmed AKT1 as a protective gene in PD, AKT proteins are less phosphorylated in the brains of PD patients compared to healthy control subjects; therefore, it seems likely that the AKT/PKB (Protein Kinase B) family is somehow implicated in PD.77,78 Whether AKT is a cause or a consequence of PD remains to be established.

**ITGA8**

The ITGA8 locus on chromosome 10p13 reached genome-wide significance level in a large meta-analysis performed on the studies available in the PDGene database.1,79 The database included 828 individual association studies from more than 50 countries and was combined with GWAS data of all currently published GWASs, which resulted in a total of 16,452 PD patients and 48,810 control subjects. Among several markers in ITGA8, one intronic SNP, rs7077361, was found to associate with PD.

**ITGA8** encodes the α8 subunit of the integrin family of receptors. The gene contains many SNPs, some of which are potentially deleterious nonsynonymous variations.80 The integrin α8 subunit forms a complex exclusively with the β1 subunit and is expressed in smooth muscle, myofibroblasts of the lung, and in the mesangial cells of the kidney glomeruli. ITGA8 knockout mice have severe kidney malformation or no kidneys at all.81,82 Moreover, ITGA8 seems to be involved in memory function: mice lacking the α8 subunit in forebrain excitatory neurons display a reduction in long-term plasticity, as measured in acute hippocampal slices.83 Despite the results of the meta-analysis, there is no obvious link between ITGA8 and PD pathophysiology.

**GWAS**

Genome-wide technology enables large-scale, hypothesis-free searches of associated SNPs in the whole genome. Although there were great expectations that these studies would rapidly identify new susceptibility loci for PD, the results of GWASs in PD have been somewhat disappointing. The first GWASs done in Parkinson cohorts were underpowered due to small sample sizes and did not reach genome-wide significance.84–86 Also, there were no replications of genetic loci between the different studies. In 2009, the first significant GWASs were published.19,20 SNCA at the PARK4 loci was confirmed to associate with PD in two independent cohorts. In the European population, MAPT was also confirmed as a risk locus for PD.20 In the replication phase of this study, Satake et al were the first to identify two new PD-associated loci by means of GWAS,19 PARK16 at 1q32 and the bone marrow stromal cell antigen 1 (BST1) at 4p15 (Table 2). Furthermore, LRRK2 at 12q12 was confirmed as a risk gene for PD. Since 2009, SNCA, MAPT, and LRRK2 have all been replicated in GWASs.13–15,18,21,22 Three other candidate loci have also been discovered through GWASs, HLA-DRA, SREBF/RAI1, and FAM47E/SCARB2 (Table 2): 1) human leukocyte antigen (HLA) locus harboring HLA-DRA expressed by antigen-presenting cells;15 2) a locus was found on chromosome 17 tagged by rs11868035, which potentially could map to SCAP SREBF chaperone (SREBF) and/or retinoic acid induced 1 (RAI1);13 3) a locus marked by rs6812193 on chromosome 4 which lies within the family with sequence similarity 47 member E (FAM47E) gene. This SNP is also close to a more interesting candidate gene, scavenger receptor class B member 2 (SCARB2). SCARB2 encodes the lysosomal integral membrane protein type 2, a protein which directs GCase to the lysosomes.13 This finding strengthens the involvement of GBA in PD, and highlights the importance of correctly functioning lysosomes in the brain.

The lack of significance in GWASs described herein has several explanations. Insufficient power in terms of too few participants has been a major problem that could be overcome with GWAS meta-analysis. It is also possible that both genetic risk factors and protection factors are different depending on

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>SNP</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARK16</td>
<td>1q32</td>
<td>rs947211</td>
<td>Satake et al, 200919</td>
</tr>
<tr>
<td>BST1</td>
<td>4p45</td>
<td>rs4538457</td>
<td>Satake et al, 200919</td>
</tr>
<tr>
<td>HLA-DRA</td>
<td>6p21</td>
<td>rs3129882</td>
<td>Hamza et al, 201015</td>
</tr>
<tr>
<td>FAM47E/SCARB2</td>
<td>4p16</td>
<td>rs6812193</td>
<td>Do et al, 201113</td>
</tr>
<tr>
<td>SREBF/RAI1</td>
<td>17p11</td>
<td>rs11868035</td>
<td>Do et al, 201113</td>
</tr>
<tr>
<td>GAK</td>
<td>4p16</td>
<td>rs1564282</td>
<td>Nalls et al, 201116</td>
</tr>
<tr>
<td>ACM5D</td>
<td>2q21.3</td>
<td>rs6710823</td>
<td>Nalls et al, 201116</td>
</tr>
<tr>
<td>STK39</td>
<td>2q24.3</td>
<td>rs2102808</td>
<td>Nalls et al, 201116</td>
</tr>
<tr>
<td>MCCCI1/LAMP3</td>
<td>3q27</td>
<td>rs1711441</td>
<td>Nalls et al, 201116</td>
</tr>
<tr>
<td>SYT11</td>
<td>1q21.2</td>
<td>rs3472695*</td>
<td>Nalls et al, 201116</td>
</tr>
<tr>
<td>CCD6C2/HIP1R</td>
<td>12q24</td>
<td>rs12817488</td>
<td>Nalls et al, 201116</td>
</tr>
<tr>
<td>SETD1A/STX1B</td>
<td>16p11.2</td>
<td>rs4889603</td>
<td>IPDGC, 201118</td>
</tr>
<tr>
<td>STBD1</td>
<td>4p21</td>
<td>rs6812193</td>
<td>IPDGC, 201118</td>
</tr>
<tr>
<td>GPNNMB</td>
<td>7p15</td>
<td>rs156429</td>
<td>IPDGC, 201118</td>
</tr>
<tr>
<td>DGKQ</td>
<td>4p16</td>
<td>rs11248060</td>
<td>Pankratz et al, 201217</td>
</tr>
<tr>
<td>RIT2</td>
<td>18q12.3</td>
<td>rs12456492</td>
<td>Pankratz et al, 201217</td>
</tr>
</tbody>
</table>

**Note:** *In proxy with Chr1:154105678.
Abbreviation: SNP, single nucleotide polymorphism; rs, reference SNP number.*
ethnic background, which could potentially have influenced results from GWASs where large sample sizes are used from cohorts with different ethnic backgrounds. Differences in PD phenotypes such as disease onset, cognitive impairments, symptom severity, and tremor should also be considered. Moreover, there is also the possibility that risk and protective factors are heterogeneous and mirror the variability of Parkinsonian phenotypes, thus indicating that the different forms of PD be considered and analyzed as different disorders. In recognition of this, many studies now exclude PD patients who carry known pathogenic mutations such as G2019S in LRRK2 and L444P in GBA. It is important not to consider results from GWASs as identification of pathogenic variations leading to PD, but rather as indicators of which loci might harbor genes and genetic variations that could lead to PD pathology.

Three meta-analysis studies performed on pre-existing GWAS datasets have identified a few new PD loci (listed in Table 2), which include 1) aminocarboxymuconate semialdehyde decarboxylase (ACMSD), a candidate gene which falls in the group of detoxifying enzymes that prevents the accumulation of the neurotoxic metabolite quinolinate, which is a product of tryptophan metabolism; 2) serine threonine kinase 39 (STK39), a kinase protein activated by cellular stress; 3) chromosomal loci 3q27, which tags either methylcrotonoyl-CoA carboxylase 1 (MCCC1) or lysosomal-associated membrane protein 3 (LAMP3); 4) synaptotagmin 11 (SYT11), which encodes a calcium-sensing protein involved in membrane trafficking in synaptic transmission and also constitutes a known substrate of Parkin; 5) chromosomal loci 12q24, which tags either coiled-coil domain containing 62 (CCDC62) or huntingtin-interacting protein 1 related (HIP1R); 6) chromosomal loci 16p11.2, which tags either SET domain containing 1A (SETD1A), a component of a histone methyltransferase, or syntaxin 1B (STX1B), a receptor for vesicle transport; 7) starch-binding domain 1 (STBD1); 8) glycoprotein nmb (GPNNMB); and 9) Ras-like without CAAX (RIT2).17

**DGKQ/GAK**

Cyclin G associated kinase (GAK) and diacylglycerol kinase theta (DGKQ) were suggested as susceptibility genes in an early GWAS where a genetic locus encompassing both genes at chromosome 4p16 reached P-values close to significance in the order of magnitude 10^-9.86 Further support for the involvement of this locus in PD susceptibility was provided by two additional GWASs.21,22 The association was finally confirmed for rs1564282 in GAK in a meta-analysis which included the discovery study material16 and in a replication study performed on an American case-control cohort.88 Since then, many replications of GAK have been reported in PD populations of different ethnicities. Recently, there have also been reports of association with rs11248060 in DGKQ17,79,89

Further studies on the PD associated locus at chromosome 4p16 should include both genes in order to be able to draw conclusions about their possible involvement in PD. GAK encodes a ubiquitously expressed protein kinase, which is active during the cell cycle.90 Of interest to PD pathophysiology is the involvement of GAK in uncoating of clathrin-coated vesicles.91 Many genes involved in vesicle transport and the endosomal-lysosomal pathway have been suggested as risk-genes for PD; ie, GBA, VPS35, and SCARB2. Functional data support the pathogenicity of GAK gene dysregulation and GAK mutations. Cathepsin D binds to GAK and expression levels of the cathepsin D gene correlate with those of GAK,92 indicating that GAK could have an impact on lysosomal function. Moreover, GAK mRNA levels are known to be decreased in SN pars compacta of PD patients compared to aged control individuals.93 In addition, the disease-associated SNP, rs1564282, is associated with increased SNCA mRNA expression in the frontal cortex. The relationship between α-synuclein and GAK has been confirmed in SNCA transfected human embryonic kidney 293 (HEK293) cells, where α-synuclein levels were increased by inhibition of GAK expression.92 This effect on α-synuclein is probably due to the decrease in lysosomal cathepsin D.

Less is known about the possible implication of DGKQ in PD. DGKQ is a member of the diacylglycerol kinase (DGK) gene family. These proteins are ubiquitously expressed and responsible for the conversion of diacylglycerol (DAG) to phosphatidic acid (PA) by phosphorylation. All DGK members share a well-conserved kinase domain responsible for the enzymatic activity, and different regulatory domains and motifs confer their specificity. The phosphorylation of DAG is an important step in phosphatidylinositol (PI) turnover, and it is also involved in lipid signaling. Both DAG and PA are biologically active signaling molecules. DGK proteins have been shown to take part in multiple cellular processes. In the central nervous system, DGK proteins are known to affect Ca^2+ signaling, as well as the trafficking and fusing of synaptic vesicles at nerve terminals.94

**PARK16**

The PARK16 locus was first associated with PD in a Japanese GWAS.95 PARK16 is located in a gene-rich area on
chromosome 1q32, and several SNPs within the locus strongly associate with disease. Using LD analysis, five candidate genes have been suggested for PARK16: 1) SLC45A3 (Solute carrier family 45 member 3); 2) NUCKSI (nuclear casein kinase cyclin-dependent kinase substrate 1); 3) RAB7LI (or RAB29); 4) SLC41A1 (Solute carrier family 41 member 1); and 5) PM20D1 (Peptidase M20 containing domain 1). Since then, PARK16 has been replicated at a genome-wide significance level in European and American PD populations. A close-up study of three genes at the PARK16 locus (NUCKSI, RAB7LI, and SLC41A1) in a British population revealed the presence of rare mutations in both RAB7LI and SLC41A1.95 One common SNP in RAB7LI was also found to be associated with an increased risk of PD, thus strengthening the hypothesis of involvement of this locus, and in particular for RAB7LI, in PD. Three rare mutations were also found in SLC41A1 in Chinese PD patients.96

It is of great importance that there are large global variations in the genotype frequencies at this locus95 and that this locus might, therefore, exert different degrees of risk of PD in different populations. The results from replication studies are indeed ambiguous as a replication study performed on two Asian cohorts, genotyping the most significant SNPs from the initial GWAS reports19,20 suggested SNPs in the PARK16 loci to have a protective effect on PD which was also observed in a Chilean and in a Taiwanese PD population.97–99 Of note, the Chilean population displays minor allele frequencies similar to those of the Asian populations. Yet other studies show a lack of association.99–101 This variability might be accounted for by differences in genetic background. Considering the replication studies, it seems likely that PARK16 is a susceptibility locus for Asian populations.

As previously described, GWASs highlight genetic loci harboring SNPs and mutations that might confer increased or decreased risk of disease; additional investigations of other SNPs within these loci are needed to confirm the involvement of PARK16 in PD patients with European descent. The differences across populations might also reflect the possibility that the true disease-causing SNP is merely in LD with the SNPs described here and that the LD blocks are different in Asian and European populations. Alternatively, the PARK16 locus might of course contain several independent risk and protective genes for PD.

SLC41A1 encodes a magnesium (Mg$^{2+}$)/sodium (Na$^{+}$) exchanger which is believed to be important to maintain Mg$^{2+}$ homeostasis. Phosphorylation of SLC41A1 by protein kinase A affects the Mg$^{2+}$ efflux, probably by modulating Mg$^{2+}$ affinity.102,103 SLC41A1 has been reported to be associated with three very different human disorders: pre-eclampsia, nephropathiasis-like phenotype, and now PD. How this gene might be involved in the pathophysiology of PD remains to be elucidated.104

Little is known about RAB7LI. RAB genes encode small GTP-binding proteins involved in membrane trafficking. The RAB7LI protein has the conserved domains important for GTP-binding and membrane association, and it is probably involved in exo- and endocytosis.105 In support of this, RAB7LI has been found to be localized to the Golgi apparatus and to membranous transport intermediates.106 RAB7LI is ubiquitously expressed, but at very low levels in the brain.107 Interestingly, RAB7LI seems to interact with LRRK2. The risk effect of the LRRK2 locus was shown to be highly dependent on the absence of the protective PARK16 allele in a reanalysis of preexisting GWAS data.94 In addition, an SNP in LD with PD risk SNP, rs947211, was found to be associated with alternative splicing of RAB7LI. Healthy carriers of this allele displayed lower RAB7LI protein levels, comparable to the decrease of RAB7LI found in PD patients that is similar between carriers and noncarriers.10 Healthy carriers of this allele displayed lower RAB7LI protein levels, comparable to the decrease of RAB7LI found in PD patients that is similar between carriers and noncarriers.10 Malfunctioning of RAB7LI is hypothesized to affect the intracellular sorting processes between Golgi and lysosomes, for which further support has been presented by knocking down RAB7LI. Attenuation of RAB7LI produces swollen lysosomes with reduced MPR (cation-independent mannos 6-phosphate receptor), which is necessary to maintain lysosomal function.107 These results do not necessarily mirror the situation in the human brain in vivo, but render RAB7LI a promising candidate gene for PD as both an LRRK2 interaction partner and as a mediator in the protein sorting and lysosomal defect pathway.

NUCKSI encodes a nuclear protein with unknown function and does not share homology with any known gene product. The gene displays characteristics of a housekeeping gene: it is highly expressed in brain and all peripheral tissues,107,108 the promoter lacks TATA and CAAT boxes, and the 5’ region is enriched in CpG pairs.107 The NUCKSI DNA sequence contains numerous putative phosphorylation sites, and the protein is readily phosphorylated by Cdk-1 (Cyclin-dependent kinase 1).109 Although NUCKSI association with PD has not yet been validated in replication studies, rs947211, an SNP in PARK16 that associates with PD, influences the expression levels of NUCKSI.19

BST1/CD157

The PD locus on chromosome 4p15 was first discovered in the same Japanese GWAS as PARK16.19 The BST1 finding
has been replicated in populations of European and Australian origin.\textsuperscript{18,21} \textit{BST1} also turned out to be associated with PD in a larger meta-analysis performed on five cohorts from Europe and USA in which the two previously mentioned studies were also included.\textsuperscript{16} Replication studies on \textit{BST1} show conflicting results, which is most likely due to low power or population heterogeneity.\textsuperscript{101,109,110} Larger replication studies seem to support \textit{BST1} as a candidate gene for PD that has protective genetic variations. The effect is more prevalent in Asian than Caucasian PD populations.\textsuperscript{79,111,112} \textit{BST1} is a cell surface protein bound to the membrane by glycosylphosphatidylinositol linkage and possesses both (Adenosine diphosphate) ADP ribosyl cyclase and cyclic ADP ribose hydrolase enzymatic activities.\textsuperscript{113,114} \textit{BST1} has also been ascribed several additional modes of action which include receptor activity. How this membrane-bound protein with extracellular enzymatic activity can affect intracellular events and signaling has not been elucidated.\textsuperscript{115} Interestingly, \textit{BST1} seems to be an important factor in the immune system, since it is highly expressed in bone marrow cells from patients with rheumatoid arthritis and facilitates immature B-cell proliferation and growth.\textsuperscript{114,115}

### The HLA-locus

Another group of genes that have been studied in PD that are related to the immune system are located at the HLA region on chromosome 6p21. This locus was initially discovered in a GWAS performed on an American PD case-control material.\textsuperscript{15} The HLA region is one of the most polymorphic regions in the human genome, containing several genes coding for class I, II, and III HLA molecules. Evidence of the HLA locus being associated with PD is quite substantial and concerns different genes within the locus. The first report of HLA involvement in PD was published before the GWAS era in 2003, when an \textit{HLA-DQB1} allele was suggested to be associated with PD.\textsuperscript{116} This was followed by a GWAS finding supporting \textit{HLA-DRA},\textsuperscript{115} and several publications on \textit{DRA-DRB}.\textsuperscript{15,117,118}

LD analysis has revealed that several of the HLA-DR loci might represent true associations with PD;\textsuperscript{119} however, published data on this locus are more in favor of \textit{HLA-DRB}, which is the more polymorphic gene. Since HLA genes are closely located on chromosome 6, it is also possible that they act in synergy.

Here, we have reviewed evidence of the association of two loci involved in the immune system: \textit{BST1} and \textit{HLA}. There is clear evidence for the involvement of inflammatory reactions in PD. \textit{HLA} genes as well as other immunological genes involved in antigen processing and presentation have been found to be differentially expressed in idiopathic PD.\textsuperscript{120} Whether inflammatory and/or immunological mechanisms are causative in PD or are merely a consequence of pathology remains to be elucidated. In any case, the evidence indicates that common variations in genes implicated in inflammation and other immunological pathways should be considered potential risk factors for PD.

### Chromosome 18q12.3/RIT2

Although the \textit{RIT2} gene on chromosome 18q12.3 was suggested as a candidate gene for PD by Do et al,\textsuperscript{13} it did not reach genome-wide significance. There is, however, conflicting evidence as to whether \textit{RIT2} is a candidate gene for PD. In a recent study, the \textit{RIT2} gene was implicated as a candidate gene in a replication meta-analysis that did not include the discovery dataset.\textsuperscript{17} In contrast, an independent replication study in a Taiwanese population did not show association. It is plausible that the variation is of no importance in the Asian populations, since allele frequencies are very different between Asian and Caucasian populations at this locus.\textsuperscript{121} \textit{RIT2} encodes RIT2, a small GTPase of the Ras family, which is neuron specific and capable of binding calmodulin.\textsuperscript{122} Singly, \textit{RIT2} has a direct connection to DA since RIT2 colocalizes with, and is capable of interacting with, the DA transporter (DAT). This interaction is specific and does not extend to other types of transporters. DAT is found at the plasma membrane, but can be internalized by protein kinase C-mediated endocytosis. It seems that this mechanism also depends on RIT2 GTPase activity.\textsuperscript{123}

The 18q12.3 genetic region also contains the gene synaptotagmin 4 (\textit{SYT4}), a potential PD risk gene through its involvement in vesicle trafficking and endocytosis, as well as its relation to \textit{SYT11}. proposed to associate with PD by Nalls et al.\textsuperscript{16} Synaptotagmins are a family of proteins that are involved in synaptic transmission. They have Ca\textsuperscript{2+} binding domains that regulate the binding of phospholipids, so they are important in the regulation of vesicle fusion and endocytosis at synaptic terminals. In particular, \textit{SYT4} and \textit{SYT11} are different from other types of synaptotagmins in that they carry a conserved point mutation within a Ca\textsuperscript{2+} binding domain which renders them incapable of Ca\textsuperscript{2+} and phospholipid binding.\textsuperscript{124} All functions of synaptotagmins are not Ca\textsuperscript{2+} dependent, however, and \textit{SYT4} and \textit{SYT11} are capable of forming dimers with other isoforms as well as interacting with other proteins. In the cortex of mice, \textit{SYT4} is localized to intracellular membranous organelles primarily in the Golgi apparatus, but also in vesicles.\textsuperscript{125} \textit{SYT4} can be found in both dendrites and axons, but not in presynaptic vesicles, thus
making it more likely to be involved in membrane transport and exocytosis than participating in synaptic transmission as does its binding partner SYT1. SYT4 knockout mice have impaired motor coordination and impaired hippocampal dependent memory function.126

Conclusion
In order to develop better ways of combating PD—or better yet, prevent clinical symptoms altogether—one must elucidate the causes of disease and take into consideration that there are different forms of the disorder, as well as that any given phenotype may be the effect of different genotypes plus environmental exposures. The identification of new genetic risk and protective factors for PD gives insight into the pathophysiological mechanisms, and thus may potentially aid in the development of novel classes of drugs based on causative factors, and eventually prevent and/or counteract disease. PD patients suffer from major physical and mental discomfort, so the development of more efficient treatments with less or no side effects is of the utmost importance. At the time of a clinical diagnosis of PD, the majority of the DA neurons in SN pars compacta have already been lost; therefore, a challenge for the future is to identify new markers which can be used for presymptomatic diagnosis of PD and for future neuroprotective treatments to become effective. The identification of genetic markers for PD is important in that these markers may allow diagnosis before disease onset and could result in lifelong preventive medications. Today, a diagnosis of PD by a neurologist can only be definitely confirmed by postmortem brain analysis. Findings from genetic research may lead to the development of diagnostic/prognostic methods based on blood samples, or the even less invasive buccal test. Results from several research groups involving genetic studies suggest that PD is not one, but several diseases. Genetic research can give rise to new drugs for treatment based on causative factors with different drug targets. In the future, we will most likely see more individual treatments for PD, based on genetic information, which should lead to a reduction in side effects and more efficient treatments.

In particular, we would like to stress the significance of the retromer associated sorting and the lysosomal pathways. The body of evidence in support of the involvement of these pathways is steadily increasing; and many of the genes discussed in the present review are implicated as risk genes by sorting and lysosomal degradation events: GBA, VPS35, RAB7L1, DGKQ, and GAK, as well as the less-studied genes SYT4, SCARB2, and LAMP3; impairment of these pathways is potentially detrimental in many ways (for example, they could cause impaired protein degradation and accumulation). Indeed, it has been suggested that improper GCase handling results in increase of α-synuclein accumulation, which also provides a link between α-synuclein pathology and the more recently identified candidate genes. α-Synuclein has been demonstrated to have a role in PD by several methods, through genetic findings in linkage studies and GWASs, as well as by the fact that it constitutes a major component of LBs and Lewy neurites. Also, LRRK2 might be involved in this retromer-lysosome pathway, since it has been demonstrated that RAB7L1 as well as VPS35 knockout mimics the LRRK2 knockout mice phenotype of lysosomal swelling and neurite collapse.

In this review, we have summarized the most recent and important results from candidate gene and genome-wide association findings in sporadic PD, as well as in linkage and next-generation sequencing studies of familial PD in order to help explain the different possible pathways leading to neurodegeneration in PD.

Acknowledgments
We would like to thank Dr Mathew Birdsell Abrams for critically reading through this review. This work was supported by the Swedish Research Council, the Swedish Brain Foundation, the Swedish Parkinson Foundation, the Swedish Brain Power Initiative, Åke Wibergs Stiftelse, and Karolinska Institutet Funds.

Disclosure
The authors report no conflicts of interest in this work.

References


