

Review

The Birth of the Modern Era of Parkinson's Disease Genetics

Andrew B. Singleton^{a,*}, John A. Hardy^b and Thomas Gasser^c

^aLaboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA

^bDepartment of Molecular Neuroscience, UCL Institute of Neurology, University College London, London, UK

^cDepartment of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany



Dr. Singleton is an NIH Distinguished Investigator within the National Institute on Aging at NIH. He identified alpha-synuclein multiplication mutations and *LRRK2* mutations in Parkinson's disease and has been a leader in the identification of risk loci for Parkinson's disease. His laboratory has identified the genetic cause of more than a dozen forms of neurological disease and has identified hundreds of risk loci for various diseases and traits. He has been awarded the Jay van Andel Award for Outstanding Achievement in Parkinson's Disease Research, the Annemarie Opprecht Foundation Award for Parkinson's Disease Research, the NIH Director's Award, and the American Academy of Neurology Movement Disorders Research Award. He has authored more than 500 publications.



Dr. Gasser is a Professor of Neurology and the Director of the Department of Neurodegenerative Diseases at the Hertie-Institute for Clinical Brain Research at the University of Tübingen, and coordinator of clinical studies of the Tübingen site of the German Center for Neurodegenerative Diseases. He discovered SGCE mutations as a cause for myoclonus-dystonia syndrome and *LRRK2* mutations in Parkinson's disease. He has received the Dingebauer-Award of the German Neurologic Society and the Zülch-Prize of the Max-Planck Society. He has authored more than 400 papers.



Dr. Hardy is Professor of Molecular Neuroscience at University College London. He shared the Annemarie Opprecht Foundation Award for Parkinson's Disease Research with Dr. Singleton. He has also won the Pritzker, the MetLife, Potamkin and Breakthrough Prizes for his work on the genetics of neurodegenerative disease. He has authored over 800 papers on these topics.

*Correspondence to: Andrew B. Singleton, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA.
E-mail: Singleta@mail.nih.gov.

Genetic understanding in Parkinson's disease (PD) has followed a path of hard won evolution occasionally punctuated by revolution. While it was suggested early on by both Leroux and Gowers that heredity had a role to play in PD, this was a view that wasn't widely enough held to even be unpopular. The dogma was that the disease was one of environmental provenance and while the evidence for this is still rather scarce, this view pervades in the minds of patients, clinicians, and scientists. Conversely the evidence linking genetics to PD is both overwhelming and growing. Here we describe the growth of genetics in PD from backwater to driving force, and the structure and shape of its future.

The localization and identification of α -synuclein mutations as a cause of PD in the mid 1990's was perhaps the first concrete and revolutionary finding in PD genetics [1]. This came about as a result of the intuition and hard work of a clinical team from New Jersey, followed by the linkage and positional cloning efforts of a genetic team at NIH, orchestrated by the then director of NINDS, Zach Hall. This effort (described by Bob Nussbaum in another article in this issue) was an extraordinary success.

The discovery of α -synuclein mutations as a rare cause of PD was an invigorating and welcome progression for myriad reasons. Most prominently, it gave us the mutation as a tool with which to attempt to understand the disease process. Perhaps more importantly, at least in the short term, it provided empirical evidence that there was a genetic basis for rare forms of the disease and because α -synuclein was a major component of all Lewy bodies, that these findings were directly relevant to all cases of PD. This in fact, prompted one of us to say, tongue in cheek "If you're not working on synuclein, you're not working on Parkinson's disease".

A BRAVE NEW WORLD

With this finding not only did the scientists pursuing a functional understanding of PD have a new tool but we as geneticists had a new 'in'. Many of us had previously worked in Alzheimer's or known monogenic disorders, but now there was a good rationale for concentrating more genetic efforts in PD. This wasn't based on a new belief that PD was a genetic disease (being geneticists, we thought everything was genetic anyway) but because the finding of

α -synuclein mutations meant that the phrase "Parkinson's is not a genetic disease" would be less evident in the grant reviews we received from then on (although sadly not entirely absent) and that there was a greater likelihood of garnering support for this work.

As geneticists the late 1990's and early 2000's was an exciting time where the major accomplishments centered almost completely on monogenic forms of disease. It is hard to convey the urgency and excitement of gene identification during this time. Finding a new genetic cause of disease was a major undertaking, requiring collaboration between geneticists and clinicians; the latter had sometimes spent decades tracking down and characterizing rare families, and the former invested enormous amounts of time, effort, and money on a somewhat unpredictable process. However, the payoff was huge, the promise of mutations were that they could provide a molecular start with which to piece together the disease, they were a tangible, early, and inarguable component in the disease process. The publication of a novel cause of disease was met with great interest and these papers were high-profile, well-cited, and often paradigm altering. A particularly compelling aspect of this search was that it truly represented the completion of a puzzle; finding a mutation and knowing that this is the cause of disease provides instant gratification - at that instant you are the only person in the World to know the solution to a particularly vexing problem.

Because the rewards were so great, and because being second to identify a particular genetic lesion yielded little, gene hunting was an extremely competitive and fast-paced world. This remained true in the late 1990's through to 2005. In thinking back over the sheer scale of work and effort that we and others placed in this area, it is perhaps surprising that only a small number of genes were identified for PD during this time; however, the influence of these findings was great. The identification of mutations in genes encoding parkin, dj1, pink1 and lrrk2 in addition to the identification of gene dosage mutation of α -synuclein represented major advances that provided insight into the genetic basis of disease and tools with which to understand the biology of that influence [2-7]. How close these races to new mutations could get is illustrated by the fact that the two papers identifying LRRK2 as a PD gene appeared back to back in the same issue of "Neuron" [5, 6].

IN RISK, LITTLE REWARD

While the main pay dirt for geneticists during this period was the cloning of monogenic forms of disease, there was also a great deal of fool's gold mined by our attempts to identify common genetic variants that influence disease risk. *APOE* in Alzheimer's disease was the example that many of us strove to replicate [8]; however the strength and frequency of this risk allele quickly proved to be the exception rather than the rule (and indeed this was not discovered through candidate analysis but really by family based linkage) [9]. The ease in which candidate gene association studies could be executed meant that anyone with a thermal cycler and a hundred samples could perform a study that would be published. This led to an overabundance of reported associations, which almost all proved to be false positives. With hindsight the view that the right gene and the right variant could be selected based on a perception of biological understanding was a little self delusional - and it is therefore not surprising that this was a largely fruitless area of research.

Although it has been argued that the exceptions to this folly were *SNCA* and *MAPT* [10, 11] these genes were really nominated as candidates for PD based on existing genetic or pathologic evidence, the former of which was based on unbiased genetic evidence (ie monogenic gene cloning) the latter from characterization of a major deposited species in neurodegenerative disease [12–15]. The one true exception to this is the finding that heterozygous *GBA* mutations predispose to PD. This observation was made through the keen clinical observation that the parents and grandparents of children with Gaucher's disease had an increased incidence of PD (odds ratio of ~4) [16, 17]. While this idea took time to become proven and accepted (perhaps we were gunshy from the previous 10 years of failed associations), it turned out to be a critical observation. First, of course, because, it focussed attention on the lysosome and other lysosomal genes have subsequently been identified as risk loci for the disease, and second because it suggested that other heterozygous loss of function mutations could predispose to late onset disease. A very similar example of this being *TREM2* mutations, which when homozygous cause a young-onset disorder with a neurological component and when heterozygous increase risk for Alzheimer's disease [18–21].

SECOND GENERATION GENETICS

The most successful approaches in genetics have centered on an unbiased assessment of variability - ie genome wide and with good enough coverage to identify or tag a variant of interest. This is, in essence, why linkage worked so well - recombinations are distant enough within families that the heredity of individual chunks of chromosomes could be easily traced with only a few hundred sentinel markers. This type of approach is analogous to rotating through objective lenses on a microscope, using the positional information from the current view to zoom in on a smaller scale area of interest. Conversely the candidate approach is more like throwing a slide on, going straight the highest magnification and hoping to be in the right place and in focus.

While linkage panels were powerful enough to detect regions of interest in families, where recombinations are few and shared segments are large, they did not possess sufficient resolution to detect regions that were identical and shared between distantly related individuals in a population, particularly on a genome wide scale. Unbiased detection of common risk loci, which were generally of small effect and by definition were 'old' alleles therefore required a very dense assay of genetic variability that could be performed in a large number of individuals. The introduction of single nucleotide polymorphism (SNP) chips met this need. Our laboratories were some of the first to use these methods and it is hard to overstate the excitement within our groups at performing these experiments and seeing the data. We were enthusiastic participants in a transition that took a good lab from a space where they could generate a few thousand genotypes a day, to one where they could generate millions. This was a true revolution in our capabilities, reminiscent of moving from southern blot to PCR amplification. It was certainly one where we recognized the excitement - with one of our laboratories working in shifts to ensure data production 24 hours a day 7 days a week over a period of 6 months (even loading chips for scanning on xmas day).

PD was one of the first diseases to be investigated using this method, with the second published genome wide association (GWA) study and the first publicly available GWA data in a disease [22, 23]. Both of these studies were unsuccessful in individually identifying risk variants for PD; however, both laid the foundation for future work through providing existing

data and by showing empirically that larger numbers (and thus collaboration) was needed.

SECOND GENERATION GENETICISTS

With changing methods came a requirement in a change in the way in which geneticists operated. The methods were, at first, unproven and extremely expensive. Array based genotyping initial cost ~\$1000 a sample and after genotyping a few hundred cases and controls an individual investigator was left with a large amount of data, a fairly rudimentary idea about how to analyze it, and a lot of hope. In the early GWA studies it became quickly clear that a few hundred or a thousand samples was simply not enough to see unequivocal association - and even with an initial genome wide significant signal, a large amount of replication was required. Thus, those of us in this area found ourselves experiencing the 7 stages of high-content genetics grief:

1. **Shock** - "I can't believe there's nothing new to see"
2. **Denial** - "there must be something new to see, we should reanalyze the data"
3. **Bargaining** - "if you reanalyze the data for me I'll give you authorship"
4. **Guilt** - "all that work and nothing to show for it"
5. **Anger** - "I knew I should have worked on single gene disorders"
6. **Depression** - "I bet my competitor has found something and it's under review somewhere awesome"
7. **Acceptance/hope** - "I'm going to have to collaborate with someone to increase my n, then we'll definitely find something"

On the plus side, geneticists as a group moved through the grieving process quite quickly. Within a year or two of the original individual GWA efforts collaborations were being formed and meta-analyses executed that indeed identified new genetic loci for disease [24, 25]. These collaborative groups formed the basis of large international consortia that now serve as discovery engines and large reference datasets in PD genetics. Thus, while we have not entirely gotten rid of the seven stages, they are certainly truncated and in large part can be efficiently dispatched.

Thus, hand in hand with the maturing approach there has been a remarkable change in culture, from

secretive and isolated single lab approaches, towards open science, data sharing and collaboration to a degree that was inconceivable 20 years ago. Innovative models of data acquisition, for example by direct to consumer genetics companies, are successfully used to leverage academic efforts and the use of cloud based infrastructure is making the sharing and democratic distribution of massive scale data a reality. Such an evolution has been a critical step in the formation of groups that aim to take advantage of the second generation genetics tools of whole exome sequencing (WES) and whole genome sequencing (WGS).

The early success of WGS and WES approaches has been in the area of identifying the cause of highly penetrant monogenic PD [26–28]. Certainly these approaches afford a more complete and much more rapid detection of the genetic causes of rare familial forms of disease. The expansion of these efforts into more population based approaches is underway [29]. These data are easily shared across sites, particularly using cloud-based infrastructure. Such sharing not only increases the power to detect genetic variants of interest but also serves to produce a reference PD dataset that the field as a whole can use.

To be clear: a continuation and expansion of data sharing and collaborations is needed to face the challenges ahead and to move towards cures. We still know only a fraction of the total heritability of PD, some of it presumably still hidden in the huge number of rare coding variants that are found by whole exome sequencing. These data will be fully interpretable only when much larger cohorts are available that allow us to look at genetic variability at much higher resolution than to date. And then, the non-coding regions of the genome, the complexities of gene-gene and gene-environment interactions will still prove a considerable but worthwhile challenge.

RELATIONSHIP STATUS: "IT'S COMPLICATED"

It should not have been a surprise that the biological interpretation and understanding of complex genetic risk factors would in turn be complicated; however the extent of the challenge that such a problem holds was perhaps unanticipated. The complexity of going from (usually) non-coding SNP associations in GWAs to a mechanistic understanding of that association is a difficult and multi-dimensional

problem. Sometimes the gene underlying the association is almost certain (for example, SNPs in both the *SNCA* and *LRRK2* locus show association with PD), but even there the effect of the risk haplotype is not immediately obvious. With *SNCA*, the association seems to reflect genetic variability in *SNCA* expression, which makes sense in terms of the pathogenic effect of gene duplications, but the precise mechanism of expression control remains unclear. With *LRRK2*, again, expression control is likely to be involved but this has been difficult to dissect and the effect may in fact be complex and on *LRRK2* splicing. With other loci, the locus dissection is more complex still, with it often being unclear as to which gene at the locus underlies the relationship to the disease. Increasingly, simple genetic analysis to prove gene associations at a locus, need to be complemented by whole genome expression studies and bioinformatic analyses designed to understand whether any particular gene fits into pathogenic pathways. Thus, data from many sources need to be integrated: genetic data (SNP associations, plus the occurrence of rare high risk variants), gene expression studies (does the risk SNP correlate with increases or decreases in expression), protein studies (is the protein a binding partner of a known PD gene), function studies (does the protein fit into a pathway already implicated in disease). Because these data are from different fields of investigation it remains a challenge to integrate them and have a consistent and reliable view of the disease mechanism as mediated at any locus. The challenge of integrating diverse datasets is one that complex genetics faces across diseases. During the resolution of these challenges, undoubtedly, mistakes will be made, but as more loci are dissected it should make dissection of subsequent loci easier just as a jigsaw puzzle gets easier the more pieces are filled.

WRONG PLACE, WRONG TIME, WRONG PEOPLE

Much (perhaps all!) of the work we describe above is aimed at understanding the chain of events that represent the disease process at the molecular, cellular, and systems level, with an eye toward developing an etiologic based intervention. In thinking about why we have thus far been unsuccessful in introducing a PD modifying therapy into the clinic it is perhaps most obvious that we may have been looking in the *wrong place* i.e. our understanding about the molec-

ular processes is so imperfect that we do not know enough to identify a good, clean therapeutic target. This has almost certainly been the case; however it is a case that will and already is changing. Therapeutics aimed at modifying and effecting *Lrrk2* or alpha-synuclein are etiologically sensible and rational and there can be little doubt that more rational drug design against etiologic targets will be forthcoming; however, unfortunately, even effective therapies against the right target can fail - if they are applied at the wrong time or in the wrong subjects. These two hypothetical barriers to success are predicated on the ideas that there is likely a critical mass that the disease process reaches after which it is difficult to halt the insidious progression, current trials may be too late in the disease process, being applied at the *wrong time*; additionally it may be that there is more than one etiologic subtype of PD, this is the heart of the precision medicine movement and it would suggest that certain etiologic-based therapies will only be successful in patients matched to those therapies - current trials, which have test PD as a whole, may have been executed in the *wrong people*.

Identifying the right patients at the right time is a considerable challenge - and one that needs to be met in parallel to understanding the etiology of disease if we are to maximize our likelihood of success in trials. This, in our opinion, is a need that must be met, and one that genetics has a large part to play in. We envision that our fields approach to identifying risk factors and causes of disease will be adapted to understand an individual's risk for disease, when they are likely to show clinical signs, and to determine whether there are indeed distinct etiologic subtypes of PD. Some of this work is already ongoing, with success in the use of clinically ignorant factors to predict disease status and to identify PD mimic presentations [30]. Much more needs to be done in this regard; clearly genetics alone will not suffice in this regard and our view is that the greatest likelihood of success in predicting and subtyping disease will come through multimodal data, including genetics, biomarkers, and longitudinal data. In the context of discovery, our efforts will use genetic tools identical to those described above but will require cohorts with much deeper phenotypic and biologic data - a considerable challenge but one that is beginning to be met. We predict that the integration of these data, along with data from basic research, will be the most efficient path to resolving the challenges of selecting the *right target*, for the *right patient*, and modulating it at the *right time*.

A NEW BRAVE NEW WORLD

The remarkable progress of the last 20 years has opened our eyes to the enormous ocean of unknowns that lies before us. As a field we have certainly come a long way, admittedly we've strayed off the correct course occasionally but we are most assuredly closer to our destination that we were 20 years ago. In many ways genetics has been the engine that has pushed us along on our voyage to understand and treat PD, we know now, more than ever that genetics is a central component to every case of PD and one for which we have the tools at hand to understand. So with this in mind we might be tempted to finish with some more illustrative hyperbole - "if you're not working with genetics, you're not working on Parkinson's disease".

ACKNOWLEDGMENTS

We would like to thank our friends, colleagues, and competitors (and some who have been all) who have contributed so much to this work; perhaps more importantly, we would like to thank the patients and their families for selflessly taking part in and supporting these efforts.

This work was supported in part by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services; project number ZO1 AG000949.

REFERENCES

- [1] Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B et al. (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*, **276**(5321), 2045-2047.
- [2] Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M et al. (2003) Alpha-synuclein locus triplication causes Parkinson's disease. *Science*, **5646**(30), 841.
- [3] Valente EM (2004) Hereditary early-onset Parkinson's disease caused by mutations in PINK1. *Science*, **304**(5674), 1158-1160.
- [4] Bonifati V, Rizzo P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, Dekker MCJ et al. (2003) Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science*, **299**(5604), 256-259.
- [5] Zimprich A, Biskup S, Leitner P, Lichtner P, Farrer M, Lincoln S, Kachergus J et al. (2004) Mutations in LRRK2 cause autosomal-dominant Parkinsonism with pleomorphic pathology. *Neuron*, **44**(4), 601-607.
- [6] Paisán-Ruiz C, Jain S, Whitney Evans E, Gilks WP, Simón J, van der Brug M, López de Munain A et al. (2004) Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron*, **44**(4), 595-600.
- [7] Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, & Shimizu N (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature*, **392**(6676), 605-608.
- [8] Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, & Roses AD (1993) Apolipoprotein E: High-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*, **90**(5), 1977-1981.
- [9] Pericak-Vance MA, Bebout JL, Gaskell PC Jr, Yamaoka LH, Hung WY, Alberts MJ, Walker AP, Bartlett RJ, Haynes CA, & Welsh KA (1991) Linkage studies in familial Alzheimer disease: Evidence for chromosome 19 linkage. *American Journal of Human Genetics*, **48**(6), 1034-1050.
- [10] Krüger R, Vieira-Saecker AM, Kuhn W, Berg D, Müller T, Kühnl N, Fuchs GA et al. (1999) Increased susceptibility to sporadic Parkinson's disease by a certain combined alpha-synuclein/apolipoprotein E genotype. *Annals of Neurology*, **45**(5), 611-617.
- [11] Maraganore DM, de Andrade M, Elbaz A, Farrer MJ, Ioannidis JP, Krüger R, Rocca WA et al. (2006) Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. *JAMA: The Journal of the American Medical Association*, **296**(6), 661-670.
- [12] Brion JP, Couck AM, Passareiro E, & Flament-Durand J (1985) Neurofibrillary tangles of Alzheimer's disease: An immunohistochemical study. *Journal of Submicroscopic Cytology*, **17**(1), 89-96.
- [13] Grundke-Iqbal I, Iqbal K, Tung YC, Quinlan M, Wisniewski HM, & Binder LI (1986) Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences of the United States of America*, **83**(13), 4913-4917.
- [14] Kosik KS, Joachim CL, & Selkoe DJ (1987) Microtubule-associated protein tau (?) Is a major antigenic component of paired helical filaments in Alzheimer disease. *Alzheimer Disease and Associated Disorders* **1**(3), 203.
- [15] Wood JG, Mirra SS, Pollock NJ, & Binder LI (1986) Neurofibrillary tangles of Alzheimer disease share antigenic determinants with the axonal microtubule-associated protein tau (tau). *Proceedings of the National Academy of Sciences of the United States of America* **83**(11), 4040-4043.
- [16] Aharon-Peretz J, Rosenbaum H, Gershoni-Baruch R (2004) Mutations in the glucocerebrosidase gene and Parkinson's disease in ashkenazi Jews. *The New England Journal of Medicine*, **351**(19), 1972-1977.
- [17] Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, Bar-Shira A et al. (2009) Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *The New England Journal of Medicine*, **361**(17), 1651-1661.
- [18] Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, Cruchaga C et al. (2013) TREM2 variants in Alzheimer's disease. *The New England Journal of Medicine*, **368**(2), 117-127.
- [19] Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, Bjornsson S et al. (2013) Variant of TREM2 associated with the risk of Alzheimer's disease. *The New England Journal of Medicine*, **368**(2), 107-116.
- [20] Guerreiro RJ, Lohmann E, Miguel Brás J, Raphael Gibbs J, Rohrer JD, Gurunlian N, Dursun B et al. (2013) Using exome sequencing to reveal mutations in TREM2 presenting as a frontotemporal dementia-like syndrome without bone involvement. *JAMA Neurology*, **70**(1), 78-84.

- [21] Paloneva J, Manninen T, Christman G, Hovanes K, Mandelin J, Adolfsson R, Bianchin M et al. (2002) Mutations in two genes encoding different subunits of a receptor signaling complex result in an identical disease phenotype. *American Journal of Human Genetics*, **71**(3), 656-662.
- [22] Fung H-C, Scholz S, Matarin M, Simón-Sánchez J, Hernandez D, Britton A, Raphael Gibbs J et al. (2006) Genome-wide genotyping in Parkinson's disease and neurologically normal controls: First stage analysis and public release of data. *Lancet Neurology*, **5**(11), 911-916.
- [23] Maraganore DM, de Andrade M, Lesnick TG, Strain KJ, Farrer MJ, Rocca WA, Krishna Pant PV, Frazer KA, Cox DR, Ballinger DG (2005) High-resolution whole-genome association study of Parkinson disease. *American Journal of Human Genetics*, **77**(5), 685-693.
- [24] Simón-Sánchez J, Schulte C, Bras JM, Sharma M, Raphael Gibbs J, Berg D, Paisan-Ruiz C et al. (2009) Genome-wide association study reveals genetic risk underlying Parkinson's disease. *Nature Genetics*, **41**(12), 1308-1312.
- [25] Satake W, Nakabayashi Y, Mizuta I, Hirota Y, Ito C, Kubo M, Kawaguchi T et al. (2009) Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nature Genetics*, **41**(12), 1303-1307.
- [26] Vilarinho-Güell C, Wider C, Ross OA, Dachsel JC, Kachergus JM, Lincoln SJ, Soto-Ortolaza AI et al. (2011) VPS35 mutations in Parkinson disease. *American Journal of Human Genetics*, **89**(1), 162-167.
- [27] Zimprich A, Benet-Pagès A, Struhal W, Graf E, Eck SH, Offman MN, Haubenberger D et al. (2011) A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *American Journal of Human Genetics*, **89**(1), 168-175.
- [28] Lesage S, Drouet V, Majounie E, Deramecourt V, Jacoupy M, Nicolas A, Cormier-Dequaire F et al. (2016) Loss of VPS13C function in autosomal-recessive Parkinsonism causes mitochondrial dysfunction and increases PINK1/Parkin-dependent mitophagy. *American Journal of Human Genetics*, **98**(3), 500-513.
- [29] Jansen IE, Ye H, Heetveld S, Lechler MC, Michels H, Seinstra RI, Lubbe SJ et al. (2017) Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing. *Genome Biology*, **18**(1), 22.
- [30] Nalls MA, McLean CY, Rick J, Eberly S, Hutten SJ, Gwinn K, Sutherland M et al. (2015) Diagnosis of Parkinson's disease on the basis of clinical and genetic classification: A population-based modelling study. *Lancet Neurology*, **14**(10), 1002-1009.