

MINIREVIEW

Copy number variation in Parkinson's disease

Mathias Toft¹ and Owen A Ross^{*2}

Abstract

A central theme of human genetic studies is to understand genomic variation and how this underlies the inherited basis of disease. Genomic variation can provide increased biological understanding of disease processes, which is necessary to develop future treatments. Recent technological advances have highlighted the role of copy number variants in normal and pathological phenotypic expression. These applications have been used in studies of Parkinson's disease, a common, late-onset, progressive neurodegenerative disorder. At present the main therapeutic approach is administration of symptom-alleviating drugs, which neither reverses the disease process nor halts its progression. However, the generation of *in vivo* model systems and development of novel disease intervention strategies for Parkinson's disease have come from research on monogenic forms of the disorder, including those caused by copy number variants. Here, we review the role of copy number variants and the mechanistic insights they have provided on the pathogenesis of Parkinson's disease.

Copy number variation

Traditional cytogenetic approaches first showed that variations in chromosome copy number could cause disease in humans. However, it is only with the recent advances in genome scanning technology that screening for human structural mutations and their role in disease susceptibility has really come into the limelight. Completion of the human genome project showed that copy number variation is a widespread and common phenomenon in humans [1]. A copy number variant (CNV) is a region of DNA in which allelic-number differences have been found by comparison of two or more independent genomes. These DNA segments may range

from less than one kilobase (kb) to several megabases (Mb) in size and can be caused by genomic rearrangements such as deletions, multiplications, inversions and translocations [2]. The vast majority of CNVs are unbalanced, may be limited to a single gene or include a contiguous set of genes and can either be inherited or caused by *de novo* events. Altered expression levels of these CNV genes may be responsible for observed phenotypic variability, complex behavioral traits and disease susceptibility. So far, CNVs have been implicated in the pathogenesis of several neurological disorders, including Parkinson's disease (PD) [3,4].

Parkinson's disease

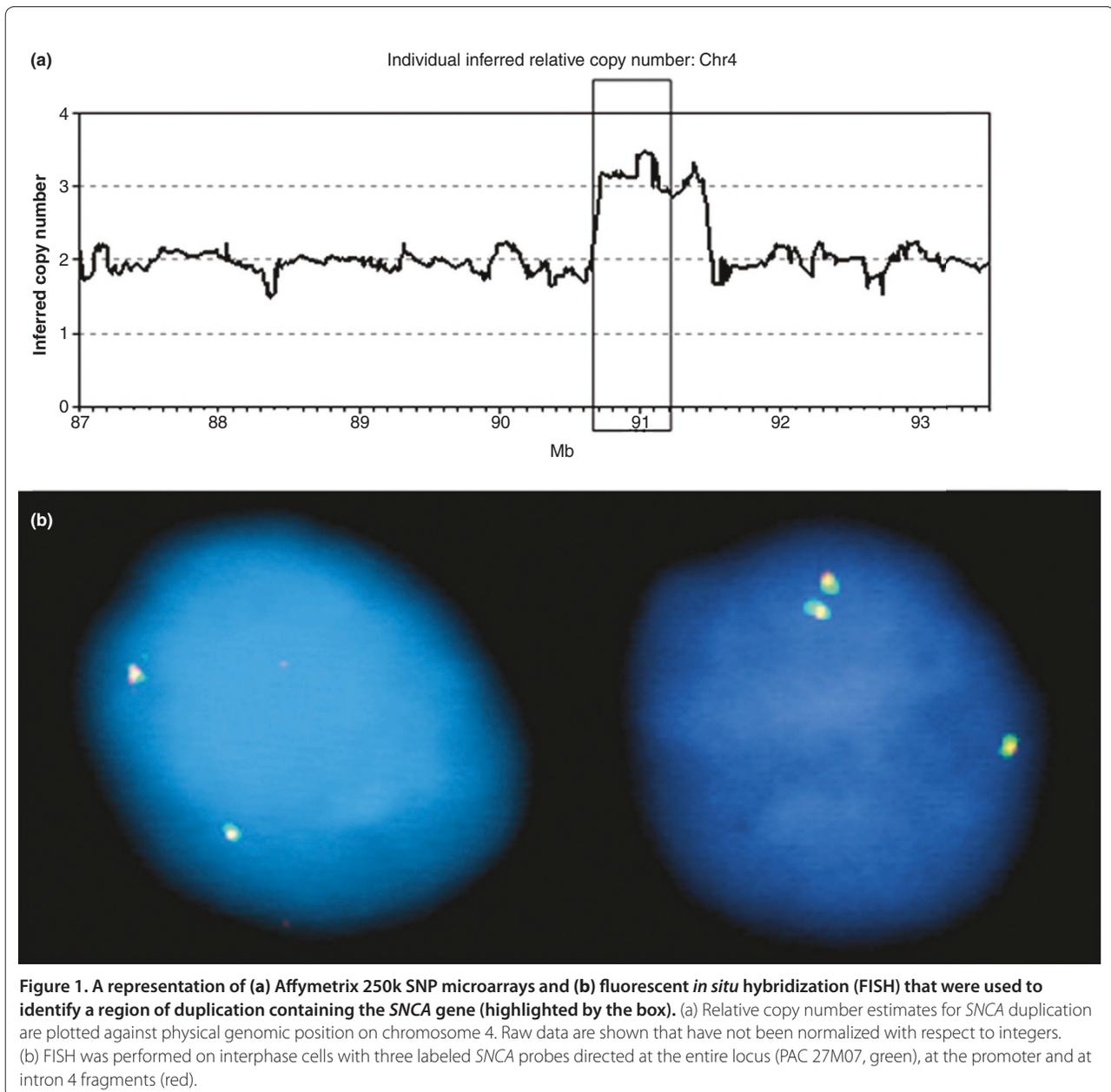
PD is a common neurodegenerative disorder affecting approximately 1% of the population aged over 60 years. The disease presents clinically as a movement disorder characterized by tremor at rest, bradykinesia, rigidity and postural instability. Neuropathological changes leading to a diagnosis of PD are dopaminergic cell loss in the substantia nigra accompanied by the formation of Lewy bodies, which are intracytoplasmic protein aggregates within the remaining neurons. The neuropathology associated with PD progresses over time, and in more advanced stages, patients develop a range of non-motor symptoms, including cognitive decline. Although drugs such as levodopa or surgical intervention (deep brain stimulation) can help alleviate the motor symptoms, they do not halt disease progression and are not effective against the non-motor aspects of the disease, such as rapid eye movement sleep behavior disorder, constipation, depression and cognitive decline.

A major breakthrough in recent years has been the mapping of chromosomal loci linked to familial PD and the cloning of five genes causing monogenic forms of the syndrome [5]. Mendelian forms of PD are relatively rare and the genetic mutations account for only a small fraction of affected individuals; however, common variation at some of these loci has been shown to confer population disease risk in association-based studies. The identification of CNVs in several of these genes has provided mechanistic insights into PD pathogenesis, generated *in vivo* model systems and driven the development of novel therapeutic intervention strategies.

*Correspondence: ross.owen@mayo.edu

²Department of Neuroscience, Mayo Clinic Jacksonville, 4500 San Pablo Road South, Jacksonville, FL 32224, USA

Full list of author information is available at the end of the article



Pathogenic copy number variation in autosomal dominant Parkinson's disease

A seminal discovery in the study of PD was the report of missense mutations in the *SNCA* gene, encoding the α -synuclein protein, in dominantly inherited disease [6]. This was the first evidence that genomic variation could result in an inherited form of PD. The subsequent identification of α -synuclein as the major protein component of the pathological substrate (Lewy bodies) placed α -synuclein at the center of PD research. However, the missense mutations reported in *SNCA* provided limited insight into the pathological mechanisms involved.

In 2003, genomic multiplications of chromosome 4q21-22 containing the *SNCA* locus were shown to cause familial PD (Figure 1). Genomic triplication of the gene causes a rapidly progressive form of PD [4]. In contrast, the clinical phenotype of families with a duplication of *SNCA* resembles idiopathic PD with late age at onset and slower disease progression and without early development of dementia. Several asymptomatic duplication carriers over 70 years of age without any signs of PD have recently been identified, indicating a reduced penetrance of disease [7]. Segmental intra-allelic duplication and interallelic recombination with unequal crossing-over

both seem to be responsible for *SNCA* multiplications [8]. Measurements of protein levels in triplication carriers confirmed the predicted doubling of α -synuclein protein in blood and increased levels of the protein in the brain. Therefore, even without sequence mutants, increased wild-type α -synuclein dosage may cause PD.

Recent studies have confirmed the association of common non-coding genetic variants at the *SNCA* locus and increased risk of PD [9]. These findings suggest that variation in regions of *SNCA*, most likely deregulating constitutive gene overexpression, may provide a therapeutic target to a substantial proportion of patients with the more frequent sporadic form of PD. Currently, studies are underway to generate *in vivo* *SNCA* multiplication model systems and identify sensitive and specific downregulators of *SNCA* gene and protein expression.

Recessive forms of PD

Whereas the *SNCA* story suggests a gain of function, several early-onset forms of PD-like disorders have demonstrated the role of loss of function in the disease. The Parkin (*PRKN*) gene is one of the largest known genes, spanning a 1.4 Mb genomic region. *PRKN* mutations are the most common cause of early-onset PD identified so far and are particularly frequent in individuals with evidence of recessive inheritance. The initial pathogenic *PRKN* mutations identified were large homozygous genomic deletions, and more than 100 different pathogenic variants in this gene have been published, including deletions, multiplications and missense and nonsense mutations [10]. The *PRKN* gene is located in one of several genomic regions of very high deletion frequency ('hotspots'), where independent rare deletions are found at frequencies of up to 100-fold higher than the average for the genome as a whole [11]. Approximately a third of all pathogenic *PRKN* variants are CNVs occurring between exons 2 and 5, forming a recombination hot spot [10].

Pathogenic mutations in the PTEN-induced kinase (*PINK1*) gene are much less common than *PRKN* mutations and probably account for only 1 to 2% of early-onset PD patients. CNVs in families with PD have been reported, including a deletion of the entire *PINK1* gene [12]. This deletion also partly involved two neighboring genes, and two highly similar *Alu* repeat sequences within these genes were found, which enclose the putative breakpoints. It is likely that the deletion resulted from an unequal crossing-over between these two sequences. Furthermore, a homozygous deletion of *PINK1* exons 4 to 8 was found in three affected siblings from a consanguineous Sudanese family with early-onset PD [13]. Breakpoint analysis revealed a complex rearrangement combining a large deletion and the insertion of a sequence duplicated from the neighboring dolichyl-

diphosphooligosaccharide-protein glycosyltransferase (*DDOST*) gene.

The *PARK7* locus on chromosome 1p36 was localized by homozygosity mapping in two consanguineous families from genetically isolated communities in the Netherlands and Italy. In one of the families, a 14 kb deletion involving the first five of seven exons in the *DJ-1* gene was identified [14]. The DJ-1 protein is involved in oxidative stress response and may be targeted to the mitochondria. *Alu* repeat elements flank the deleted sequence on both sides, suggesting that unequal crossing-over was likely at the origin of this genomic rearrangement. Very few mutations in this gene have been reported, and mutation of *DJ-1* accounts for less than 1% of early-onset PD cases. A heterozygous duplication involving the first five exons of *DJ-1* in a single patient is the only other CNV published so far [15].

Genome-wide association studies

Genome-wide association studies (GWASs) have helped to elucidate the genetic basis for a number of complex disorders. GWASs use microarrays with up to one million single nucleotide polymorphisms (SNPs) to capture a significant amount of an individual's genetic variation. In addition to identification of SNPs associated with disease, these studies also allow quantitative assessment of the genome. A total of five GWASs of PD have been published so far. However, analyses of structural genetic variation have been published from only a relatively small study of 273 patients [16]. No new regions associated with PD were identified, but several deletions and duplications in the *PRKN* gene were observed. The potential power of this approach in PD and other neurodegenerative disorders has yet to be fully used.

Conclusions

A recent study of eight common diseases (including coronary artery disease, rheumatoid arthritis and diabetes) concluded that common CNVs are unlikely to have a major pathogenic role [17]. However, this does not exclude the possible existence of several rare CNVs causing the same disorders in a proportion of patients. The number of individuals carrying a given CNV at a known gene locus is frequently small. Nevertheless, discovery of these changes by GWASs or family-based studies are important.

Limited data from GWASs have been published on the extent of CNVs in PD, even though previous family-based studies have highlighted CNVs in both familial and sporadic PD. A large proportion of functional and translational PD research today is based on hypotheses provided by the identification of mutations in familial forms of the disorder, including CNVs. The most common genetic cause of PD is point mutations in the leucine-rich

repeat kinase-2 (*LRRK2*) gene and CNVs have yet to be identified at this locus. The hypothesized toxic gain of function (kinase activity) would be supported if *LRRK2* multiplications are identified. The *SNCA* story in PD has shown how the identification of a single CNV can provide insights into pathogenic mechanisms and drive therapeutic development, and although it remains unclear what proportion of genetic disease is caused by CNVs, it is likely that many of those affecting risk of PD are still to be identified.

Abbreviations

CNV, copy number variant; GWAS, genome-wide association study; PD, Parkinson's disease; SNP, single nucleotide polymorphism.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MT and OAR drafted the manuscript. Both authors read the final manuscript and agreed its content before publication.

Acknowledgements

Work in the authors' laboratory was supported by grants from the South-Eastern Norway Regional Health Authority and the family of Carl and Susan Bolch.

Author details

¹Department of Neurology, Oslo University Hospital, Rikshospitalet, Sognsvannsveien, 0027 Oslo, Norway. ²Department of Neuroscience, Mayo Clinic Jacksonville, 4500 San Pablo Road South, Jacksonville, FL 32224, USA.

Published: 6 September 2010

References

- Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M: **Large-scale copy number polymorphism in the human genome.** *Science* 2004, **305**:525-528.
- Hastings PJ, Lupski JR, Rosenberg SM, Ira G: **Mechanisms of change in gene copy number.** *Nat Rev Genet* 2009, **10**:551-564.
- Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A, Dumanchin C, Feuillette S, Brice A, Vercelletto M, Dubas F, Frebourg T, Campion D: **APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy.** *Nat Genet* 2006, **38**:24-26.
- Singleton AB, Farrer M, Johnson J, Singleton A, Hague S, Kachergus J, Hulihan M, Peuralinna T, Dutra A, Nussbaum R, Lincoln S, Crawley A, Hanson M, Maraganore D, Adler C, Cookson MR, Muentner M, Baptista M, Miller D, Blacato J, Hardy J, Gwinn-Hardy K: **alpha-Synuclein locus triplication causes Parkinson's disease.** *Science* 2003, **302**:841.
- Farrer MJ: **Genetics of Parkinson disease: paradigm shifts and future prospects.** *Nat Rev Genet* 2006, **7**:306-318.
- Polymeropoulos MH, Lavedan C, Leroy E, Ide SE, Dehejia A, Dutra A, Pike B, Root H, Rubenstein J, Boyer R, Stenroos ES, Chandrasekharappa S, Athanassiadou A, Papapetropoulos T, Johnson WG, Lazzarini AM, Duvoisin RC, Di Iorio G, Golbe LI, Nussbaum RL: **Mutation in the alpha-synuclein gene identified in families with Parkinson's disease.** *Science* 1997, **276**:2045-2047.
- Nishioka K, Ross OA, Ishii K, Kachergus JM, Ishiwata K, Kitagawa M, Kono S, Obi T, Mizoguchi K, Inoue Y, Imai H, Takanashi M, Mizuno Y, Farrer MJ, Hattori N: **Expanding the clinical phenotype of SNCA duplication carriers.** *Mov Disord* 2009, **24**:1811-1819.
- Ross OA, Braithwaite AT, Skipper LM, Kachergus J, Hulihan MM, Middleton FA, Nishioka K, Fuchs J, Gasser T, Maraganore DM, Adler CH, Larvor L, Chartier-Harlin MC, Nilsson C, Langston JW, Gwinn K, Hattori N, Farrer MJ: **Genomic investigation of alpha-synuclein multiplication and parkinsonism.** *Ann Neurol* 2008, **63**:743-750.
- Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, Paisan-Ruiz C, Lichtner P, Scholz SW, Hernandez DG, Kruger R, Federoff M, Klein C, Goate A, Perlmutter J, Bonin M, Nalls MA, Illig T, Gieger C, Houlihan H, Steffens M, Okun MS, Racette BA, Cookson MR, Foote KD, Fernandez HH, Traynor BJ, Schreiber S, Arepalli S, Zonozzi R, et al.: **Genome-wide association study reveals genetic risk underlying Parkinson's disease.** *Nat Genet* 2009, **41**:1308-1312.
- Hedrich K, Eskelson C, Wilmot B, Marder K, Harris J, Garrels J, Mejia-Santana H, Vieregge P, Jacobs H, Bressman SB, Lang AE, Kann M, Abbruzzese G, Martinelli P, Schwinger E, Ozelius LJ, Pramstaller PP, Klein C, Kramer P: **Distribution, type, and origin of Parkin mutations: review and case studies.** *Mov Disord* 2004, **19**:1146-1157.
- Bradley WE, Raelson JV, Dubois DY, Godin E, Fournier H, Prive C, Allard R, Pinchuk V, Lapalme M, Paulussen RJ, Belouchi A: **Hotspots of large rare deletions in the human genome.** *PLoS ONE* 2010, **5**:e9401.
- Marongiu R, Brancati F, Antonini A, Ialongo T, Ceccarini C, Scarciolla O, Capalbo A, Benti R, Pezzoli G, Dallapiccola B, Goldwurm S, Valente EM: **Whole gene deletion and splicing mutations expand the PINK1 genotypic spectrum.** *Hum Mutat* 2007, **28**:98.
- Cazeneuve C, San C, Ibrahim SA, Mukhtar MM, Kheir MM, Leguern E, Brice A, Salihi MA: **A new complex homozygous large rearrangement of the PINK1 gene in a Sudanese family with early onset Parkinson's disease.** *Neurogenetics* 2009, **10**:265-270.
- Bonifati V, Rizzu P, van Baren MJ, Schaap O, Breedveld GJ, Krieger E, Dekker MC, Squitieri F, Ibanez P, Joosse M, van Dongen JW, Vanacore N, van Swieten JC, Brice A, Meco G, van Duijn CM, Oostra BA, Heutink P: **Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism.** *Science* 2003, **299**:256-259.
- Macedo MG, Verbaan D, Fang Y, van Rooden SM, Visser M, Anar B, Uras A, Groen JL, Rizzu P, van Hilten JJ, Heutink P: **Genotypic and phenotypic characteristics of Dutch patients with early onset Parkinson's disease.** *Mov Disord* 2009, **24**:196-203.
- Simon-Sanchez J, Scholz S, Matarin Mdel M, Fung HC, Hernandez D, Gibbs JR, Britton A, Hardy J, Singleton A: **Genomewide SNP assay reveals mutations underlying Parkinson disease.** *Hum Mutat* 2008, **29**:315-322.
- Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, Vukcevic D, Barnes C, Conrad DF, Giannoulatos E, Holmes C, Marchini JL, Stirrups K, Tobin MD, Wain LV, You C, Aerts J, Ahmad T, Andrews TD, Arbury H, Attwood A, Auton A, Ball SG, Balmforth AJ, Barrett JC, Barroso I, Barton A, Bennett AJ, Bhaskar S, Blaszczyk K, et al.: **Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls.** *Nature* 2010, **464**:713-720.

doi:10.1186/gm183

Cite this article as: Toft M, Ross OA: Copy number variation in Parkinson's disease. *Genome Medicine* 2010, **2**:62.