JAMA Neurol. Author manuscript; available in PMC 2015 March 17.

Published in final edited form as:
JAMA Neurol. 2014 September ; 71(9): 1162–1171. doi:10.1001/jamaneurol.2014.994.

A 6.4MB duplication of the alpha-synuclein locus causing fronto-temporal dementia and parkinsonism - phenotype-genotype correlations

Eleanna Kara, MD, MSc[1], **Aoife P Kiely, PhD**[1,2], **Christos Proukakis, MD, PhD**[3], **Nicola Giffin, MD**[5], **Seth Love, MD**[6], **Jason Hehir, BSc**[1], **Khadija Rantell, PhD**[7,8], **Amelie Pandraud, MSc**[1], **Dena G Hernandez, MS**[1,9], **Elizabeth Nacheva, MD, PhD**[10], **Alan M Pittman, PhD**[1], **Mike A Nalls, PhD**[9], **Andrew B Singleton, PhD**[9], **Tamas Revesz, MD**[1,2], **Kailash P Bhatia, MD**[4], **Niall Quinn, MD**[4], **John Hardy, PhD**[1], **Janice L Holton, MD, PhD**[1,2], and **Henry Houlden, MD, PhD**[1,11,*]

[1]Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK

[2]The Queen Square Brain Bank, UCL Institute of Neurology, London, UK

[3]Department of Clinical Neuroscience, UCL Institute of Neurology, London, UK

[4]Sobell Department of Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, London, UK

[5]Department of Neurology, Frenchay Hospital, Bristol, UK

[6]Department of Neuropathology, Frenchay Hospital, Bristol, UK

[7]Biomedical Research Centre, UCL, London, UK

[8]Education Unit, UCL Institute of Neurology, Queen Square, London, UK

[9]Laboratory of Neurogenetics, NIA, NIH, Bethesda, USA.

[10]Department of Academic Haematology, Royal Free Campus, UCL

[11]MRC Centre for Neuromuscular Diseases, UCL Institute of Neurology, London, UK

Abstract

Importance—*SNCA* locus duplications are associated with variable clinical features and reduced penetrance but the reasons underlying this variability are unknown.

[*]**Corresponding author:** Henry Houlden, Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, WC1N 3BG, London, UK. Tel: +44 2034484068. h.houlden@ucl.ac.uk.

AUTHOR CONTRIBUTIONS:

Study concept and design: Kara, Rantell, Singleton, Hardy, Holton, Houlden
Acquisition of data: Kara, Kiely, Proukakis, Giffin, Love, Hehir, Pandraud, Hernandez, Nacheva, Bhatia, Quinn
Analysis and interpretation of data: Kara, Kiely, Proukakis, Giffin, Love, Hehir, Rantell, Hernandez, Nacheva, Pittman, Nalls, Singleton, Revesz, Bhatia, Quinn, Hardy, Holton, Houlden
Drafting of the manuscript: Kara, Kiely, Holton, Houlden
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Kara, Rantell
Obtained funding: Proukakis, Singleton, Revesz, Hardy, Holton, Houlden
Administrative, technical, and material support: Hehir, Rantell, Pandraud, Hernandez, Nacheva, Pittman, Nalls, Singleton
Study supervision: Proukakis, Love, Rantell, Hernandez, Pittman, Nalls, Singleton, Revesz, Bhatia, Quinn, Hardy, Holton, Houlden

Europe PMC Funders Author Manuscripts

Objective—1) To report a novel family carrying a heterozygous 6.4Mb duplication of the *SNCA* locus with an atypical clinical presentation strongly reminiscent of frontotemporal dementia (FTD) and late-onset pallidopyramidal syndromes. 2) To study phenotype-genotype correlations in *SNCA* locus duplications.

Design, Setting, Participants and Data sources—We report the clinical and neuropathologic features of a family carrying a 6.4Mb duplication of the *SNCA* locus. To identify candidate disease modifiers, we undertake a genetic analysis in the family and conduct statistical analysis on previously published cases carrying *SNCA* locus duplication using regression modelling with robust standard errors to account for clustering at the family level.

Main outcome measures—To assess whether length of the *SNCA* locus duplication influences disease penetrance and severity, and whether extra-duplication factors have a disease-modifying role.

Results—We identified a large 6.4Mb duplication of the *SNCA* locus in this family. Neuropathological analysis showed extensive α-synuclein pathology with minimal phospho-tau pathology. Genetic analysis showed an increased burden of PD-related risk factors and the disease-predisposing H1/H1 *MAPT* haplotype. Statistical analysis of previously published cases suggested that there is a trend towards increasing disease severity and disease penetrance with increasing duplication size. The corresponding odds ratios (95% CI) from the univariate analyses were 1.17 (0.81 to 1.68) and 1.34 (0.78 to 2.31) respectively. Gender was significantly associated with both disease risk and severity; males compared to females had increased disease risk and severity and the corresponding odds ratios (95% CI) from the univariate analyses were 8.36 (1.97 to 35.42) and 5.55 (1.39 to 22.22) respectively.

Conclusions and relevance—These findings further expand the phenotypic spectrum of *SNCA* locus duplications. Increased dosage of genes located within the duplicated region probably cannot increase disease risk and disease severity without the contribution of additional risk factors. Identification of disease modifiers accounting for the substantial phenotypic heterogeneity of patients with *SNCA* locus duplications could provide insight into molecular events involved in α-synuclein aggregation.

Keywords

fronto-temporal dementia and parkinsonism; alpha-synuclein; duplication; gene

INTRODUCTION

α-Synuclein is a protein central to the pathogenesis of Parkinson's disease (PD), a role illustrated by the identification of α-Synuclein as the principal component of Lewy Bodies (LBs) [1], the pathological hallmark of PD, and of α-*Synuclein (SNCA)* point mutations and multiplications as rare causes of PD [2-4]. Disease severity correlates with the number of *SNCA* alleles in multiplication carriers. Patients with *SNCA* triplication present with an early-onset, aggressive form of PD, with dementia, psychiatric features and dysautonomia [5]. On the other hand, *SNCA* duplications are not fully penetrant and are associated with variable clinical features ranging from late-onset sporadic PD to presentations

indistinguishable from those in triplication carriers but the reasons underlying this variability are unknown.

Here, we report a new kindred carrying the largest non-chromosomal duplication of the *SNCA* locus identified to date presenting with clinical features reminiscent of frontotemporal dementia (FTD) and widespread Lewy bodies. We undertake a comprehensive assessment of candidate disease modifiers in carriers of *SNCA* locus duplications based on genetic analysis of the proband and on statistical analysis of previously published cases.

METHODS

Genetics

MAPT was screened through Sanger sequencing in cases III:1 and III:3 (figure 1A) because of clinical similarity to FTD [6]. Because of clinical similarities to *SNCA* multiplication cases and widespread α-synuclein pathology, a genomic DNA sample from case III:1 was assessed for whole gene *SNCA* multiplications with the Multiplex-ligation dependent probe amplification (MLPA) kit P051C (MRC Holland, the Netherlands) [7]. The approximate breakpoints of the duplication were determined through an Agilent custom 8×60k dense comparative genomic hybridization (CGH) array designed as part of a larger on-going study with dense spacing of probes throughout the *SNCA* genomic region (CP, details available on request).

In an attempt to identify risk factors potentially influencing the clinical features, a DNA sample from case III:1 was genotyped on the Immunochip, an Illumina array of custom content with a specific focus on immune-related and PD/parkinsonism genes [8-10]. We focused on risk variants identified through previously published PD genome wide association studies (GWAS) and on mutations and risk factors located within PD genes (see supplementary information for methods).

Neuropathology

The brain from case III:3 was fixed by suspension in buffered 20% formalin and sampled extensively for neuropathology. Paraffin-embedded sections (8μm) from cortical, subcortical, brainstem and cerebellar regions were stained immunohistochemically for standard pathological markers including Aβ (DAKO, M0872) AT8 (Tau Ser202/Thr205, Source Bioscience 90206), TDP-43 (Abnova, H00023435-M01) and α-synuclein (Abcam ab15530). For Aβ and α-synuclein immunohistochemistry, sections required pre-treatment with formic acid for 10 minutes before antigen retrieval. Antigen retrieval was achieved by heating sections in a pressure cooker for 10 minutes in boiling citrate buffer (pH 6.0). Endogenous peroxidase activity was blocked with methanol/0.3% H_2O_2 followed by incubation in dried milk solution (10% in PBS, 30 minutes) to block non-specific antibody binding. Primary antibodies diluted in PBS were applied for 60 minutes at room temperature and staining was visualised by the streptavidin-biotin-peroxidase method (Vector laboratories, Peterborough, UK), using 3,3′-diaminobenzidine as the chromogen with Mayer's haematoxylin counterstaining.

Europe PMC Funders Author Manuscripts

Statistical analysis of previously published cases

We reviewed previously published articles on carriers of *SNCA* locus duplications in order to assess whether the length of the duplication influences penetrance and disease severity and whether longer duplication sizes are related to a younger age at onset (AAO), indicating a contribution of increased dosage of other gene(s) to the development of disease. We used a duplication size threshold of 5Mb as an inclusion criterion which was chosen on the basis of the following arguments. Firstly, we hypothesised that shorter duplications are more frequently non-penetrant and that as duplication size increases, the individual exhibits a more severe phenotype ranging from unaffected, sporadic PD, PD-Dementia, to indistinguishable from triplication carriers. Secondly, we also hypothesised that, above a certain duplication length threshold, the individual will exhibit a more complex phenotype with features unusual for *SNCA* multiplications. Consequently, two individuals with such complex presentations, one described by Garraux et al (2012) [11] and the case reported herein carrying duplications larger than 5Mb were excluded from this analysis. Published individuals were also excluded if any of the following were true: a) there was no information on the length of the duplication, b) there was insufficient clinical information, c) clinical presentation was clearly inconsistent with parkinsonism, d) they were affected relatives of probands with *SNCA* duplications but their carrier status wasn't assessed (efigure 1). Data on all eligible individuals from each family carrying the duplications were included in the study to maximise study power (eTable 1). Review of the articles' eligibility and cataloguing of the information was undertaken by the main author (EK) and was repeated twice.

Disease severity was measured by deriving a composite score for each affected individual from the published data. Each individual was assigned a score according to a clinical presentation point system as follows: Typical PD/parkinsonism = 1, young onset (before or at 40 years) = 1, dementia = 1, hallucinations = 1, dysautonomia = 1, depression = 1, any additional features (dystonia, epilepsy, sleep disturbances, myoclonus, psychiatric features) = 1. These scores were then summed up to form a composite score, which ranged from 0 (unaffected) to 6 (severely affected).

Numerical data were summarised using mean and standard deviation or median and range depending on data distribution. Categorical data were summarised using count and percentages. We used Spearman's rank correlation to quantify strength of monotonic association between duplication size and each of AAO and disease severity. The main outcomes of interest were disease status (binary), disease severity (ordinal) and age at onset (binary). We modelled binary outcomes using logistic regression with robust sandwich estimation of the variance to account for the clustering effect at the family level [12]. We modelled ordinal outcomes using ordinal regression with the proportional odds assumption, adjusted for clusters at the family level [13]. This approach has been shown to be more suitable than multilevel approaches in the presence of small number of clusters with low number of individuals [14]. We investigated the individual effect of the covariates on each of the outcomes using univariate analyses. We also investigated their combined effect using multivariable analyses.

Europe PMC Funders Author Manuscripts

The functional form of the relationship between duplication size and each outcome was also evaluated graphically. Model fit was assessed using the Akaike Information Criterion (AIC) [15], where smaller AIC is preferred.

All statistical analyses were carried out using Stata [16]. Results from modelling are presented as estimates (95 % confidence interval). Plots were generated with the statistical package R version 3.0.2 [17] using the package ggplot2 [18].

RESULTS

CASE REPORT

Clinical description—The proband, an English, Caucasian woman (patient III:3, figure 1A), first developed lifelong and progressive symptoms of extreme anxiety, panic disorder and hallucinations from age 8 that were managed adequately with medication up to her mid-20's when she had to quit her job. At age 38 years, she developed right arm tremor followed by a progressive akinetic-rigid syndrome, plus a classic pill-rolling tremor, complicated by worsening obsessive behaviour, walking disturbances and falls, increased salivation, REM sleep behaviour disorder, personality changes, short-term memory impairment and poor self-care. She also had facial, head and tongue tremor accompanied by blepharospasm, dystonic neck and dysarthria. Levodopa treatment improved her walking but increased her falls. Neuropsychometry testing showed frontal and temporal deficits though neuroimaging was normal. There was a history of recurrent unexplained blackouts or seizures. EEG showed traces of alpha-rhythm, and autonomic testing indicated cardiovascular autonomic failure. On examination, she exhibited akinesia and rigidity of all four limbs, rest tremor and hypometric saccades. From the age of 42, the motor and cognitive aspects of her disorder declined further, with poor concentration, frontal release signs, wandering, repetitive speech and a profound increase in appetite particularly for sweet food, and at age 46 she was given a clinical diagnosis of possible FTDP-17 but no *MAPT* mutations were detected. She became bedbound and died at the age of 49 years.

In her family history, her father (II:3), paternal grandmother (I:1) and two paternal great-aunts (I:2, I:3) suffered from PD with no documented dementia. Her paternal cousin (III:1) was diagnosed as having possible FTDP-17 in his early 50's when he presented with memory problems, hallucinations and falls, complicated by alcoholism. This progressed with increasing cognitive decline and disorientation and rigidity at which point his progressive condition was managed by local services.

Pathology—Neuropathological examination showed neuronal loss which was most severe in the substantia nigra, moderate in the locus coeruleus and dorsal motor nucleus of the vagus and mild in the nucleus basalis of Meynert and cerebellar Purkinje cells. Neuronal loss was not observed in other regions including the hippocampus. Widespread Lewy bodies and Lewy neurites affecting brainstem, limbic and neocortical regions corresponding to neocortical Lewy body pathology, Braak stage 6[19,20] (figure 1B, a-d) were detected. The hippocampus showed mild accumulation of α–synuclein inclusions with severe neuritic pathology in the CA2 region (figure 1B, e, f). Occasional oligodenrdrocytes contained α–synuclein immunoreactive inclusions with coiled body morphology. No inclusions

resembling glial cytoplasmic inclusions were observed (figure 1B, g, h). Minimal phospho-tau pathology was present in the transentorhinal cortex corresponding to Braak and Braak stage I (figure 1B, I, j) and there were sparse diffuse cortical amyloid β deposits. TDP-43 pathology was not present.

Genetics—Patient III:1 was negative for *MAPT* point mutations but was found to carry a heterozygous *SNCA* duplication on MLPA (figure 2A). CGH arrays showed that the duplication extended over approximately 6.4Mb (chr4:88349207-94751141, corresponding to the first and the last duplicated probes respectively, human build 37) and contained 37 genes including *SNCA* (figure 2B,C)(table 1). The first and last non-duplicated probes are A_14_P119434 (chr4:88295927-88295986) and A_14_P124103 (chr4:94859877-94859936) respectively (figure 2D). The breakpoints are located centromeric to intron 1 of the *NUDT9* gene (NM_198038.2), and telomeric to the *ATOH1* gene (figure 2E).

The patient was homozygous for the wild type allele for all rare SNPs or probes encoding mutations on the Immunochip apart from 1 heterozygous *parkin* variant that is probably non-pathogenic (rs1801334) (MAF 0.04 in European populations as catalogued in ensembl) (eTable 2) and had a *MAPT* haplotype of H1/H1. The patient carried a total of 18/46 risk alleles for 23 GWAS SNPs (eTable 3) placing him in the 3rd risk quintile shown to confer approximately a 77% increase in PD risk[21-23].

Statistical analysis—Table 2 describes the demographic and clinical characteristics of individuals included in the statistical analysis. Overall, 73% (27/37) of individuals were affected, with a higher proportion of females compared to males in the unaffected group (90% versus 10 %). The median (range) age at onset of disease was 47 (31-71) years. The median (range) composite score was 2 (0-6). The median (range) duplication size for affected and unaffected individuals was 0.63 (0.2-5) and 0.6 (0.4-3.47) Mb respectively. There was a weak positive monotonic association between duplication size and composite score (rho=0.17, p=0.314) and between duplication size and age at onset (rho=0.12, p=0.560).

Table 3 provides estimates of the effect of duplication size and gender on disease status, age at onset and disease severity. Overall, only the effect of gender on disease status and disease severity reached statistical significance for both univariate and multivariate models ($p<0.05$). The odds of being affected were more than 8-fold for males compared to females. The corresponding OR (95% CI) was 8.36 (1.97 to 35.42). The odds of being affected increased by approximately 34% for each unit increase in duplication size. The corresponding OR (95% CI) was 1.34 (0.78 to 2.31). Males had a 77% greater chance of developing an early onset form of disease in comparison to females with an OR of 0.23 (95% CI: 0.04 to 1.28), whereas increase in duplication size by one unit resulted in a 6% decrease in the chance of developing late-onset disease (OR, 95% CI: 0.94, 0.71 to 1.24). Males had over 5-fold increase in the chance of developing a more severe disease, and disease severity increased by 17% for each unit increase in duplication size. The corresponding OR (95% CI) were 5.55 (1.39, 22.22) and 1.17 (0.81, 1.68) respectively. (All ORs quoted in-text refer to univariate models).

DISCUSSION AND CONCLUSION

A total of 29 kindreds carrying *SNCA* locus duplications have been reported in the literature [24], with pathology information on 4 duplication patients (3 belonging to the same family) [25-29] (eTable 4). The size of the duplicated region has been studied in 16 kindreds/sporadic cases (eTable 4) and has been found to vary greatly, from 0.2Mb [30]to 41.2Mb [11] containing from 2 to 150 genes (figure 2E,F, eTable 5). The case described here carries the largest 'submicroscopic' duplication of the *SNCA* locus reported to date.

Even though we were not able to determine the exact breakpoints of the duplication, in accordance with Ibanez et al (2009) [31], we showed that these lay within regions rich in repetitive elements and especially transposable elements (long interspersed elements-LINEs, short interspersed elements-SINEs, and long terminal repeat-LTR retrotransposons) facilitating the insertion of additional copies of flanking sequences in the genome [32], allowing non-homologous recombination events leading to genomic structural rearrangements. The 4q22 region is known to be inherently prone to disruption as a fragile site (FRA4F), which interestingly was originally described as a 7 Mb region corresponding very closely to the duplicated region in this report [33]. Subsequent work has extended the fragile site from 88200000-99100000Mb [34] but the centromeric boundary still corresponds to the centromeric breakpoint in our report. Other reported copy number variants (CNVs) are mostly within the confines of the fragile site, although a breakpoint centromeric to the proposed FRA4F boundary has been documented in some cases (figure 2F, G).

This report further emphasises the clinical variability and expands the phenotypic spectrum associated with *SNCA* duplications and represents the first case with such a complex presentation and substantial overlap with other dementing and psychiatric disorders. *SNCA* duplications are characterised by incomplete penetrance and variable clinical presentation [27,35] occasionally being indistinguishable from sporadic PD [31] or having the core clinical features of young-onset PD-dementia (YOPD-dementia) with dysautonomia, depression and hallucinations, resembling the phenotype of *SNCA* triplications [5]. Our case exhibits some of the main features associated with *SNCA* multiplications (YOPD-dementia, hallucinations, dysautonomia) which are, however, part of a more complicated phenotype reminiscent of FTD. Obsessive compulsive disorder (OCD) and anxiety have not been previously reported in any patient with an *SNCA* duplication, nor was it present in any other member of our family. Thus although it is possible these were causally related in our case III:3, it is more likely a coincidental occurrence. However, this does make it difficult to establish with certainty the exact age at onset of the patient (possibly age 8 or probably age 38 years). Intrafamilial variability is marked as two members had an FTD-like phenotype and 4 family members from older generations apparently had typical PD. CaseII:1, an obligate carrier, was an alcoholic, as was his son III:1, and died at age 70 without documented parkinsonism or dementia.

Neuropathological examination showed features corresponding to idiopathic PD and *SNCA* duplication cases, with widespread Lewy pathology extending into neocortical regions (Braak stage 6), and no TDP-43 pathology [25,27-29]. There was only minimal Alzheimer pathology with tau pathology limited to Braak and Braak stage I and sparse diffuse amyloid

β deposition in the cortex. Although previous studies have suggested 'cross-seeding' between α-synuclein and tau [36], the relatively low accumulation of tau compared to α-synuclein containing inclusions suggests that this tau accumulation was a normal consequence of aging. Unlike other reported cases of *SNCA* duplication, neuronal loss was restricted and did not affect the hippocampus [25,28]. In further contrast, oligodendroglial inclusions with coiled body morphology were observed, but there were no structures resembling glial cytoplasmic inclusions[25]. However, as has been reported, hippocampal pathology, particularly neuritic α–synuclein deposition was severe, especially in the CA2 region.

Given that *SNCA* duplications are frequently non-penetrant [24,27] it is possible that mutations in another gene are responsible for the unusual clinical presentation of the proband. However, there are several lines of evidence supporting the pathogenicity of the duplication: the widespread Lewy body pathology and other pathological features are consistent with *SNCA* duplications, the pattern of inheritance of the disease in the family is autosomal dominant with possible reduced penetrance, and other affected relatives have phenotypes that have been previously associated with *SNCA* duplication. Thus, even though we were unable to confirm complete segregation due to the lack of DNA samples, the *SNCA* duplication is most likely to be responsible for the disease in this family.

The reasons underlying this clinical variability associated with *SNCA* duplication are currently unknown, although several preliminary hypotheses emerge from our analyses that require validation in future studies on larger cohorts of *SNCA* duplication carriers. We found trends for association between duplication size and disease status, age at onset and disease severity that didn't reach statistical significance though association with gender did. Even though our study was underpowered to detect association (n=37) [37,38], it is possible that larger duplication size may not independently produce disease without the contribution of additional risk factors. Gender-specific risk factors are an intriguing possibility as regression model fitting was improved after adjusting for gender (table 3), and males with *SNCA* duplications had an 8-fold increase in risk for developing PD. A predilection for male involvement is well known for sporadic PD [39,40]. Alternatively, an increased burden of other risk factors could have a disease-modifying role; the patient we report carried one single heterozygous *parkin* rare variant, had an H1H1 *MAPT* haplotype and an accumulation of risk alleles in loci identified through GWAS. Similarly, Itokawa et al have reported a patient with a *SNCA* duplication carrying a PD-risk factor (p.G2385R) within *LRRK2* [35]. Other unexplored possibilities include increased dosage of common or rare variants with a regulatory effect on expression or other properties of α-synuclein [41,42], mosaicism [43], environmental factors [44], or stochasticity [45].

In conclusion, patients in whom complex presentations of onset in early adulthood with dementia and parkinsonism complicated by multiple cognitive, psychiatric and motor features should be considered for genetic testing for *SNCA* multiplication. Understanding of the genetic modifiers influencing the phenotype of *SNCA* duplications may be important in elucidating the mechanisms underlying α-synuclein accumulation and the formation of Lewy bodies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

AKNOWLEDGMENTS

The authors would like to thank the patients and their families for their help and sample donation. This study was supported by the Parkinson's disease foundation (HH, EK), and the Wellcome Trust/MRC Joint Call in Neurodegeneration award (WT089698) to the UK Parkinson's Disease Consortium whose members are from the UCL Institute of Neurology, the University of Sheffield, and the MRC Protein Phosphorylation Unit at the University of Dundee. JHolton is supported by the Reta Lila Weston Institute for Neurological Studies, the Multiple System Atrophy Trust, Alzheimer's Research UK and Parkinson's UK. AK is supported by the Multiple System Atrophy Trust. 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 ZO1 .AG000949-08. The authors are grateful to the Queen Square Brain Bank and the South West Dementia Brain Bank (SWDBB) for provision of human tissue. This study was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. EK and HH had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Kailash P Bhatia received funding for travel from GlaxoSmithKline, Orion Corporation, Ipsen, and Merz Pharmaceuticals, LLC; serves on the editorial boards of Movement Disorders and Therapeutic Advances in Neurological Disorders; receives royalties from the publication of Oxford Specialist Handbook of Parkinson's Disease and Other Movement Disorders (Oxford University Press, 2008); received speaker honoraria from GlaxoSmithKline, Ipsen, Merz Pharmaceuticals, LLC, and Sun Pharmaceutical Industries Ltd.; personal compensation for scientific advisory board for GSK and Boehringer Ingelheim; received research support from Ipsen and from the Halley Stewart Trust through Dystonia Society UK, and the Wellcome Trust MRC strategic neurodegenerative disease initiative award (Ref. number WT089698), a grant from the Dystonia Coalition and a grant from Parkinson's UK (Ref. number G-1009). The funding organisations or sponsors weren't involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Footnotes

Conflict of interest: Kailash P Bhatia received funding for travel from GlaxoSmithKline, Orion Corporation, Ipsen, and Merz Pharmaceuticals, LLC; serves on the editorial boards of Movement Disorders and Therapeutic Advances in Neurological Disorders; receives royalties from the publication of Oxford Specialist Handbook of Parkinson's Disease and Other Movement Disorders (Oxford University Press, 2008); received speaker honoraria from GlaxoSmithKline, Ipsen, Merz Pharmaceuticals, LLC, and Sun Pharmaceutical Industries Ltd.; personal compensation for scientific advisory board for GSK and Boehringer Ingelheim; received research support from Ipsen and from the Halley Stewart Trust through Dystonia Society UK, and the Wellcome Trust MRC strategic neurodegenerative disease initiative award (Ref. number WT089698), a grant from the Dystonia Coalition and a grant from Parkinson's UK (Ref. number G-1009).

URLs

UCSC genome browser http://genome.ucsc.edu/

R http://www.r-project.org/

R package ggplot2 http://cran.r-project.org/web/packages/ggplot2/index.html

STATA package gologit2 http://www3.nd.edu/~rwilliam/gologit2/

Europe PMC Funders Author Manuscripts

REFERENCES

1. Spillantini MG, Schmidt ML, Lee VM, Trojanowski JQ, Jakes R, Goedert M. Alpha-synuclein in Lewy bodies. Nature. Aug 28; 1997 388(6645):839–840. [PubMed: 9278044]
2. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science. Jun 27; 1997 276(5321):2045–2047. [PubMed: 9197268]
3. Singleton AB, Farrer M, Johnson J, et al. alpha-Synuclein locus triplication causes Parkinson's disease. Science. Oct 31.2003 302(5646):841. [PubMed: 14593171]
4. Ibanez P, Bonnet AM, Debarges B, et al. Causal relation between alpha-synuclein gene duplication and familial Parkinson's disease. Lancet. Sep 25; Oct 25; 2004 364(9440):1169–1171. [PubMed: 15451225]
5. Gwinn K, Devine MJ, Jin LW, et al. Clinical features, with video documentation, of the original familial lewy body parkinsonism caused by alpha-synuclein triplication (Iowa kindred). Mov Disord. Sep; 2011 26(11):2134–2136. [PubMed: 21656851]
6. Kara E, Ling H, Pittman AM, et al. The MAPT p.A152T variant is a risk factor associated with tauopathies with atypical clinical and neuropathological features. Neurobiol Aging. Sep; 2012 33(9):2231e2237–2231 e2214. [PubMed: 22595371]
7. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res. Jun 15.2002 30(12):e57. [PubMed: 12060695]
8. Plagnol V, Nalls MA, Bras JM, et al. A two-stage meta-analysis identifies several new loci for Parkinson's disease. PLoS Genet. Jun.2011 7(6):e1002142. [PubMed: 21738488]
9. Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. Nov 1; 2012 491(7422):119–124. [PubMed: 23128233]
10. Faraco J, Lin L, Kornum BR, et al. ImmunoChip study implicates antigen presentation to T cells in narcolepsy. PLoS Genet. 2013; 9(2):e1003270. [PubMed: 23459209]
11. Garraux G, Caberg JH, Vanbellinghen JF, et al. Partial trisomy 4q associated with young-onset dopa-responsive parkinsonism. Arch Neurol. Mar; 2012 69(3):398–400. [PubMed: 22410449]
12. Williams RL. A note on robust variance estimation for cluster-correlated data. Biometrics. Jun; 2000 56(2):645–646. [PubMed: 10877330]
13. Harrell FE Jr. Margolis PA, Gove S, et al. WHO/ARI Young Infant Multicentre Study Group. Development of a clinical prediction model for an ordinal outcome: the World Health Organization Multicentre Study of Clinical Signs and Etiological agents of Pneumonia, Sepsis and Meningitis in Young Infants. Stat Med. Apr 30; 1998 17(8):909–944. [PubMed: 9595619]
14. Austin PC. Estimating multilevel logistic regression models when the number of clusters is low: a comparison of different statistical software procedures. Int J Biostat. 2010; 6(1) Article 16.
15. Chirwa ED, Griffiths PL, Maleta K, Norris SA, Cameron N. Multi-level modelling of longitudinal child growth data from the Birth-to-Twenty Cohort: a comparison of growth models. Ann Hum Biol. Oct 11.2013
16. StataCorp.. Stata Statistical Software: Release 12. StataCorp LP; College Station, TX: 2011.
17. Team, RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: 2013. URL http://www.R-project.org/
18. Wickham, H. ggplot2: elegant graphics for data analysis. Springer; New York: 2009.
19. Braak H, Alafuzoff I, Arzberger T, Kretzschmar H, Del Tredici K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. Acta Neuropathol. Oct; 2006 112(4):389–404. [PubMed: 16906426]
20. McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. Neurology. Dec 27; 2005 65(12):1863–1872. [PubMed: 16237129]

21. Nalls MA, Plagnol V, Hernandez DG, et al. Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet. Feb 19; 2011 377(9766):641–649. [PubMed: 21292315]
22. Hernandez DG, Nalls MA, Ylikotila P, et al. Genome wide assessment of young onset Parkinson's disease from Finland. PLoS One. 2012; 7(7):e41859. [PubMed: 22911860]
23. Kara E, Xiromerisiou G, Spanaki C, et al. Assessment of Parkinson's disease risk loci in Greece. Neurobiol Aging. Sep 27.2013
24. Elia A, Petrucci S, Fasano A, et al. Alpha-synuclein gene duplication: Marked intrafamilial variability in two novel pedigrees. Mov Disord. 2013
25. Ikeuchi T, Kakita A, Shiga A, et al. Patients homozygous and heterozygous for SNCA duplication in a family with parkinsonism and dementia. Arch Neurol. Apr; 2008 65(4):514–519. [PubMed: 18413475]
26. Wakabayashi K, Hayashi S, Ishikawa A, et al. Autosomal dominant diffuse Lewy body disease. Acta Neuropathol. Aug; 1998 96(2):207–210. [PubMed: 9705138]
27. Nishioka K, Hayashi S, Farrer MJ, et al. Clinical heterogeneity of alpha-synuclein gene duplication in Parkinson's disease. Ann Neurol. Feb; 2006 59(2):298–309. [PubMed: 16358335]
28. Obi T, Nishioka K, Ross OA, et al. Clinicopathologic study of a SNCA gene duplication patient with Parkinson disease and dementia. Neurology. Jan 15; 2008 70(3):238–241. [PubMed: 18195271]
29. Ross OA, Braithwaite AT, Skipper LM, et al. Genomic investigation of alpha-synuclein multiplication and parkinsonism. Ann Neurol. Jun; 2008 63(6):743–750. [PubMed: 18571778]
30. Nishioka K, Ross OA, Ishii K, et al. Expanding the clinical phenotype of SNCA duplication carriers. Mov Disord. Sep 15; 2009 24(12):1811–1819. [PubMed: 19562770]
31. Ibanez P, Lesage S, Janin S, et al. Alpha-synuclein gene rearrangements in dominantly inherited parkinsonism: frequency, phenotype, and mechanisms. Arch Neurol. Jan; 2009 66(1):102–108. [PubMed: 19139307]
32. Ivics Z, Izsvak Z. Repetitive elements and genome instability. Semin Cancer Biol. Aug; 2010 20(4):197–199. [PubMed: 20851335]
33. Rozier L, El-Achkar E, Apiou F, Debatisse M. Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer. Oncogene. Sep 9; 2004 23(41):6872–6880. [PubMed: 15286716]
34. Fungtammasan A, Walsh E, Chiaromonte F, Eckert KA, Makova KD. A genome-wide analysis of common fragile sites: what features determine chromosomal instability in the human genome? Genome Res. Jun; 2012 22(6):993–1005. [PubMed: 22456607]
35. Itokawa K, Sekine T, Funayama M, et al. A case of alpha-synuclein gene duplication presenting with head-shaking movements. Mov Disord. Mar; 2013 28(3):384–387. [PubMed: 23124679]
36. Guo JL, Covell DJ, Daniels JP, et al. Distinct alpha-Synuclein Strains Differentially Promote Tau Inclusions in Neurons. Cell. Jul 3; 2013 154(1):103–117. [PubMed: 23827677]
37. Bonett DG, Right TAW. SAMPLE SIZE REQUIREMENTS FOR ESTIMATING PEARSON, KENDALL AND SPEARMAN CORRELATIONS. PSYCHOMETRIKA. 2000; 65(1):23–28.
38. Hsieh FY. Sample size tables for logistic regression. Stat Med. Jul; 1989 8(7):795–802. [PubMed: 2772439]
39. Wooten GF, Currie LJ, Bovbjerg VE, Lee JK, Patrie J. Are men at greater risk for Parkinson's disease than women? J Neurol Neurosurg Psychiatry. Apr; 2004 75(4):637–639. [PubMed: 15026515]
40. Taylor KS, Cook JA, Counsell CE. Heterogeneity in male to female risk for Parkinson's disease. J Neurol Neurosurg Psychiatry. Aug; 2007 78(8):905–906. [PubMed: 17635983]
41. Spencer CC, Plagnol V, Strange A, et al. Dissection of the genetics of Parkinson's disease identifies an additional association 5′ of SNCA and multiple associated haplotypes at 17q21. Hum Mol Genet. Jan 15; 2011 20(2):345–353. [PubMed: 21044948]
42. Rhinn H, Qiang L, Yamashita T, et al. Alternative alpha-synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. Nat Commun. 2012; 3:1084. [PubMed: 23011138]

43. Proukakis C, Houlden H, Schapira AH. Somatic alpha-synuclein mutations in Parkinson's disease: Hypothesis and preliminary data. Mov Disord. May 14.2013
44. Baranzini SE, Mudge J, van Velkinburgh JC, et al. Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature. Apr 29; 2010 464(7293):1351–1356. [PubMed: 20428171]
45. Xiromerisiou G, Houlden H, Sailer A, Silveira-Moriyama L, Hardy J, Lees AJ. Identical twins with Leucine rich repeat kinase type 2 mutations discordant for Parkinson's disease. Mov Disord. Sep 1.2012 27(10):1323. [PubMed: 22488887]
46. Simon-Sanchez J, Scholz S, Fung HC, et al. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. Hum Mol Genet. Jan 1; 2007 16(1):1–14. [PubMed: 17116639]
47. Simon-Sanchez J, Schulte C, Bras JM, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet. Dec; 2009 41(12):1308–1312. [PubMed: 19915575]
48. Zhang X, Zhao J, Li C, et al. DSPP mutation in dentinogenesis imperfecta Shields type II. Nat Genet. Feb; 2001 27(2):151–152. [PubMed: 11175779]
49. Xiao S, Yu C, Chou X, et al. Dentinogenesis imperfecta 1 with or without progressive hearing loss is associated with distinct mutations in DSPP. Nat Genet. Feb; 2001 27(2):201–204. [PubMed: 11175790]
50. Lorenz-Depiereux B, Bastepe M, Benet-Pages A, et al. DMP1 mutations in autosomal recessive hypophosphatemia implicate a bone matrix protein in the regulation of phosphate homeostasis. Nat Genet. Nov; 2006 38(11):1248–1250. [PubMed: 17033625]
51. Mochizuki T, Wu G, Hayashi T, et al. PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. Science. May 31; 1996 272(5266):1339–1342. [PubMed: 8650545]
52. Hills LB, Masri A, Konno K, et al. Deletions in GRID2 lead to a recessive syndrome of cerebellar ataxia and tonic upgaze in humans. Neurology. Sep 27.2013

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

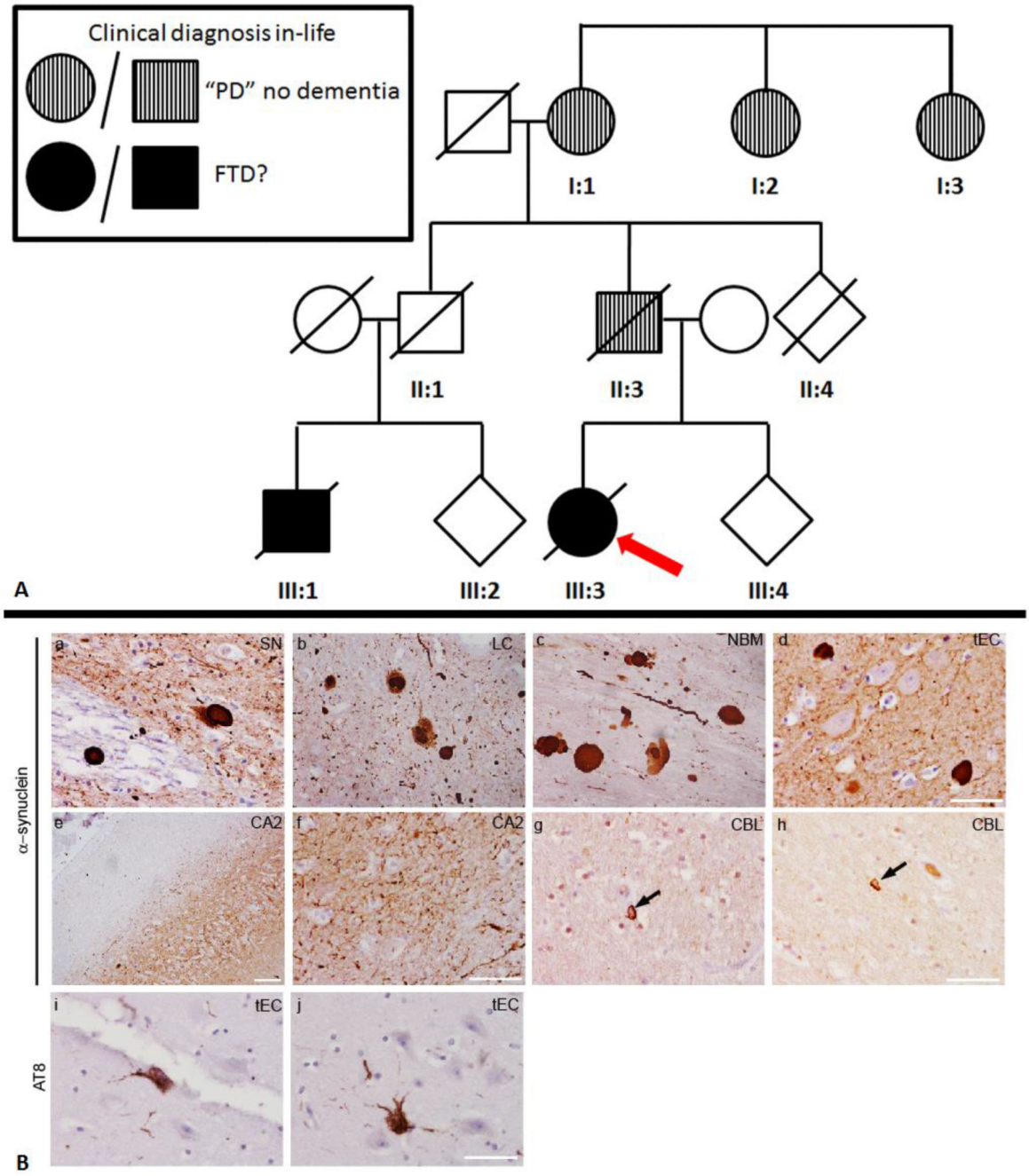

Figure 1. 'Family tree and neuropathology findings'

A) Family tree. Red arrow depicts the proband. Information has been omitted for deidentification purposes. **B)** Representative images of α-synuclein immunoreactive inclusions within (a) substantia nigra (SN), (b) locus coeulus (LC), (c) nucleus basalis of Meynert (NBM) and (d) transentorhinal cortex (tEC). Severe neuritic α-synuclein deposition is shown in the CA2 region of the hippocampus at low and high magnification (e, f). α–Synuclein coiled bodies shown in the white matter of the cerebellum (g,h, CBL). Sparse AT8 (phospho-tau, Ser 202 and Thr205) immunoreactive neurofibrillary tangles were

observed in the transentorhinal cortex (i,j, tEC). Scale bars represent 50 μm.
PD=Parkinson's disease; FTD=Frontotemporal dementia

Europe PMC Funders Author Manuscripts

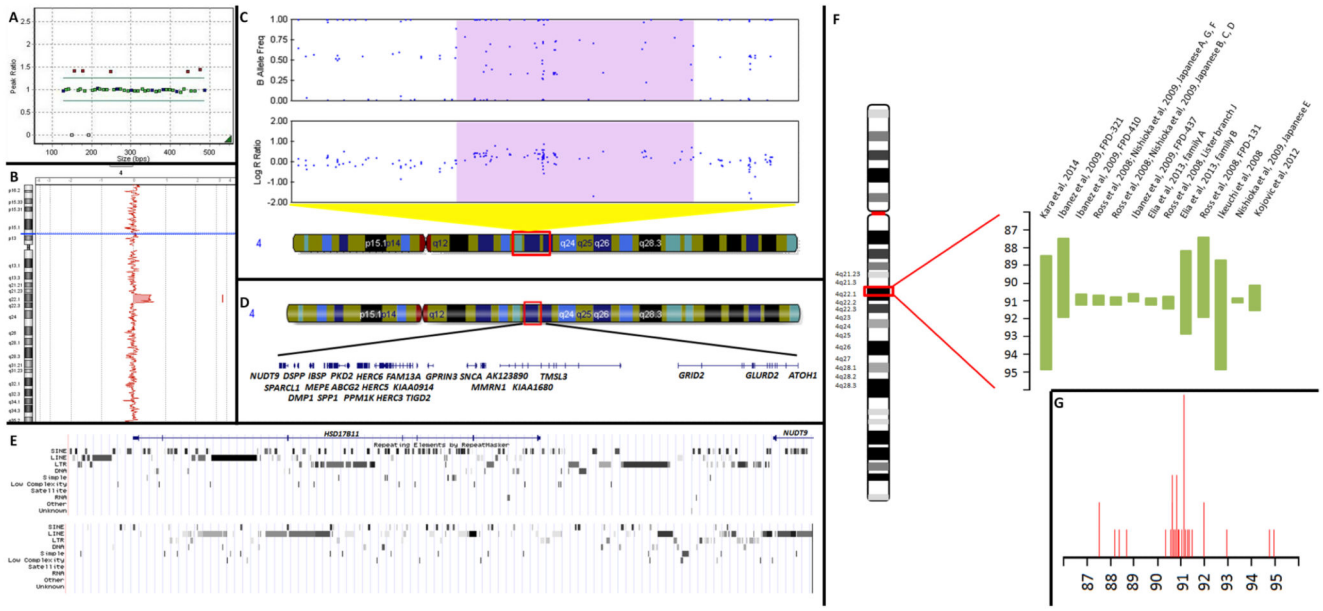

Figure 2. 'Genetic findings'

A) MLPA results depicting the duplication. **B)** Minimum duplicated region as identified by the CGH array (chr4:88,349,207-94,751,141, HB37). **C)** Minimum duplicated region as identified by the Immunochip (chr4: 88,231,429-94,816,144 HB37) with an increased logR ratio and probes forming 4 distinct genotype clusters [46]. Region depicted in figure is chr4: 83,666,791-97,713,411 (HB37). **D)** Genes included within the minimum duplicated region (chr4:88,349,207-94,751,141, HB37) as identified by the CGH array. **E)** Breakpoint regions and repetitive sequences identified through UCSC repeat masker. Upper panel centromeric (chr4:88,249,207-88,349,207), lower panel telomeric (chr4: 94,751,141-94,851,141). **F)** Comparison of the length of the duplicated/triplicated region between our case and published cases (band sizes are approximate). Sizes represent minimum length based on Human Build 36. **G)** Location of breakpoints. Height of peak corresponds to the number of cases with a breakpoint in the particular region. Breakpoint location could correspond to regions of recombination hotspots [21,47]. The case reported by Garraux et al (2012) [11] was excluded from both images.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

JAMA Neurol. Author manuscript; available in PMC 2015 March 17.

Table 1

List of genes included in the duplicated region in case III:1. Generated with NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/projects/mapview/).

List of genes included within duplication region in case III:1				
Start[1]	**Stop**[1]	**Gene Symbol**	**Description**	**Known role in disease**
88343728	88380606	*NUDT9*	nudix (nucleoside diphosphate linked moiety X)-type motif 9	NA
88394487	88450655	*SPARCL1*	SPARC-like 1 (hevin)	NA
88529681	88538025	*DSPP*	dentin sialophosphoprotein	dentinogenesis imperfecta-1[48,49]
88571454	88585513	*DMP1*	dentin matrix acidic phosphoprotein 1	Autosomal recessive hypophosphatemia[50]
88707075	88707514	*CHCHD2P7*	coiled-coil-helix-coiled-coil-helix domain containing 2 pseudogene 7	NA
88720702	88733601	*IBSP*	integrin-binding sialoprotein	NA
88742563	88767969	*MEPE*	matrix extracellular phosphoglycoprotein	NA
88812995	88815167	*HSP90AB3P*	heat shock protein 90kDa alpha (cytosolic), class B member 3, pseudogene	NA
88896802	88904563	*SPP1*	secreted phosphoprotein 1	NA
88928799	88998931	*PKD2*	polycystic kidney disease 2 (autosomal dominant)	Autosomal dominant polycystic kidney disease (ADPKD) [51]
89011416	89152474	*ABCG2*	ATP-binding cassette, sub-family G (WHITE), member 2	NA
89084691	89085162	*RPL31P24*	ribosomal protein L31 pseudogene 24	NA
89178761	89205983	*PPM1K*	protein phosphatase, Mg2+/Mn2+ dependent, 1K	NA
89299891	89364249	*HERC6*	HECT and RLD domain containing E3 ubiquitin protein ligase family member 6	NA
89378268	89427321	*HERC5*	HECT and RLD domain containing E3 ubiquitin protein ligase 5	NA
89428154	89431638	*LOC728333*	nuclear receptor coactivator 4 pseudogene	NA
89442129	89444952	*PIGY*	phosphatidylinositol glycan anchor biosynthesis, class Y	NA
89442129	89444952	*PYURF*	PIGY upstream reading frame	NA
89448713	89449318	*LOC100129137*	CD53 molecule pseudogene	NA
89513647	89629686	*HERC3*	HECT and RLD domain containing E3 ubiquitin protein ligase 3	NA
89605999	89606105	*RNU6-33P*	RNA, U6 small nuclear 33, pseudogene	NA
89617066	89619023	*NAP1L5*	nucleosome assembly protein 1-like 5	NA
89630940	89651254	*FAM13A-AS1*	FAM13A antisense RNA 1	NA
89647105	89978346	*FAM13A*	family with sequence similarity 13, member A	NA
90031864	90033424	*LOC731282*	uncharacterized LOC731282	NA

Europe PMC Funders Author Manuscripts

List of genes included within duplication region in case III:1				
Start[1]	**Stop**[1]	**Gene Symbol**	**Description**	**Known role in disease**
90033968	90036052	*TIGD2*	tigger transposable element derived 2	NA
90165429	90229161	*GPRIN3*	GPRIN family member 3	NA
90645250	**90759447**	***SNCA***	**synuclein, alpha**	PD [2]
90757552	90763142	*LOC644248*	uncharacterized LOC644248	NA
90816052	90875780	*MMRN1*	multimerin 1	NA
91048684	92523370	*CCSER1*	coiled-coil serine-rich protein 1	NA
91759636	91760266	*TMSB4XP8*	thymosin beta 4, X-linked pseudogene 8	NA
93103426	93105206	*LOC133083*	peptidase (mitochondrial processing) alpha pseudogene	NA
93225550	94693649	*GRID2*	glutamate receptor, ionotropic, delta 2	Cerebellar ataxia [52]
93623496	93623786	*MTND1P19*	MT-ND1 pseudogene 19	NA
93743117	93744164	*LOC100422562*	guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1 pseudogene	NA
94750078	94751142	*ATOH1*	atonal homolog 1 (Drosophila)	NA

[1]Base pairs, HB37

Europe PMC Funders Author Manuscripts

Table 2

Demographic and clinical characteristics of study sample by disease status

Demographic and clinical characteristics of individuals included in the statistical analysis.

	Disease status		Overall
Variable	**Unaffected**	**Affected**	
Individuals [n (%)]	10 (27%)	27 (73%)	37 (100%)
Gender			
Male [n (%)]	1 (7.14%)	13 (92.86%)	14 (37.84%)
Female [n (%)]	9 (39.13%)	14 (60.87%)	23 (62.16%)
Members per family [median (range)]	NA	NA	2 (1-7)
Age at Onset [median (range)]	NA	47 (31-71)	47 (31-71)
Duplication size [median (range)][2]	0.6 (0.4-3.47)	0.63 (0.2-5)	0.6 (0.2-5)
Composite score [median (range)]	0 (0-0)	2 (0-6)	2 (0-6)

[2]Mean (standard deviation-SD): 1.17 (1.22) (unaffected), 2.09 (2.12) (affected), 1.84 (1.95) (overall)

Europe PMC Funders Author Manuscripts

Table 3

Estimates for univariate (A) and multivariate models (B). OR represents the odds of affected disease status or late age at onset (> 40 years) for models 1 and 2 respectively (logistic regression models); OR >1 represents the odds of being affected or having a late AAO. OR represents the odds of severe disease outcome compared to mild and moderate for model 3 (ordinal logistic regression), where OR > 1 indicates worse outcome. OR=odds ratio, 95% CI=95% confidence intervals, AIC= Akaike Information Criterion.

A. Estimates for Univariate models									
	Model 1			**Model 2**			**Model 3**		
Outcome	**Disease status** (Logistic regression with robust standard errors)			**Age at Onset** (Logistic regression with robust standard errors)			**Disease severity** (Ordinal logistic regression with robust standard errors)		
Independent variable	**OR (95% CI)**	**p-value**	**AIC**	**OR (95% CI)**	**p-value**	**AIC**	**OR (95% CI)**	**p-value**	**AIC**
Duplication size	1.34 (0.78, 2.31)	0.285	45.31	0.94 (0.71, 1.24)	0.662	39.49	1.17 (0.81, 1.68)	0.395	83.78
Gender									
female	base			base			base		
male	8.36 (1.97, 35.42)	0.004**	41.99	0.23 (0.04, 1.28)	0.094	36.49	5.55 (1.39, 22.22)	0.016*	78.09
B. Estimates for Multivariate models									
Duplication size	1.35 (0.77, 2.37)	0.297	42.40	0.95 (0.69, 1.31)	0.766	38.43	1.12 (0.78, 1.62)	0.532	79.61
Gender									
female	base			base			base		
male	8.26 (1.81, 37.75)	0.006**		0.24 (0.04, 1.29)	0.095		5.27 (1.29, 21.50)	0.021*	