Linkage stratification and mutation analysis at the parkin locus identifies mutation positive Parkinson’s disease families

W C Nichols, N Pankratz, S K Uniacke, M W Paiciulo, C Halter, A Rudolph, P M Conneally, T Foroud, and the Parkinson Study Group

Parkinson’s disease (PD) is one of the most common neurological disorders in humans with an overall prevalence of 1:1000 with the incidence increasing to as high as 3.4% among people aged 75 years. The clinical phenotype includes resting tremor, muscular rigidity, bradykinesia, and postural instability. The signs and symptoms of the disease are the consequence of a striatal deficiency of dopamine resulting from neuronal death in the substantia nigra. Genetic linkage analyses have been performed by the presence of the Lewy body, an intracytoplasmic inclusions body found in many brain regions which is not entirely specific to, but is a highly sensitive marker for, Parkinson’s disease.

The pathogenesis of idiopathic Parkinson’s disease is unknown. For the overwhelming majority of PD patients, the disease has previously been thought to occur sporadically. However, there is increasing evidence of a genetic contribution to the disorder. Recently, two studies have investigated familial aggregation of PD using large, population based, case-control studies. Elbaz et al. reported an odds ratio of 3.2 for the presence of PD in first degree relatives (parents and sibs) of 175 cases as compared to 481 controls. Analyses stratified by age showed this aggregation to be stronger for younger PD patients. Familial aggregation of PD in Iceland was studied to the disorder. In an effort to identify additional genetic factors contributing to Parkinson’s disease, we have initiated the collection and analysis of a large panel of affected sib pairs. Families reporting at least two living sibs diagnosed or considered likely to be diagnosed with PD were ascertained through a variety of sources, including 52 US and Canadian centres participating in the Parkinson Study Group (PSG). All study participants completed a uniform clinical evaluation that consisted of parts II and III of the Unified Parkinson Disease Rating Scale (UPDRS).

In an effort to identify additional genetic factors contributing to Parkinson’s disease, we have initiated the collection and analysis of a large panel of affected sib pairs. Families reporting at least two living sibs diagnosed or considered likely to be diagnosed with PD were ascertained through a variety of sources, including 52 US and Canadian centres participating in the Parkinson Study Group (PSG). All study participants completed a uniform clinical evaluation that consisted of parts II and III of the Unified Parkinson Disease Rating Scale (UPDRS). In addition, a diagnostic checklist was developed based on the results of clinicopathological studies of Parkinson’s disease so as to reach an acceptable degree of diagnostic specificity and sensitivity. All study participants were then classified as either verified PD (VPD) or non-verified PD (NVPD), based on the results of the diagnostic checklist. The sample to date consists of 230 participants with VPD and 64 participants classified as NVPD giving a total of 162 sib pairs from 148 families and 94 affected sib pairs classified as verified PD from 86 families. Our sample consists of 95% whites and 5% Hispanics with an average age of onset of 60.4 years. Peripheral blood was obtained from all subjects after appropriate written informed consent approved by each individual institution’s IRB. DNA was prepared using standard methods.

As mutations in parkin are the most common inherited defect identified in PD to date, our initial efforts were focused on looking for evidence of linkage to the parkin locus. DNA samples from all 162 sib pairs were genotyped using 21 chromosome 6 dinucleotide repeats which are part of the ABI Prism Linkage Mapping Set (Applied Biosystems, Foster City, CA), including D6S305 located within intron 7 of the parkin gene. Briefly, 30 ng of genomic DNA was PCR amplified using each individual marker in a 10 µl reaction. After PCR, the PCR products were pooled using equal amounts of each PCR reaction. One µl of this multiplexed mix was added to 10 µl formamide containing the GeneScan-400HD ROX size standard (Applied Biosystems, Foster City, CA). Genotypes were determined using the ABI 3700 DNA Analyzer (Applied
of intronic PCR primers (sequences available upon request). For direct sequence analysis, all 12 exons of the parkin gene were PCR amplified for all 113 subjects. The resulting PCR products were purified using the QIAquick 96 PCR purification kit (QIAGEN, Santa Clara, CA) and sequenced on an ABI 377 DNA analyzer using the Applied Biosystems BigDye Terminator version 2.0 kit. We confirmed segregation of the mutations within families, and excluded the presence of the mutations in a panel of 182 normal chromosomes, by PCR amplification of the relevant exon followed by either mutation specific restriction fragment length polymorphism (RFLP) analysis or direct sequencing.

Fluorescent dosage PCR was performed as described by Yau et al. Briefly, 125 ng of genomic DNA was PCR amplified using a fluorescently labelled forward PCR primer. Multiplex PCR reactions were performed for exons 2-12 of the parkin gene in two different groups with group 1 containing exons 4, 6, 7, 8, 9, and 12 and group 2 containing exons 2, 3, 5, 10, and 11. PCR products were electrophoresed on an ABI 377 DNA Analyzer. Data were analysed using GeneScan and Genotyper software to produce electropherograms showing the size in base pairs of the peaks and areas under the peaks representing the amount of PCR product present. To determine gene dosage for each exon, samples were compared to each other, as well as to control samples, to obtain the dosage quotients. Each sample was repeated three times. Any individual exon sample giving an ambiguous or uninterpretable result was repeated an additional three times.

**RESULTS**

Results of the direct sequence and fluorescent dosage analysis are shown in table 1 and fig 1. Seventeen different parkin mutations were identified in 16 of 56 families analysed. Five of the mutations were detected more than once with the missense mutation Arg275Trp (three families) and the exon 8 duplication (3 families) accounting for ~40% of mutation positive families. It is not known whether these recurrent mutations are the result of a shared haplotype or represent independent mutational events. Three (16%) of the mutation positive families are of Hispanic origin; however, none of them carries the same mutation. The remaining 12 mutations were each detected in a single family. Three of the 17 mutations are missense mutations resulting in single amino acid substitutions in the parkin peptide. The remaining 14 are either deletions ranging in size from one base (54delA) up to entire exons or duplications of entire exons. All of the deletion and duplication mutations would predict a frameshift, except for the exon 3 and 4 deletion that predicts an in frame deletion of...

### Table 1 Parkin mutations identified in 16 Parkinson disease families

<table>
<thead>
<tr>
<th>Family</th>
<th>No.</th>
<th>Lod score</th>
<th>Age of onset</th>
<th>Mutation 1</th>
<th>Exon</th>
<th>AA change</th>
<th>Zygosity</th>
<th>Mutation 2</th>
<th>Exon</th>
<th>AA change</th>
</tr>
</thead>
<tbody>
<tr>
<td>33007</td>
<td>2</td>
<td>0.4031</td>
<td>28, 44</td>
<td>Duplication</td>
<td>7</td>
<td></td>
<td>Hetero</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05007</td>
<td>4</td>
<td>0.7164</td>
<td>45, 45, 45</td>
<td>Deletion</td>
<td>3, 4</td>
<td>Gln57fs ([+35 aa]†</td>
<td>Homo</td>
<td>Leu112fs ([+50 aa]</td>
<td>Hetero</td>
<td>4, 5, 6</td>
</tr>
<tr>
<td>15007</td>
<td>2</td>
<td>0.5376</td>
<td>38, 38</td>
<td>c.337-377del*</td>
<td>3</td>
<td></td>
<td>Hetero</td>
<td>Gln57fs ([+328 aa]</td>
<td>Homo</td>
<td></td>
</tr>
<tr>
<td>47007</td>
<td>2</td>
<td>0.1719</td>
<td>50, 51</td>
<td>Deletion</td>
<td>8</td>
<td></td>
<td>Hetero</td>
<td>Arg505fs ([+28 aa]</td>
<td>Homo</td>
<td></td>
</tr>
<tr>
<td>68007</td>
<td>4</td>
<td>1.4682</td>
<td>25, 31, 50</td>
<td>Deletion</td>
<td>3</td>
<td></td>
<td>Hetero</td>
<td>Deletion</td>
<td>2, 3, 4</td>
<td></td>
</tr>
<tr>
<td>79007</td>
<td>2</td>
<td>0.3599</td>
<td>24, 30</td>
<td>Deletion</td>
<td>3, 4</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>71107</td>
<td>2</td>
<td>0.4999</td>
<td>32, 54</td>
<td>c.154delA A</td>
<td>2</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>73107</td>
<td>2</td>
<td>0.3197</td>
<td>34, 36</td>
<td>C823T</td>
<td>7</td>
<td></td>
<td>Hetero</td>
<td>Lys211Asn</td>
<td>2, 3</td>
<td></td>
</tr>
<tr>
<td>93107</td>
<td>2</td>
<td>0.1424</td>
<td>48, 50</td>
<td>Duplication</td>
<td>8</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>74107</td>
<td>2</td>
<td>0.547</td>
<td>47, 62</td>
<td>Duplication</td>
<td>5</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>56107</td>
<td>2</td>
<td>0.5147</td>
<td>45, 72</td>
<td>Duplication</td>
<td>3</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>79107</td>
<td>3</td>
<td>0.9522</td>
<td>38, 38</td>
<td>A633T</td>
<td>6</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>93207</td>
<td>2</td>
<td>0.5538</td>
<td>35, 45</td>
<td>C823T</td>
<td>7</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>26207</td>
<td>2</td>
<td>0.5348</td>
<td>59, 65</td>
<td>Duplication</td>
<td>8</td>
<td></td>
<td>Hetero</td>
<td>Arg275Trp</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>60307</td>
<td>2</td>
<td>0.3792</td>
<td>22, 38</td>
<td>C823T</td>
<td>7</td>
<td></td>
<td>Hetero</td>
<td>C1310T</td>
<td>12</td>
<td>Pro437Leu</td>
</tr>
<tr>
<td>01307</td>
<td>2</td>
<td>0.3999</td>
<td>35, 44</td>
<td>Deletion</td>
<td>6</td>
<td></td>
<td>Hetero</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Nucleotides are numbered according to GenBank Accession AB009973 with the A of the initiator ATG numbered as +1.
† Frameshifts are denoted by the amino acid and its number at which the frameshift occurs as recommended by Dunnen and Antonarakis.

BioSystems, Foster City, CA) and GeneScan 3.5, Genotyper 3.6, and GeneMapper 1.1 software. All genotypic data were checked for Mendelian inheritance of marker alleles with the program Pedcheck. The marker genotypic data, including genotypic data from other parts of the genome, were used to verify the full sib relationships among the subjects using the computer program RELATIVE.

Observed allele frequencies in the subjects genotyped for the genome screen were used. Marker order and map positions were obtained from the Marshfield Medical Research Foundation (http://research.marshfieldclinic.org/genetics/). Multipoint affected sib pair linkage analysis was performed using the program Mapmaker/Sib and included dominance variance in the model. Lod scores were computed at 1 cM intervals along chromosome 6, using all possible sib pairs formed from families with more than two affected sibs. The parkin gene was previously reported to result in autosomal recessive PD; therefore, parametric linkage analysis was also performed using the computer program Allegro and an autosomal recessive, fully penetrant disease model with the disease allele having a frequency of 0.01. Analyses were performed using two models of affection: (1) including both verified and non-verified PD as affected, and (2) including only verified PD as affected.

Initial analyses of chromosome 6 using all available sib pairs resulted in a maximum lod score of 0.7 at marker D6S305. The highest peak in the region was 30 cM centromeric at D6S292 (lod=1.2). When analyses were restricted to only those sib pairs with VPD, the maximum lod score at the marker D6S305 (lod=1.9). Parametric linkage analyses performed using all available subjects also excluded linkage to this region using an autosomal recessive (lod=−6.85) mode of inheritance. When affection was limited to only those subjects with verified PD, the region was still excluded (lod=−38.4). These parametric linkage results were evaluated for each family, and 56 of the 148 families analysed had positive lod scores at D6S305, ranging from 0.0111 to 1.667. The ethnicity of these 56 families is 90% white and 10% Hispanic.

The positive lod scores at the parkin locus in this subset of 56 families prompted analysis of the parkin gene in the 113 affected subjects from these 56 families using both direct sequencing and fluorescent dosage analysis. By BLAST similarity searching of GenBank (http://www.ncbi.nlm.nih.gov/) and Celera (http://www.celera.com/) databases using a full length parkin cDNA sequence (GenBank accession number AB009973), the intron/exon boundaries of the gene were determined, which in turn enabled the design of fluorescence primer sets for exons 2-12 of the parkin gene, which were subsequently used to amplify intronic and exonic PCR products. The PCR products were electrophoresed on an ABI 377 DNA Analyzer. Data were analysed using GeneScan and Genotyper software to produce electropherograms showing the size in base pairs of the peaks and areas under the peaks representing the amount of PCR product present. To determine gene dosage for each exon, samples were compared to each other, as well as to control samples, to obtain the dosage quotients. Each sample was repeated three times. Any individual exon sample giving an ambiguous or uninterpretable result was repeated an additional three times.

The marker was previously reported to result in autosomal recessive PD; therefore, parametric linkage analysis was also performed using the computer program Allegro and an autosomal recessive, fully penetrant disease model with the disease allele having a frequency of 0.01. Analyses were performed using two models of affection: (1) including both verified and non-verified PD as affected, and (2) including only verified PD as affected.

Initial analyses of chromosome 6 using all available sib pairs resulted in a maximum lod score of 0.7 at marker D6S305. The highest peak in the region was 30 cM centromeric at D6S292 (lod=1.2). When analyses were restricted to only those sib pairs with VPD, the maximum lod score at the marker D6S305 (lod=1.9). When analyses were restricted to only those sib pairs with VPD, the maximum lod score at the marker D6S305 (lod=1.9). When analyses were restricted to only those sib pairs with VPD, the maximum lod score at the marker D6S305 (lod=1.9). When analyses were restricted to only those sib pairs with VPD, the maximum lod score at the marker D6S305 (lod=1.9). When analyses were restricted to only those sib pairs with VPD, the maximum lod score at the marker D6S305 (lod=1.9).
The five novel mutations reported here indicated that some of the point mutations reported previously by Abbas et al. Recently reported that haplotype analysis of 48 families carrying 14 distinct mutations indicated that some of the point mutations had been identified. Whether these patients are truly heterozygous or compound heterozygotes with the other allele carrying an undetected mutation remains to be determined. Our method of direct sequencing and fluorescent dosage would clearly miss any mutations that are not within the coding region, that is, those in the promoter or introns. Heterozygosity for parkin mutations in PD has been previously reported, and it has also been hypothesised that these subjects probably carry an undetected mutation on their other parkin allele.23,24

**DISCUSSION**

Of the 17 mutations identified in this study, 12 have been previously reported.23-25 The five novel mutations reported here consist of one missense mutation, three exon duplications, and one exon deletion. While we have no evidence in our study whether the recurrent mutations represent common founders or independent mutational events, Periquet et al. recently reported that haplotype analysis of 48 families carrying 14 distinct mutations indicated that some of the point mutations seem to arise from a common founder while the events involving whole exons occurred independently. The Arg275Trp mutation has now been identified in at least 12 unrelated families.25 An additional 116 affected subjects from 58 families who did not have positive parkin lod scores were screened by PCR and restriction endonuclease digestion for some of the mutations reported here as well as other previously reported mutations which were determined to change a restriction site. In this manner, two additional families were identified with mutations, one each with the 219-220insGT mutation reported by Abbas et al.25 No additional mutations were detected in any other affected subjects/families. Subsequent analyses of only those families without identified parkin mutations resulted in lod scores <0.1 for this entire region of chromosome 6, suggesting that all families with mutations had been identified.

The approach we are taking to identify additional genetic loci shows the synergy of linkage stratification and direct sequencing to identify mutation positive families. Our initial analysis of either the full sample of 162 affected sib pairs or the subset of 94 sib pairs with verified PD yielded no evidence of linkage to the parkin region. While stratification of the sample by families with suggestive evidence of allele sharing, or linkage to, the parkin gene enabled the identification of 17 parkin mutations in 16 of 56 (29%) families analysed. While all other reports have stratified samples based on age of onset of PD of 45 years or less for mutation detection,23 our strategy to analyse those families with evidence of positive lod scores at the parkin locus yielded the highest percentage of parkin mutations of any study to date.23-24 The ages of onset of those affected subjects in whom parkin mutations were identified in this study ranged from 22-69 years. However, four of the 16 mutation positive families had ages of onset greater than 45 years for all affected subjects and would have thus been missed if we had used age of onset as our stratification criteria. As well, identification and removal of subjects with mutations in known genes from subsequent genome wide analyses will decrease the heterogeneity in this sample and probably increase the power of this sample of affected sib pairs to detect linkage to additional chromosomal regions harbouring genetic loci contributing to the susceptibility to PD.

**ACKNOWLEDGEMENTS**

This project was supported by ROI NS37167. We thank the subjects for their participation in this research study. Support from NINDS NS37167 is gratefully acknowledged.


**Authors’ affiliations**

W C Nichols, S K Uniacke, M W Paucible, Division of Human Genetics, Children’s Hospital Medical Center, Cincinnati, OH, USA

N Pankratz, C Halter, P M Conneally, T Foroud, Department of Medical and Molecular Genetics, Indiana University Medical Center, Indianapolis, IN, USA

---

**Figure 1** Location of identified parkin mutations within the parkin gene. The 12 exons of the parkin gene are depicted as rectangles numbered 1-12. Locations of the initiator ATG and terminator TAG are as shown. Missense and small insertion/deletion frameshift mutations are depicted above the gene while mutations resulting in deletions/duplications of entire exons are indicated below the gene. All mutations were identified in a single family except those with a number in parentheses beside the mutation. The number in parentheses indicates the number of families in which those mutations were identified. An asterisk indicates mutations that were detected as heterozygous changes.
A Ruddolph, Department of Neurology, University of Rochester, Rochester, NY, USA

Correspondence to: Dr W C Nichols, Division of Human Genetics, Children’s Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229, USA; bill.nichols@chmcc.org

REFERENCES