

## Contig Map of the Parkinson's Disease Region on 4q21-q23

Christian LAVEDAN,<sup>1</sup> Anindya DEHEJIA,<sup>1</sup> Brian PIKE,<sup>2</sup> Amalia DUTRA,<sup>2</sup> Elisabeth LEROY,<sup>1</sup> Susan E. IDE,<sup>1</sup> Holly ROOT,<sup>2</sup> Jeffrey RUBENSTEIN,<sup>1</sup> Rebecca L. BOYER,<sup>1</sup> Settara CHANDRASEKHARAPPA,<sup>2</sup> Izabela MAKALOWSKA,<sup>3</sup> Robert L. NUSSBAUM,<sup>1</sup> and Mihael H. POLYMERPOULOS<sup>1,\*</sup>

*Laboratory of Genetic Disease Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892,<sup>1</sup> Laboratory of Gene Transfer, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892,<sup>2</sup> and Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892<sup>3</sup>*

(Received 7 January 1998)

### Abstract

We have constructed a yeast artificial chromosome contig (YAC) map of human chromosome 4q21-q23 across the Parkinson's disease region by combining molecular and fluorescence in situ hybridization techniques. This map contains 55 YACs and 51 molecular markers, including 23 polymorphic markers. We have also isolated one P1 and 33 bacterial artificial chromosomes located within this contig. Plasmid libraries were generated from 11 of these BAC and P1 clones, and 614 random plasmid clones were sequenced for a total of about 200 kb. This contig allowed us to precisely determine the location of 18 transcripts within the D4S2460-D4S2986 interval, including the alpha-synuclein gene found to be mutated in some families with Parkinson's disease.

**Key words:** Parkinson's disease; chromosome 4; contig; map; synuclein

### 1. Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects 1% to 2% of the population over age 50. The major clinical features of PD include resting tremor, bradykinesia and muscular rigidity.<sup>1,2</sup> Linkage analysis in a large Italian kindred has demonstrated for the first time the existence of a gene responsible for PD on human chromosome 4q21-q23.<sup>3</sup> The location of this gene was first established between polymorphic markers D4S2361 and D4S421,<sup>3</sup> and later narrowed down to a ~6 cM region between polymorphic markers D4S2371 and D4S2986 (unpublished data). In an effort to identify this PD gene by positional cloning, we constructed a high-resolution physical map of the region between markers D4S2371 and D4S2986 which we also extended on the centromeric side to include marker D4S2460. The map was constructed using the Centre d'Etude du Polymorphisme Humain (CEPH)/Génethon,<sup>4</sup> MIT/Whitehead,<sup>5</sup> and Stanford ([http://shgc.stanford.edu/Mapping/phys\\_map/index.html](http://shgc.stanford.edu/Mapping/phys_map/index.html)) databases. High throughput screening of yeast and bacterial artificial chromosome libraries (YACs and BACs) and one P1 library was performed, as well as fiber and metaphase fluorescence *in situ* hybridization (FISH)

analyses. These data were used to construct a map that includes the location and order of polymorphic markers, transcripts, and YAC and BAC clones. In addition, construction of 11 plasmid libraries and random sequencing of clones were performed to offer additional data on the gene content within this chromosomal region. This contig map provides useful mapping information in the 4q21-q23 chromosomal region.

### 2. Materials and Methods

#### 2.1. Oligonucleotides

Specific oligonucleotides were designed: For the ADH5 gene ADH5-5F (5'-CTTTTATAAGGCATTGCTGC-3') and ADH5-5R (5'-CAAGAGAATCACTGGGTTTC-3') amplify a 498-bp fragment that includes exon 5 (accession # M81116), and ADH5-8F (5'-TGTGATTTCTTTTAGGATGG-3') and ADH5-9R (5'-TAGCTGTGAAGCTCTACGAG-3') amplify a 415-bp fragment that includes exon 8 and part of exon 9 (accession # M81118). MPK-F (5'-TTGTTGCAGGTGACTAGCCG-3') and MPK-R (5'-TTAGGCTTGATTCTCTCCC-3') amplify a 149-bp fragment of the MAP kinase gene (accession # U07620). All oligonucleotides and Map Pairs were purchased from Research Genetics (Huntsville, AL, USA).

Communicated by Kenichi Matsubara

\* To whom correspondence should be addressed. Tel. +1-301-402-2119, Fax. +1-301-402-2170

## 2.2. Libraries

YAC clones from the CEPH library were grown in 25 ml culture in AHC medium (Bio 101, Vista, CA, USA) at 30 °C for 2–3 days. One milliliter of culture was transferred to a deep 96-well plate (Beckman Instruments) and aliquoted into MicroAmp Optical 96-well reaction plates (Perkin Elmer, Foster City, CA, USA) using a Hydra 96 (Robbins Scientific). Twenty-microliter PCR reactions were carried out on 2  $\mu$ l of liquid culture, in a 9600 thermocycler (Perkin Elmer) under standard conditions in the presence of 0.1 mM cresol and 12% sucrose. Amplification products were migrated in 3% Ultra Pure Agarose (Gibco BRL) and visualized by ethidium bromide staining. Fifty- and 100-bp ladders (Gibco BRL) were used as molecular weight standards. BAC and P1 libraries purchased from Research Genetics and Genome Systems (St. Louis, MO, USA) were screened as described earlier.<sup>6</sup>

BAC and YAC DNAs were prepared using Qiagen kit (Qiagen, Santa Clarita, CA, USA), or the Puregene kit (Gentra Systems, Research Triangle, NC, USA) respectively.

## 2.3. Sizing of clone inserts

BAC and P1 clone DNA was digested with *Not* I and *Sfi* I, as recommended by the manufacturer (Boehringer Mannheim, Indianapolis, IN, USA). Digestion products were subjected to a clamped homogeneous electric field (CHEF) gel electrophoresis, carried out in 1% agarose gel, 0.5  $\times$  TBE, at 6 volts/cm for 16 h at 14 °C with a 50- to 90-s ramp, using a CHEF-DRIII system (BioRad, Hercules, CA, USA). Size standards used were low and mid range pulsed-field gel (PFG) markers and lambda PFG marker (New England BioLabs, Beverly, MA, USA). DNA was visualized by SYBER Green I staining (FMC BioProducts, Rockland, ME, USA).

## 2.4. Fluorescence in situ hybridization

Fluorescence *in situ* hybridization (FISH) experiments were performed as previously described<sup>7</sup> with the following modifications. Colcemid was added after 72 h of mitogen stimulation, and bromodeoxy uridine (BrdU) was added to cultures at 72 h. On each slide, 1 mg of labeled BAC or YAC DNA was applied. Non-unique and nonspecific DNA hybridizations were blocked by pre-annealing the probes with a ten fold of human Cot-1 DNA. The signals were detected with either Cy5-Avidin (Biological Detection Systems Inc., Pittsburgh, PA, USA) or anti-digoxigenin conjugated to fluorescein isothiocyanate (FITC) (Oncor, Gaithersburg, MD, USA). Slides were counterstained with 4',6-Diamine-2'-phenylindole dihydrochloride (DAPI), 250 ng/ml (Boehringer Mannheim) with antifade solution. For fiber-FISH experiments, one million cells of Epstein-Barr virus (EBV)-transformed

lymphoblast lines were washed twice in phosphate-buffered saline (PBS, pH 7.4, Quality Biological, Inc., Gaithersburg, MD, USA). Four drops were placed on each slide. After air drying, the slides were immersed in lysis buffer (2 M NaCl, 25 mM Tris-HCl, 1% Triton X) for 20 min at room temperature. The buffer was then removed and the slides placed in a vertical position for 2 min to allow the chromatin fibers to extend with the flow of the liquid. The slides were then fixed in a 3 : 1 mixture of methanol : acetic acid (v/v) for 15 min and a FISH protocol was performed as described above.

## 2.5. Analysis of random sequences

Shotgun plasmid library construction from BAC inserts, arraying, and single-pass sequencing of random clones were carried out under commercial contract at SeqWright (Houston, TX, USA). The sequences were analyzed for potential microsatellite repeating elements. Sequences were compared to GenBank database sequences (nr and dblast) using BLAST.<sup>8</sup> Sequences have been submitted to the GSS division of GenBank (Table 3).

# 3. Results and Discussion

## 3.1. Contig map

As a first step toward constructing a contig map of the PD critical region, we searched the following databases to identify YAC clones and molecular markers known to map near the critical 4q21-q23 interval; MIT/Whitehead: [http://www.genome.wi.mit.edu/cgi-bin/contig/phys\\_map](http://www.genome.wi.mit.edu/cgi-bin/contig/phys_map), Stanford: [http://shgc.stanford.edu/Mapping/phys\\_map/index.html](http://shgc.stanford.edu/Mapping/phys_map/index.html), CEPH/G  n  thon: <http://www.cephb.fr/ceph-genethon-map.html>. A set of 95 YAC clones was selected and typed by polymerase chain reaction (PCR) for 136 markers. The markers were then ordered so as to minimize the number of clones with a non-contiguous set of markers. A contig map formed of 55 YAC clones (Fig. 1) was generated, that contained 18 transcripts and 33 anonymous markers, including 23 polymorphic markers.

Of these, a minimum set of 12 YACs formed a contig which covered the entire critical PD region from marker D4S2460 to marker D4S2986 (Fig. 1). YAC addresses were also identified for 34 additional 4q21-q23 markers, including 25 ESTs, which could not unambiguously be placed within this minimum contig (Table 1).

## 3.2. BAC and P1 clones

Our strategy was to develop sequence based polymorphisms near the existing polymorphic markers and known genes. Towards that goal, BAC libraries were screened for 28 sequence-tagged sites (STSs) of the YAC contig (Table 2). A P1 library was also screened for marker D4S1647, since no BAC or YAC library clone could be identified with this marker. A total of 33 BAC and one

**Table 1.** YAC addresses and characteristics of additional chromosome 4q21-q23 markers.

locus	accession number	characteristics	YAC addresses
D4S1534	Z23406	anonymous	804C3, 742G3
D4S2340	G02750	anonymous	804C3
D4S2446	G02764	anonymous	804C3, 742G3
D4S2462	*433424	anonymous	775D1
D4S2622	*685245	anonymous	775D1, 915E6
D4S2785	Z38358	MAP kinase	804C3, 742G3
D4S2828	G03886	anonymous	845B7, 922G11, 950H12
D4S2844	G04419	anonymous	775D1, 915E6
D4S3132	J04765	Osteopontin precursor	950H12
D4S3161	G06757	MAP kinase	804C3, 742G3
D4S3193	G11664	anonymous	775D1, 915E6, 890D3, 316G5, 840F2, 915E6
D4S3294	*734284	unidentified transcript	804C3, 742G3
D4S3381	Z43301	unidentified transcript	845B7, 922G11, 772D7
D4S3386	*4588566	unidentified transcript	804C3
D4S3388	Z67188	polymorphic	845B7, 922G11, 950H12, 772D7, 873F11
D4S3390	F02339	unidentified transcript	775D1, 915E6
D4S3392	R85643	unidentified transcript	775D1, 915E6
D4S3393	*4567323	unidentified transcript	775D1, 915E6
D4S3401	H09684	unidentified transcript	775D1, 915E6
D4S3402	H73415	unidentified transcript	775D1
D4S3405	*4564689	unidentified transcript	804C3, 742G3
D4S3408	*4566231	Highly similar to hypothetical 50.6-kD protein in RPL14B-GPA1 intergenic region [Saccharomyces cerevisiae]	845B7, 950H12, 873F11
D4S3410	*4573473	unidentified transcript	922G11, 950H12
D4S3414	R39502	unidentified transcript	775D1
D4S3417	H19314	unidentified transcript	804C3
D4S3419	F03124	unidentified transcript	775D1, 915E6
D4S3423	T78871	pM5 protein; highly similar to ferritin light chain [Equus caballus]	922G11, 950H12
D4S3429	T25852	unidentified transcript	845B7, 950H12, 753A3
D4S3430	R64619	unidentified transcript	775D1, 915E6
D4S3432	R92462	unidentified transcript	775D1
D4S3433	R92579	unidentified transcript	775D1, 915E6
D4S3439	G07227	protein tyrosine phosphatase (PTP-BAS, type 1)	775D1, 915E6, 61F6, 224C10, 925F8, 409A8, 853F12, 845B7, 915E6
D4S3440	T33104	human homologue to SC1	854G12, 922A10
ADH5	*118978	Alcohol dehydrogenase 5	

Accession # refer to Genbank ID # or GDB ID # (\*)

**Table 2.** P1 and BAC clones isolated with chromosome 4q21-q23 markers.

Marker		Clone		Not I	Sfi I
name	accession number	address	origin		
D4S414	Z16838	281O4	GS		
D4S423	Z17012	<b>23E8</b>	RG		
D4S1089	*195022	<b>335P21</b>	RG	2	2
D4S1557	Z23331	66A9	GS		
D4S1559	Z23754	<b>358P16</b>	RG		
D4S1560	Z23785	458M3	RG	1	
D4S1578	Z23953	<b>358P16</b>	RG		
D4S1647	G09195	<b>54F1</b>	RG		
D4S2361	*684066	125E7	GS		
D4S2371	G08356	<b>43M21</b>	GS		1
		102J7	GS		
D4S2380	G08405	<b>358M14</b>	RG		1
D4S2404	G08337	62A7	GS	1	
D4S2460	*424508	<b>355N11</b>	RG		1
D4S2461	*424511	<b>225H6</b>	GS		
D4S2667	*458591	151H19	GS		
		151H20	GS		
D4S2785	*588068	293H15	GS		
D4S2845	G03961	58O13	GS		
		82M18	GS		
D4S2909	*602954	113O20	GS		
D4S2929	Z52588	146F9	GS		
		16K18	GS		
D4S2986	*611523	<b>94E7</b>	RG		
D4S3006	*612582	171E7	GS	1	2
D4S3224	*678907	43J3	GS	1	1
		204M7	GS		
D4S3294	*734284	273G23	RG		
		367B9	RG		
D4S3368	*4564647	120M16	GS		
		21M20	GS		
D4S3369	*1221141	262C20	GS		
D4S3377	*4582796	219N21	GS		
D4S3414	R39502	53D3	GS		
		169O16	GS		
ADH5	M81118	360G16	RG		

54F1 is a P1 clone; all the other clones are BACs. Accession # refer to GenBank ID # or GDB ID # (\*). RS and GS refer to Research Genetics and Genome Systems, respectively. Plasmid libraries were constructed for clones shown in bold. Insert sizes are given in kilobase pairs; ND: not determined. The number of restriction sites present within the clone insert is indicated.

P1 library clones were isolated (Table 2), and their insert size was determined by PFG electrophoresis after two independent enzymatic digestions (*Not* I, and *Sfi* I). The presence of these restriction sites within the insert of clones 335P21, 458M3, 43M21, 358M14, 62A7, and 355N11 (Table 2), visualized as multiple digestion fragments, indicates the presence of possible CpG islands.<sup>9</sup>

### 3.3. Fluorescence in situ hybridization

Metaphase and fiber-FISH analysis of YAC, BAC and P1 clones was carried out from marker D4S2460 to marker D4S2986. Metaphase FISH analysis allowed us to orient the map with respect to the centromere and telomere (Fig. 1). YACs were also evaluated by FISH analysis for chimerism; only one of the 12 YAC clones from the minimum contig (Fig. 1) was apparently chimeric. Clone 966A6 hybridized to chromosome 4q22 as well as to 6p22 (data not shown). YAC clones 888B9, and 783H11 which were previously found by PCR to be chimeric with chromosome 5 (Whitehead Institute/MIT Center for Genome Research) appeared to hybridize only with chromosome 4q22 (data not shown).

Furthermore, fiber-FISH experiments were carried out to confirm the order of clones previously determined by the STS-EST contig-map. BAC clones 293H15 and 360G16 localized the MAP kinase and the ADH5 genes outside the D4S2460-D4S2986 interval, on the centromeric and telomeric side, respectively (data not shown).

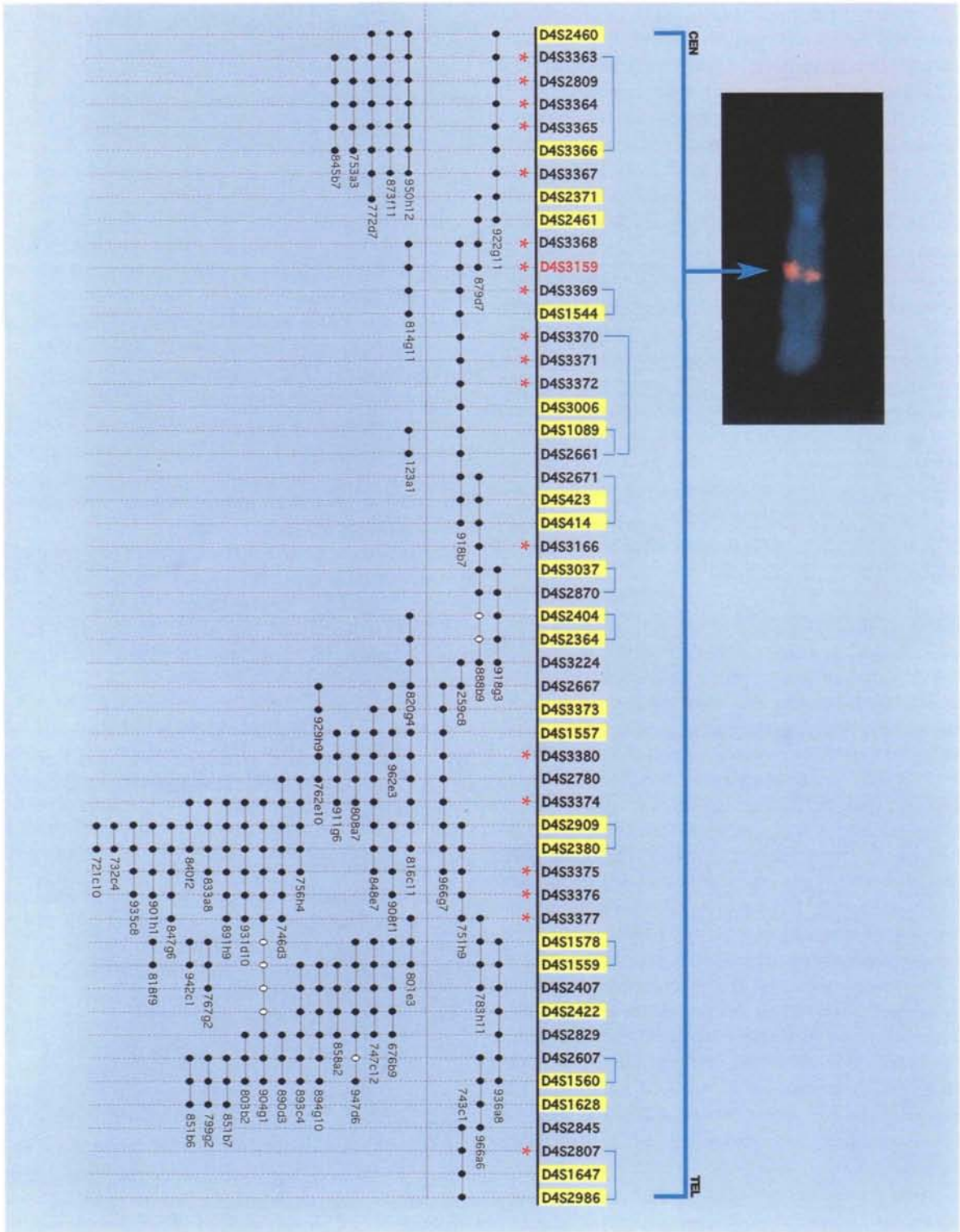
Finally, by combining estimated distances between BAC clones obtained by fiber-FISH, and sizes of YAC and BAC clones, we evaluated the D4S2460-D4S2986 interval to span about 6 Mb.

### 3.4. Random sequences

Plasmid libraries were generated for 10 BACs and for 1 P1: from 7 of these 11 libraries, a total of 614 random clones covering a total of about 200 kb were sequenced. Sequences were compared to GenBank databases (nr and dbest) using BLAST.<sup>8,10</sup> Sequences were submitted to GenBank (Table 3) and search results are available at <http://www.nhgri.nih.gov/DIR/projects.html>.

**Table 3.** Clone sequences from the PD region on chromosome 4q21-q23.

Clone name	Number of sequences analyzed	GenBank accession numbers
102J7	102	AF010608 -> AF010709
225H6	102	AF010710 -> AF010811
335P21	18	AF010812 -> AF010829
355N11	102	AF010830 -> AF010931
358M14	94	AF010932 -> AF011025
54F1	101	AF011026 -> AF011126
94E7	95	AF011127 -> AF011221



**Figure 1.** Map of the human chromosome 4q21-q23 between markers D4S2460 and D4S2986. **Top:** Metaphase FISH analysis showing the location of YAC 922g11 on the human chromosome 4. **Bottom:** YAC contig. Only unambiguous results on YAC clones with at least 2 contiguous markers are shown. Markers are designated by their D4S number. YACs are represented by horizontal lines with black or white circles for markers present or absent, respectively. The YAC addresses are also indicated. Polymorphic markers are highlighted in yellow. The location of the alpha-synuclein gene is indicated by marker D4S3159 (in red). Brackets show blocks of markers which were not ordered. \* indicates transcripts. The minimum contig of 12 YAC clones is shown above the thin horizontal line.

### 3.5. Conclusions

We have constructed a high-resolution 6-Mb physical map of the human chromosome 4q21-q23 region between markers D4S2460 and D4S2986. This contig map contains 55 YACs and 51 molecular markers, resulting in an average density of about one marker per 120 kb. Out of 18 transcripts that we mapped within this PD critical region, only one (D4S3159) corresponded to a known gene: alpha-synuclein or non-A4 component of amyloid precursor.<sup>11</sup> This gene was found to carry a mutation in four families with PD.<sup>12</sup>

YAC addresses were determined for 34 additional STSs, and we have re-defined the location of 42 transcripts including the alpha-synuclein gene. Thirty-three BAC and one P1 clone were isolated using 28 different markers. Plasmid libraries were generated from 11 of these clones, and about 200 kb of raw sequence was obtained. This contig map provides physical mapping tools as well as DNA sequence information for gene discovery in the 4q21-q23 region.

### References

1. Parkinson, J. 1817, *An Essay on the Shaking Palsy*, London.
2. Gowers, W. R. A Manual of diseases of the nervous system, In: edited by Blackinson, Philadelphia: Blackinson, 1893, p. 6366-6657.
3. Polymeropoulos, M. H., Higgins, J. J., Golbe, L. I. et al. 1996, Mapping of a gene for Parkinson's disease to chromosome 4q21-q23, *Science*, **274**, 1197-1199.
4. Chumakov, I. M., Rigault, P., Le Gall, I. et al. 1995, A YAC contig map of the human genome, *Nature*, **377**, 175-297.
5. Hudson, T. J., Stein, L. D., Gerety, S. S. et al. 1995, An STS-based map of the human genome, *Science*, **270**, 1945-1954.
6. Couch, F. J., Castilla, L. H., Xu, J. et al. 1995, A YAC-, P1-, and cosmid-based physical map of the BRCA1 region on chromosome 17q21, *Genomics*, **25**, 264-273.
7. Dutra, A. S., Mignot, E., and Puck, J. M. 1996, Gene localization and syntenic mapping by FISH in the dog, *Cytogenet Cell Genet*, **74**, 113-117.
8. Altschul, S. F., Madden, T. L., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990, Basic local alignment search tool, *J Mol Biol*, **215**, 403-410.
9. Lindsay, S. and Bird, A. P. 1987, Use of restriction enzymes to detect potential gene sequences in mammalian DNA, *Nature*, **327**, 336-338.
10. Zhang, J. and Madden, T. L. 1997, PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation, *Genome Res*, **7**, 649-656.
11. Ueda, K., Fukushima, H., Masliah, E. et al. 1993, Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease, *Proc Natl Acad Sci USA*, **90**, 11282-11286.
12. Polymeropoulos, M. H., Lavedan, C., Leroy, E. et al. 1997, Mutation in the alpha-synuclein gene identified in families with Parkinson's disease, *Science*, **276**, 2045-2047.

