

RAPID COMMUNICATION

Christian Lavedan · Elisabeth Leroy · Anindya Dehejia
Stephanie Buchholtz · Amalia Dutra
Robert L. Nussbaum · Mihael H. Polymeropoulos

Identification, localization and characterization of the human γ -synuclein gene

Received: 6 February 1998 / Accepted: 8 April 1998

Abstract We have identified and characterized a new member of the human synuclein gene family, γ -synuclein (SNCG). This gene is composed of five exons, which encode a 127 amino acid protein that is highly homologous to α -synuclein, which is mutated in some Parkinson's disease families, and to β -synuclein. The γ -synuclein gene is localized to chromosome 10q23 and is principally expressed in the brain, particularly in the substantia nigra. We have determined its genomic sequence, and established conditions for sequence analysis of each of the exons. The γ -synuclein gene, also known as BCSG1, was recently found to be overexpressed in advanced infiltrating carcinoma of the breast. Our survey of the EST database indicated that it might also be overexpressed in an ovarian tumor.

Introduction

The synuclein gene family comprises genes that are highly conserved between species. The first synuclein gene was isolated from the electric organ of *Torpedo californica* (Maroteaux et al. 1988), and soon after rat orthologs were described (Maroteaux and Scheller 1991). In humans two members of this gene family, the α - and β -synuclein genes (SNCA and SNCB) have been characterized (Jakes et al.

1994). Although the function of the synuclein proteins remains largely unknown, it has been shown that the ortholog to α -synuclein in the zebra finch, synelfin, might play a role in song learning (George et al. 1995).

We previously identified a mutation (Ala53Thr) in the α -synuclein gene in families with the autosomal dominant form of Parkinson's disease (Polymeropoulos et al. 1997). We hypothesized that this Ala53Thr mutation might favor the aggregation of the α -synuclein protein, ultimately resulting in cell death, and development of the disorder. Recently, another mutation (Ala30Pro) has been found in a German Parkinson's disease family (Krüger et al. 1998). Our study of 52 Parkinson's disease patients with the sporadic form of the disease indicated that the Ala53Thr mutation is not likely to account for the majority of cases with Parkinson's disease (Polymeropoulos et al. 1997). Evidence for locus heterogeneity in Parkinson's disease (Gasser et al. 1997; Scott et al. 1997) suggests the involvement of different genetic factors in other familial Parkinson's disease cases. It is conceivable that such genetic factors could include genes and gene products interacting with and/or similar to α -synuclein. In a search for other genes that could be mutated in patients with Parkinson's disease we have identified a third member of the synuclein gene family, which we termed γ -synuclein, that is highly expressed in the substantia nigra. In this paper we present the genomic structure of the gene and its localization to chromosome 10q23. We also found that the γ -synuclein gene sequence is almost identical to the sequence of a recently discovered gene (BCSG1) that is overexpressed in advanced infiltrating carcinoma of the breast (Ji et al. 1997). Survey of the EST (expressed sequence tag) database revealed that the γ -synuclein gene might also be overexpressed in ovarian tumors.

C. Lavedan · E. Leroy · A. Dehejia · S. Buchholtz · R. L. Nussbaum
M. H. Polymeropoulos (✉)
Laboratory of Genetic Disease Research, National Human Genome
Research Institute, National Institutes of Health, 49 Convent Drive,
49/4A66, Bethesda, MD 20892, USA
Tel.: +1-301-402-2119, Fax: +1-301-402-2170

C. Lavedan
Johns Hopkins Bayview Medical Center,
Baltimore, MD 21224, USA

A. Dutra
Laboratory of Gene Transfer, National Human Genome Research
Institute, National Institutes of Health, Bethesda, MD 20892, USA

Materials and methods

Database searches

Sequence homology and EST clone searches were conducted using the BLAST package of the sequence analysis program (Altschul et al.

1990). Searches were performed against the dbEST and GenBank databases at the following sites:

dbEST http://www2.ncbi.nlm.nih.gov/dbST/dbest_query.html
Genbank <http://www2.ncbi.nlm.nih.gov/cgi-bin/genbank>

Sequencing and sequence analysis

Polymerase chain reaction (PCR) products were purified through Microcon 100 (Amicon, Beverly, Mass.) and 5 ng per 100 bp was used for sequencing with the ABI Prism Dye Terminator Cycle Sequencing kit as recommended (Perkin Elmer, Foster City, Calif., USA). Sequencing reactions were purified with Centrisep spin columns (Princeton Separation, Adelphia, N.J., USA) and run on the ABI 373A sequencer. DNA and amino acid sequences were analyzed using the EditSeq, SeqMan and Megalign Programs software package (DNASTAR, Madison, Wis., USA).

Oligonucleotides and PCRs

All oligonucleotides (Table 1) were purchased from Research Genetics (Huntsville, Ala., USA). Polymerase chain reactions (20 or 100 µl) were carried out in a Perkin Elmer 9600 thermocycler using 15–50 ng of DNA as template. Three different conditions were used: (1) 94°C, 5 min; 94°C, 30 s, 55°C, 30 s, 72°C, 30 s for 30 cycles; 72°C, 10 min, and then 4°C, in 50 mM KCl, 10 mM TRIS-HCl, pH 8.3, 1.5 mM MgCl₂, 0.01% gelatin, 200 µM dNTP (Pharmacia, Piscataway, N.J., USA), with 60 ng of each primer, and 0.3 U of AmpliTaq polymerase (Perkin Elmer); (2) same as (1) except that the annealing temperature was changed to 60°C; condition (1) was used with primer pairs γ18-γ19, γ24-26 and γ10-γ38, while condition (2) was used with primer pairs γ3-γ6, and γ25-γ37; long PCR was performed with the Expand long-template PCR system 1 (Boehringer Mannheim, Indianapolis, Ind., USA) using the following cycle: 94°C, 2 min, six cycles of 92°C, 10 s; 67°C, 30 s temperature decreased by 0.5°C at each cycle, 68°C, 6 min; six cycles of 92°C, 10 s, 64°C, 30 s, 68°C, 6 min; 23 cycles of 92°C, 10 s, 63°C, 30 s, 68°C, 6 min extension increased by 20 s at each cycle; 68°C, 7 min.

Screening of bacterial artificial chromosome (BAC) library

A BAC library purchased from Genome Systems (St. Louis, Mo., USA) was screened as described earlier (Couch et al. 1995). BAC

Table 1 Oligonucleotides used to determine the structure of the human γ-synuclein gene

Name	5'→3' sequence	Name	5'→3' sequence
γ1	atgtcttcaagaaggcttc	γ26	tagacaagccctgtgtgca
γ2	ccttggtcttctcagctct	γ31	agcgtggatgacctgaagag
γ3	aggagatccagctccgtct	γ32	agcacaggtggacaggccaag
γ6	accctttggccaccactgt	γ33	gtggtgcgaaggaggacttg
γ10	actcagggtggcccattaag	γ34	ctcttcaggatccacgct
γ18	ttatggtctgcagggtctgtag	γ35	aggtgaccgcgatgttctccg
γ19	accgaccccacagttgtcca	γ37	tgactccagcaggcctgctt
γ22	ccagccagtgtctcccata	γ38	aaggcagaggagcgtcttca
γ24	aggagggaagcagtgctgt	γ39	aagagcgtctctgctctg
γ25	tggggtgtggagtggagtgat	γ40	cagcagcataagtggggctc

DNA was prepared using the Qiagen DNA extraction kit (Qiagen, Santa Clarita, Calif.).

Radiation hybrid panel and fluorescence in situ hybridization (FISH)

Mapping was performed by PCR analysis of the Genebridge 4 radiation hybrid panel (Walter et al. 1994). Statistical analysis of the data was performed using the RHMAPPER software package (D. Slonim, L. Stein, L. Kruglyak, E. Lander) at:

<http://www.genome.wi.mit.edu/ftp/distribution/software/rhmapper/>

The FISH experiments were performed as previously described (Dutra et al. 1996).

Northern blot analysis

Northern blots (Clontech Laboratories, Palo Alto, Calif.) were used as recommended by the manufacturer. A γ39-γ40 PCR product (Table 1) was amplified from BAC clone 174P13. A γ39-γ40 probe was generated by asymmetric PCR of 10 ng of the γ39-γ40 PCR product using primer γ40 (1 µM), in the presence of [³²P]dCTP (50 µCi). The probe was purified on a Sephadex G-50 column (Pharmacia) and used at 10⁶

Table 2 Human γ-synuclein clones. Accession numbers should be used when searching GenBank or dbEST. (NA not available in GenBank)

Clone designation in dbEST	Accession numbers associated with clone		Source (library name)
	5'	3'	
274172	H49348	NA	Fetal liver, spleen (Soares 1NFLS)
377642	AA056035	AA055968	Fetal heart (Soares NbHH19W)
400003	NA	AA722407	Fetal heart (Soares NbHH19W)
724046	AA410737	AA235594	Ovarian tumor (Soares NbHOT)
724129	NA	AA411094	Ovarian tumor (Soares NbHOT)
726391	AA394097	AA293803	Ovarian tumor (Soares NbHOT)
739149	NA	AA421708	Ovarian tumor (Soares NbHOT)
741179	AA402564 ^a	AA402624	Ovarian tumor (Soares NbHOT)
771303	NA	AA443638	Ovarian tumor (Soares NbHOT)
811164	AA486472	NA	Ovarian tumor (Soares NbHOT)
858268	NA	AA633976	Differentiated, postmitotic hNT neurons ^b
HRBAA27	M79265	NA	Whole adult brain
EST68G11	W22518	NA	Adult retina
EST75830	AA365132	NA	Pineal gland II

^aIndicates a sequence not homologous to synuclein genes

^bHuman neurons derived from a teratocarcinoma

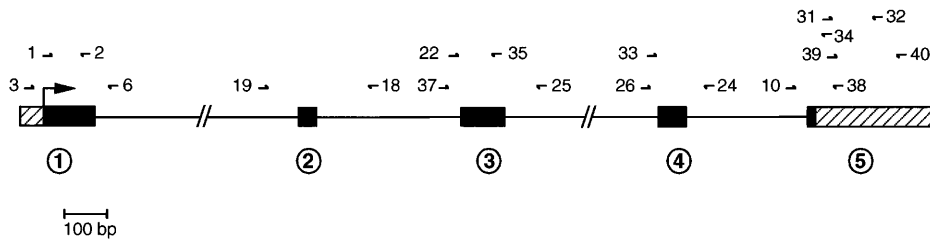


Fig. 1 Human γ -synuclein gene structure. Exons are numbered and are represented by *boxes* (black for the coding region). Introns are represented by a *horizontal line, interrupted* when the entire intron is not shown. The γ oligonucleotides used (Table 1) are shown above the gene. The direction of transcription is indicated by an *arrow* at the site of the AUG start codon

dpm/ml. A β -actin cDNA fragment, labeled with the Ready To Go kit (Pharmacia), was used as a control as recommended by the manufacturer (Clontech). Hybridizations were carried out at 42°C for 16 h in 5 \times SSPE, 10 \times Denhardt's, 100 μ g/ml denatured sheared salmon sperm DNA, 50% formamide and 2% SDS. Membranes were washed for 40 min in 2 \times SSC, 0.05% SDS at room temperature, and for 30 min in 0.1 \times SSC, 0.1% SDS at 50°C, and a Kodak X-OMAR AR film was exposed at -80°C. Between each hybridization, membranes were stripped as recommended by the manufacturer and the film was exposed overnight.

Results

Database search

We used DNA sequences of the α - and β -synuclein genes, (GenBank accession numbers L08850 and S69965, respectively) to search the GenBank and dbEST databases. Three homologous human sequences, M79265, AA394097 and AA293803, were identified that originated from two independent human cDNA clones, HRBAA27 and 726391 (Table 2). These clones were homologous to both α - and β -synucleins, more precisely to the position of the gene corresponding to exons 3 and 5, respectively, of the α -synuclein sequence. We used the M79265, AA394097 and AA293803 sequences to search the same databases, and found 14 additional human DNA sequences that were all homologous to clone 726391 (Table 2); one clone, 858268 (sequence accession number AA633976), was homologous to the M79265, AA394097 and AA293803 sequences, indicating that clones HRBAA27 and 726391 were likely to be derived from a single new gene that was different than the α - and β -synucleins. We also identified one rat mRNA and 16 mouse cDNA sequences (Lavedan et al. in preparation) matching M79265, AA394097 and AA293803. The 5' sequence of clone 741179 (AA402564) did not match the synuclein sequence, implying that this clone was probably chimeric (Table 2).

Localization

For mapping we used information derived from sequences M79265, AA394097 and AA293803. Specific human genomic markers were created by PCR. Amplicons γ 1- γ 2 from

sequence M79265 and γ 31- γ 32 from sequences AA394097 and AA293803 (Fig. 1), were used on the Genebridge 4 radiation hybrid panel (Walter et al. 1994): both mapped to chromosome 10q23 between marker WI-5226 and marker AFM225YD12 (data not shown). In order to obtain a clone containing the entire gene, we also screened a BAC library with the same primer pairs. A clone (174P13) was isolated, and shown to contain both amplicons (data not shown). Furthermore, this BAC clone was used as a probe in FISH experiments. Two hybridization signals were seen on each metaphase examined, both on chromosome 10 (data not shown): a strong signal on 10q23 and a fainter signal in the centromeric region. We believe this is due to the presence of repeated sequences or to chimerism of the BAC clone, which would contain two independent regions of chromosome 10. Since α - and β -synucleins have been previously mapped to 4q21-q23 (Chen et al. 1995; Shibasaki et al. 1995; Spillantini et al. 1995) and 5q35 (Spillantini et al. 1995), respectively, these results favored the hypothesis of the existence of a new gene of the synuclein family on chromosome 10q23.

Structure of the γ -synuclein gene

Additional oligonucleotides were designed (Table 1) and long PCR was performed using DNA from BAC 174P13 as a template (Fig. 1). The size and the sequence of the amplified fragments γ 1- γ 35 (~1.4 kb), γ 22- γ 24 (~2.8 kb), γ 33- γ 34 (~0.42 kb) confirmed that sequences M79265, AA394097 and AA293803 were part of a single gene. The exon-intron organization (Fig. 2) was deduced from the alignment and comparison of the genomic sequence with the homologous EST sequences (Table 2), including sequences from the α - and β -synuclein genes, and the presence of consensus intron-exon splice-junction sequences. The AUG start codon was identified based on the sequence homology with the α - and β -synuclein proteins and the presence of an excellent Kozak (1991) sequence: CCCACCATGG. This new member of the synuclein gene family, the γ -synuclein gene, spans ~5.0 kb and is composed of five exons (Figs. 1, 2). We have developed intronic oligonucleotides and PCR conditions specifically to amplify and sequence each exon. By comparing the sequence obtained from BAC clone 174P13, four CEPH DNA controls, and the reported ESTs, we found four single nucleotide polymorphisms (SNPs) within the the γ -synuclein gene sequence (Fig. 2): one SNP (C \rightarrow A) was observed at -19 from the AUG start codon. Another one (C \rightarrow T) was detected in the 3' untranslated region, 122 bp after the UAG stop codon. One SNP changes the third nucleotide (G or C) of codon 65. More interestingly, a

Fig. 2 Genomic sequence of the human γ -synuclein gene. The entire sequence is available in GenBank: accession number AF044311. The protein sequence is shown below the coding nucleotide sequence (*bold*). Exons are delimited by *brackets*. The Kozak sequence is *boxed*, and the stop codon is *underlined*. Asterisks indicate single-nucleotide polymorphisms. The sequence of intron 3 is interrupted (/)

```

AGGGAGATCCAGCTCCGTCCTGCCTGCAGCAGCCCAACCTGCACACCCACCATGGATGTCTTCAAGAAGGGCTTCCATCGCCAAGGAGGGCGTGGTG
Met Asp Val Phe Lys Lys Gly Phe Ser Ile Ala Lys Glu Gly Val Val
GGTGGCGGTGGAAAAGACCAAGCAGGGGGTGACGGAAGCAGCTGAGAAGACCAAGGAGGGGGTCAATGTATGGCTAAGTGGGGCATGGCAGGGTGGGACA
Gly Ala Val Glu Lys Thr Lys Gln Gly Val Thr Glu Ala Ala Glu Lys Thr Lys Glu Gly Val Met Tyr Val
GTGTGGTGGCCAAAGGGTGAGTGCCAGTTACCTTCGCCAGACCTTACTCCCCAGCCCCAGGGAGGCATTTGGGAGGGGGCGAGGCCCTGGCTATCAAG
GTGGGGTCTCCACACCTCGAGCACCACAATGCCCTGTGCACACCTATGTGTGTTGTCTTTGGCTCCTCGGGGCTCTGGGTGTGAGAGACAGCAGA
CAGGGCTGGCTACCTGTCTGTGCACGCACACACATTCCTCAAGCATACACCCCTCCCTGAGCGTGGAGCCCTGAAGCCATGAGCAGCCTGTGCTCAG
GTGGCCCCACCTCTCCACACGGGAGCGGTACAGCCAGGTACCGATCCCCCTCCCTCCAGAGAGAAGGGGCGAGGCTGGGGATGAAACCTAGGCTC
AGTGTTCCTCCCGGCATCCTCTCCTGGCACTCTCCAGAGGAGGAAGGGGAGGTCAAGCCAACTGACTCAGCTCTGGCCATCTGTCTGTTGCTGCTT
CTGAGGGCCGGCCACCCGGGAGGGGTGGACCTGGGTTAGCCAGTGTCCCTACCTCAGGCCCTGCTCTCTCTGTCCCCACATTCTGTCTGTCTC
CCTTCCCATCCATCCACTTCTTCCAGACACAGCAGGAAGAGGCCCTCTGAAGGGGCCGGCCGCCCCAGACACCATCTTACCCCGCCACCGCCACCA
GTTTGTCCAGCTGTTCTGTTGTGTTGTCTGACCGCCCCAACCTCGAGGGAGGTCTGGGCTGACAGCTCCATTTCTCCCGAGCCAAAGACCA
Gly Ala Lys Thr Lys
GGAGAATGTTGTACAGCGCTGACCTCACTGTGAGAAGCCCCAGGGCCAGGGGACACATGGGGATAGGACCCCTGGGGCTCTGCATCTAGTGTCTGGG
Glu Asn Val Val Gln Ser Val Thr Ser
CTCAAACCTAGAGTCTGCCTTACCCCAACTGGGTCCCAAGCCCTACAGACCCCTGCAGACCATAAGGATAAAGTAGGGTGGGCGTGTCTTTACCCCG
ACCAGCATCAGAGGTGCCCTGGAGTGCAGGGGAGCAGGGGAGGGTCCAGCAGGGCCAGGGCTCTGAGCTCTGGGAAGGGGGTGGCGCTGACTCCAG
CAGGCCCTGCCTGGGGCTGGGGCTGGGGTGGAGGCCAGCCAGTGTCTCCCATAGGGCCGAGAAGACCAAGGAGCAGGCCAACCGGGTGAGCGAGGGCTG
Val Ala Glu Lys Thr Lys Glu Gln Ala Asn Ala Val Ser Glu Ala
TGGTGAGCAGCGTCAACACTGTGGCCACCAAGACCGTGGAGGAGCGGAGAACATCGCGGTACCTCCGGGGTGGTGGCCAAGGTGAGCCCCGGCCCTCA
Val Val Ser Ser Val Asn Thr Val Ala Thr Lys Thr Val Glu Glu Ala Glu Asn Ile Ala Val Thr Ser Gly Val Val Arg Lys
AACCTGCCAGTCTCTCTCCGGGCCAGAAAGGCTGTGGGGGAGGGCCATCACTCCACTCCACACCCAGGAAACAACAGGGGGGCTGGCGTGGC
TTCAGTTTCAGTACCCTGAGCGCCGCATGGGGTGGGGTCCATGCTTGTGTTGCACCTGCATTGGGGTCTGGAATCACCACAACCCCTTCGGGG
GACTTGGGAGCCATCTCATAGACAAGGAAAGTGGGCTCAAACCCACATCTCCGGAGGGCTCTCAGCACTAAGAACACCCCACTTTTATAAGTGAAA
ACAGGATCCACACGGTGACAGGCACCTCCCGGCGAGTGGTACCAGGCCGGGGGCTGAAT/TCTTGTAGGGCTACTGAGGTGGGGTCAGACCTGGGG
CTTCAGGGCTTTTGCCATTTTAAACCCATTCCACCTTACCCACCCAGGTTCTTCATCAGTGAAGCCAACTAGGGGCTCCAGAGACTGGGGACC
AGCTTGAAGGGTGAAGCTGGGCTCCAGGGTACTTGGTGCCCTGGAAGGCCCTGGGCATCATCAGAGCCCCGGTATTGTCTGAATTGCACGCCCAATA
CCTCCAGCCCTCCAAAGTACAACAAGGCCACGAGGGCCCTGTGCTCTTCTTGAGGCCAGGGTAGACAAGGCCCTGGTTCAGAGGCAGCTAGGGGTCTCT
GCATTCCTCCACATGGTCTCATGCCCTTTTGTCCCTACAGAGGACTTGAGGCCATCTGCCCCCAACAGGAGGGT*GGCATCAAAGAGAAAGAGG
Glu Asp Leu Arg Pro Ser Ala Pro Gln Gln Glu Gly Val Ala Ser Lys Glu Lys Glu
AAGTGGCAGAGGAGCTTAGBAGCTGGGCTCCTGGGGTGCACCATGGGGGTTCTTGGTAGGGACCCATCCCCACAGAGCCTGCTTCCCTCCTGGG
Glu Val Ala Glu Glu
TCCAGGGCCAGAGGGGCTCCTGACCTTCCACAGCCCTACAGGGACTGTGTACAGGGCTAACCTGAACCTGAGTGGGAGGTCCTCCACGGATGAC
CCCTCAGGCATGGGGCAGAGAAGACAACCTTTCTCAGCCCACTCAGGGTGGCCATTAAGGCTGGGGCTGGAGCTGGGTGTGAGGTCATTCTCTCTCC
CAGCCCAGAGTGGGGAGACTAGAGGGCTACAGGCCAGCGTGGATGACCTGAAGAGCGCTCTCTGCCTTGAGACACCATCCCTCCTAGCACAAGGAGT
Ala Gln Ser Gly Gly Asp
GCCGCTTGAAGTACATCGGGCTGCCACGCTCCTGCCCTGTCTTCTGGCCACCTTGGCTGTCCACCTGTGCTGTGACCAACCTCACTGCCCTC
CCTGGGGCCACCCACCTCTGGTCTTCTGACCCCACTTATGCTGTGTGAATTTTTTTTTTAAATGATTCCAATAAAACTTGAGCCACTCCT

```

GTG→GAG variation within exon 4 results in an amino acid change: Val110Glu. According to their respective DNA sequences, EST clones 858268 and 400003 as well as the BCSG1 gene encode Glu110, while EST clones 726391 and EST68G11, and BAC clone 174P13 encode Val110. Both alleles of position 110 were found in CEPH control individuals.

Comparison of human γ -synuclein, with α - and β -synucleins

The γ -synuclein gene is predicted to be transcribed as an ~0.9-kb mRNA, not including the poly(A) tail, and to en-

code a 127 amino acid protein (Figs. 2, 3). Comparison of human protein sequences revealed that γ -synuclein shares 55.9 and 54.3% similarity, respectively, with the α - and β -synucleins. The most conserved regions are in the N-terminal portion of the protein, which contains repeated domains that display variations of a KTKEGV consensus sequence (Fig. 3). The mutation that we reported in the α -synuclein gene of patients with familial Parkinson's disease, between repeat domains IV and V, changed an alanine at position 53 to a threonine (Polymeropoulos et al. 1997). Interestingly, while β -synuclein also has an alanine at position 53, the human γ -synuclein normally has a threonine (Fig. 3), as do the rat and mouse orthologs of the γ -synuclein protein (Lavedan et al. in preparation).

Fig. 3 Alignment of human α -, β -, and γ -synuclein proteins. Residues that match exactly the consensus sequence are boxed. The consensus repeated domains are shown below the sequences (black boxes). For each protein, the amino acid number is indicated to the right. The number above the alignment refers to the consensus sequence (not shown)

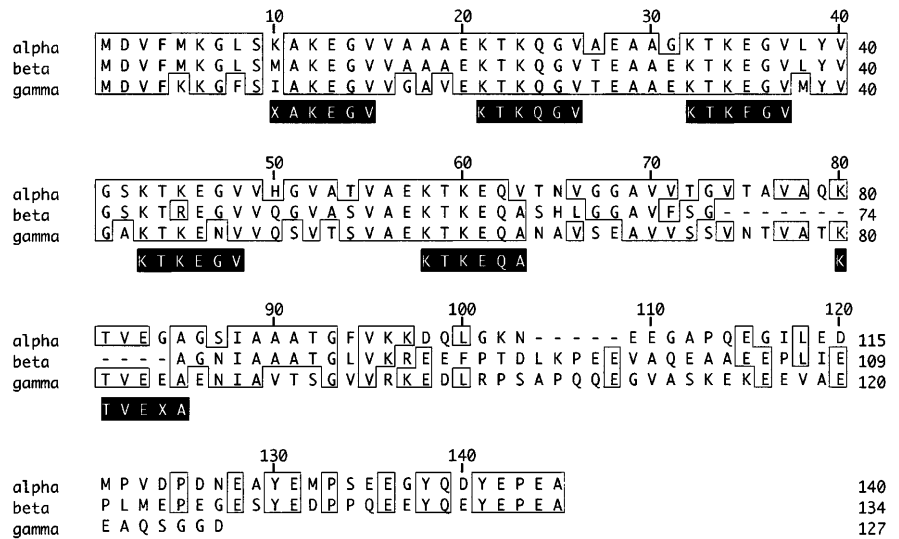
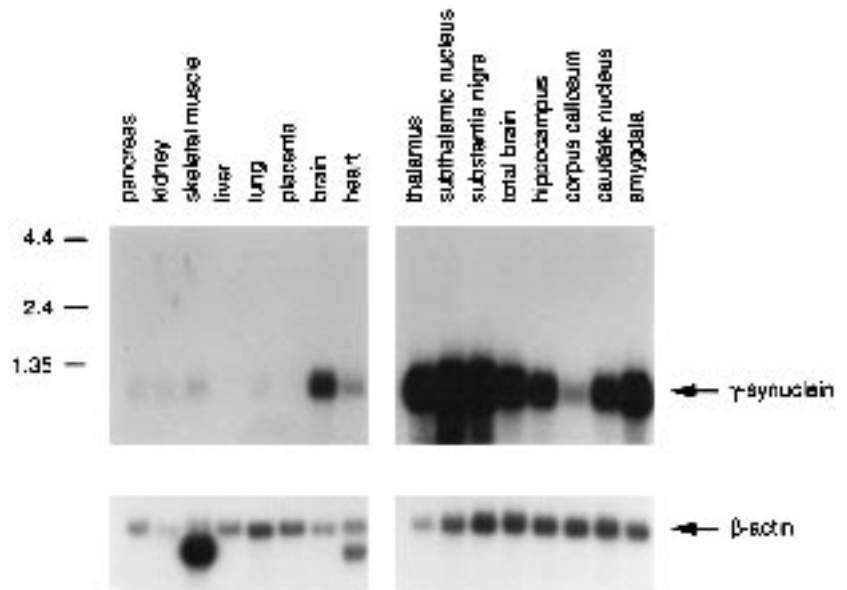


Fig. 4 Northern blot analysis of γ -synuclein; for details see Materials and methods. The sizes in kilobase pairs of a molecular weight standard RNA are indicated to the left



Tissue expression

We studied the expression of the human γ -synuclein gene by northern analysis, using a 197-bp DNA probe (γ 39- γ 40) from the 3' untranslated region that is not homologous to the α - or β -synuclein genes. The mRNA detected appeared to be ~1.0 kb, with no apparent splice variant (Fig. 4). The γ -synuclein gene is highly expressed in the substantia nigra as well as other regions of the brain: the thalamus, the subthalamic nucleus, the hippocampus, the caudate nucleus, and the amygdala. It is also moderately expressed in the corpus callosum, the heart and the skeletal muscle, and at much lower levels in pancreas, kidney, and lung. No signal was detected in liver or placenta. In dbEST, three cDNAs originated from the brain, one from the fetal heart, one from the retina, and one from a mixture of fetal liver and spleen (Table 2). Surprisingly, seven other clones came from the same library (NbHOT), made from an ovarian tumor (Table

2). Six of these seven clones are definitely independent, since their insert sizes are different and their sequences overlap (data not shown). Interestingly, when using the entire γ -synuclein sequence in a BLAST search, we found an almost perfect match with a recently released sequence of a gene called BCSG1, which is overexpressed in advanced infiltrating carcinoma of the breast (Ji et al. 1997). Furthermore, our database search indicated that the mouse γ -synuclein ortholog is also expressed in mammary gland (Lavedan et al. in preparation). The size of the mRNA as well as the pattern of expression of BCSG1 seem to be similar to that of γ -synuclein (Ji et al. 1997). BCSG1 transcripts of ~1.0 kb have been observed principally in brain but also in heart, ovary, testis, colon and spleen (Ji et al. 1997). We found five nucleotide differences between our consensus γ -synuclein sequence and the entire 550-bp BCSG1 sequence (Genbank accession number AF010126), with no difference in the 3' untranslated region, indicating that BCSG1

and γ -synuclein are likely to be the same gene. Only two of these differences resulted in an amino acid change: two glutamic acid residues present in γ -synuclein at positions 13 and 68 are lysines in BCSG1. These changes, E13K and E68K, were not present in the four CEPH control individuals, in BAC 174P13, or in clone HRBAA27. The glutamic acid at position 13 is indeed contained within the first repeated domain and is conserved in all known synuclein sequences (α , β , and γ) including human (Fig. 3), rodent, bovine, canary, and electric ray. Similarly, sequences of clone 858268 and EST68G11, as well as of the four CEPH control DNAs, predict a glutamic acid at position 68 as in the rodent ortholog (Lavedan et al. in preparation).

Discussion

We have identified and characterized a new member of the synuclein gene family, the γ -synuclein gene. This gene is highly homologous to the α - and β -synucleins. By a combination of radiation hybrid mapping and metaphase FISH, we have localized this new gene to chromosome 10q23.

We have sequenced all five coding exons of the human γ -synuclein gene, and developed exon-specific PCR assays. The 1.0-kb mRNA observed on northern blots was consistent in size with that predicted from the alignment of all ESTs and PCR product sequences.

The γ -synuclein gene seems to be identical to a gene (BCSG1) overexpressed in breast cancer specimens (Ji et al. 1997). Our survey of the EST database indicated that the human γ -synuclein gene might also be overexpressed in an ovarian tumor. Several types of genes have been shown to be involved in both breast and ovarian cancers through distinct mechanisms: mutations in tumor suppressor genes TP53 (Kupryjanczyk et al. 1993; Teneriello et al. 1993), BRCA1 (Miki et al. 1994), BRCA2 (Wooster et al. 1995), and MTS1 (Kamb et al. 1994); amplification and/or overexpression of proto-oncogenes such as *HER-2/neu* (Slamon et al. 1989) and AIB1 (Anzick et al. 1997). The observation of up-regulation of BCSG1 in malignant breast epithelial cells of infiltrating carcinoma mainly of the Comedo type (Ji et al. 1997) suggests that this gene might also be a proto-oncogene. Furthermore, it is noteworthy that one of the cloned γ -synuclein cDNAs, EST75830, originated from the pineal gland, which is thought to influence the production of gonadotrophic hormones known to be implicated in breast cancer. These observations suggest that it would be important to investigate whether γ -synuclein shares a common pathway with any of the proteins involved in tumor development and/or progression of breast, ovarian, and testicular cancers, in particular its potential relationship with the estrogen pathway.

There are currently about 100 independent sequences in the GenBank databases with high homology to α -synuclein. All synuclein sequences available to date can be assigned to three distinct proteins, suggesting that the synuclein family contains only three members: α -, β -, and γ -synuclein. The high homology between γ -synuclein and

α - and β -synucleins suggests a similar function that remains unknown. As a first step toward determining the role of these proteins, their patterns of expression have been examined. It has been shown that both α - and β -synucleins are concentrated in presynaptic nerve terminals (Maroteaux et al. 1988; Nakajo et al. 1993; Jakes et al. 1994). The rat ortholog of γ -synuclein has previously been reported as being a sensory neuron protein highly expressed in dorsal root ganglia, present also in heart and spleen and at a lower level in liver and kidney (Akopian and Wood 1995). We have shown that the human γ -synuclein gene is highly expressed in various areas of the brain, moderately in the corpus callosum, the heart, and skeletal muscle, and at a much lower level in pancreas, kidney, and lung. This result is consistent with the data reported for BCSG1, which also appears to be expressed at low levels in ovary, testis and colon (Ji et al. 1997). The high level of expression of γ -synuclein in the substantia nigra, the main region of neuronal degeneration in the brains of patients with Parkinson's disease, and its high homology to α -synuclein, which is mutated in some familial cases of Parkinson's disease (Polymeropoulos et al. 1997), points to γ -synuclein as a new candidate gene for Parkinson's disease.

Knowing the localization, sequence, and structure of the γ -synuclein gene, including oligonucleotide sequences and conditions for PCR amplification of each exon, will facilitate the search for mutations in this gene and allow examination of its involvement in Parkinson's disease and in other neurodegenerative disorders, as well as in the development and/or progression of breast and ovarian tumors.

Acknowledgements The authors would like to thank Dr. S. F. Suchy for critical review of the manuscript, and Jeffrey Rubenstein for his technical assistance. This research was funded in part by a grant from the National Parkinson's Foundation, and the intramural program of the National Human Genome Research Institute of the National Institutes of Health. The experiments comply with the current laws of the USA.

References

- Akopian AN, Wood JN (1995) Peripheral nervous system-specific genes identified by subtractive cDNA cloning. *J Biol Chem* 270:21264–21270
- Altschul SF, Madden TL, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Anzick SL, Kononen J, Walker RL, Azorsa DO, Tanner MM, Guan XY, Sauter G, Kallioniemi OP, Trent JM, Meltzer PS (1997) AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* 277:965–968
- Chen X, Silva HA de, Pettenati MJ, Rao PN, George-Hyslop P St, Roses AD, Xia Y, Horsburgh K, Ueda K, Saitoh T (1995) The human NACP/alpha-synuclein gene: chromosome assignment to 4q21.3–q22 and TaqI RFLP analysis. *Genomics* 26:425–427
- Couch FJ, Castilla LH, Xu J, Abel KJ, Welch P, King SE, Wong L, Ho PP, Merajver S, Brody LC (1995) A YAC-, P1-, and cosmid-based physical map of the BRCA1 region on chromosome 17q21. *Genomics* 25:264–273
- Dutra AS, Mignot E, Puck JM (1996) Gene localization and syntenic mapping by FISH in the dog. *Cytogenet Cell Genet* 74:113–117

- Gasser T, Müller-Myhsok B, Wszolek ZK, Dürr A, Vaughan JR, Binifati V, Meco G, Bereznaï B, Oehlmann R, Agid Y, Brice A, Wood N (1997) Genetic complexity and Parkinson's disease. *Science* 277:388–389
- George JM, Jin H, Woods WS, Clayton DF (1995) Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron* 15:361–372
- Jakes R, Spillantini MG, Goedert M (1994) Identification of two distinct synucleins from human brain. *FEBS Lett* 345:27–32
- Ji H, Liu YE, Jia T, Wang M, Liu J, Xiao G, Joseph BK, Rosen C, Shi YE (1997) Identification of a breast cancer-specific gene, BCSG1, by direct differential cDNA sequencing. *Cancer Res* 57:759–764
- Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigian SV, Stockert E, Day RS, Johnson BE, Skolnick MH (1994) A cell cycle regulator potentially involved in genesis of many tumor types. *Science* 264:436–440
- Kozak M (1991) An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol* 115:887–903
- Krüger R, Kuhn W, Müller T, Woitalla D, Graeber M, Kosel S, Przuntek H, Epplen JT, Schols L, Riess O (1998) Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nat Genet* 18:106–108
- Kupryjanczyk J, Thor AD, Beauchamp R, Merritt V, Edgerton SM, Bell DA, Yandell DW (1993) p53 gene mutations and protein accumulation in human ovarian cancer. *Proc Natl Acad Sci USA* 90:4961–4965
- Maroteaux L, Scheller RH (1991) The rat brain synucleins; family of proteins transiently associated with neuronal membrane. *Mol Brain Res* 11:335–343
- Maroteaux L, Campanelli JT, Scheller RH (1988) Synuclein: a neuron-specific protein localized to the nucleus and presynaptic nerve terminal. *J Neurosci* 8:2804–2815
- Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hussey C, Tran T, McClure M, Frye C, Hattier T, Phelps R, Haugen-Strano A, Katcher H, Yakumo K, Gholami Z, Shaffer D, Stone S, Bayer S, Wray C, Bogden R, Dayanath P, Ward J, Tonin P, Narod S, Bristow P, Norris F, Helvering L, Morrison P, Rosteck P, Lai M, Barrett JC, Lewis C, Neuhausen S, Cannon-Albright L, Goldgar D, Wiseman R, Kamb A, Skolnick MH (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266:66–71
- Nakajo S, Tsukada K, Omata K, Nakamura Y, Nakaya K (1993) A new brain-specific 14-kDa protein is a phosphoprotein. Its complete amino acid sequence and evidence for phosphorylation. *Eur J Biochem* 217:1057–1063
- 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, Iorio G Di, Golbe LI, Nussbaum RL (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 276:2045–2047
- Scott WK, Stajich JM, Yamaoka LH, Speer MC, Vance JM, Roses AD, Pericak-Vance MA, and the Deane Laboratory Parkinson Disease Research Group (1997) Genetic complexity and Parkinson's disease. *Science* 277:387–388
- Shibasaki Y, Baillie DA, Clair D St, Brookes AJ (1995) High-resolution mapping of SNCA encoding alpha-synuclein, the non-A beta component of Alzheimer's disease amyloid precursor, to human chromosome 4q21.3–q22 by fluorescence in situ hybridization. *Cytogenet Cell Genet* 71:54–55
- Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A (1989) Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244:707–712
- Spillantini MG, Divane A, Goedert M (1995) Assignment of human alpha-synuclein (SNCA) and beta-synuclein (SNCB) genes to chromosomes 4q21 and 5q35. *Genomics* 27:379–381
- Teneriello MG, Ebina M, Linnoila RI, Henry M, Nash JD, Park RC, Birrer MJ (1993) p53 and Ki-ras gene mutations in epithelial ovarian neoplasms. *Cancer Res* 53:3103–3108
- Walter MA, Spillet DJ, Thomas P, Weissenbach J, Goodfellow PN (1994) A method for constructing radiation hybrid maps of whole genomes. *Nat Genet* 7:22–28
- Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G, Barfoot R, Hamoudi R, Pate S, Rice C, Biggs P, Hashim Y, Smith A, Connor F, Arason A, Gudmundson J, Ficenec D, Kelsell D, Ford D, Tonin P, Bishop DT, Spurr NK, Ponder BAJ, Eeles R, Peto J, Devilee P, Cornelisse C, Lynch H, Narod S, Lenoir G, Egilsson V, Barkadottir RB, Easton DF, Bentley DR, Futreal PA, Ashworth A, Stratton MR (1995) Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378:789–792