

## FAMILIAL AGGREGATION OF PARKINSON'S DISEASE IN ICELAND

SIGURLAUG SVEINBJÖRNSDÓTTIR, M.D., ANDREW A. HICKS, PH.D., THORLÁKUR JÓNSSON, PH.D., HJÖRVAR PÉTURSSON, B.S., GRÉTAR GUÐMUNDSSON, M.D., MICHAEL L. FRIGGE, PH.D., AUGUSTINE KONG, PH.D., JEFFREY R. GULCHER, M.D., PH.D., AND KÁRI STEFÁNSSON, M.D., PH.D.

**ABSTRACT**

**Background** The role of genetics in early-onset Parkinson's disease has been established, but whether there is a genetic contribution to the more common, late-onset form remains uncertain.

**Methods** We reviewed the medical records and confirmed the diagnosis of Parkinson's disease in 772 living and deceased patients in whom the disease had been diagnosed during the previous 50 years in Iceland. With the use of an extensive computerized data base containing genealogic information on 610,920 people in Iceland during the past 11 centuries, several analyses were conducted to determine whether the patients were more related to each other than random members of the population (control subjects).

**Results** Patients with Parkinson's disease, including a subgroup of 560 patients with late-onset disease (onset at >50 years of age), were significantly more related to each other than were subjects in matched groups of controls, and this relatedness extended beyond the nuclear family. The risk ratio for Parkinson's disease was 6.7 (95 percent confidence interval, 4.3 to 9.6) for siblings, 3.2 (95 percent confidence interval, 1.2 to 7.8) for offspring, and 2.7 (95 percent confidence interval, 1.6 to 3.9) for nephews and nieces of patients with late-onset Parkinson's disease.

**Conclusions** Late-onset Parkinson's disease has a genetic component as well as an environmental component. (N Engl J Med 2000;343:1765-70.)

©2000, Massachusetts Medical Society.

**P**ARKINSON'S disease is an important neurodegenerative disorder affecting middle-aged and elderly persons. Its causes are largely unknown, but there is evidence that the disease has a genetic component. In a few large families with early-onset Parkinson's disease (onset at  $\leq$ 50 years of age) or juvenile Parkinson's disease (onset during childhood), the disease is transmitted as an autosomal dominant or recessive trait resulting from mutations in the genes encoding  $\alpha$ -synuclein and parkin, respectively.<sup>1-8</sup> However, in the majority of families affected by Parkinson's disease, the disease appears to skip generations, irrespective of the age of onset. Therefore, Parkinson's disease appears to be a complex, multifactorial disease resulting from interaction between one or more genes and the environment.

Although the disease is considered to be sporadic in most patients, persons with a family history of Parkinson's disease are at increased risk. Among first-degree relatives of patients, the risk is 2 to 14 times

the risk in members of unaffected families.<sup>9,10</sup> The increase in risk among first-degree relatives may result not only from genetic susceptibility, however, but also from ascertainment bias (i.e., relatives of patients with a given disease may be more likely than average to seek medical attention for that disease) or shared environmental factors. In several, but not all, studies of twins, the rate of concordance was no higher among monozygotic twins than among dizygotic twins.<sup>11-16</sup> This, together with the discovery that parkinsonism may be caused by toxic agents such as 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine, has led to increased emphasis on the role of environmental factors, especially in patients with late-onset Parkinson's disease. However, studies of late-onset Parkinson's disease in twins are hampered by age dependency, since the disease may develop later in the second twin or the second twin may die of another cause before the onset of symptoms. Studies in families are limited because often little is known about the proband's genealogic background and the health status of relatives outside the nuclear family.

Population-based studies, coupled with genealogic information, may represent a more complete method for assessing genetic contributions to common diseases. We studied a group of patients with Parkinson's disease, including the majority of patients in Iceland in whom Parkinson's disease had been diagnosed during the previous 50 years, and assessed their relatedness with use of a comprehensive genealogic data base of most Icelanders who have ever lived to adulthood to look for further evidence of a genetic component of the disease.

**METHODS****Patients**

This epidemiologic study, in which we used encrypted medical information, was approved by the National Bioethics Commission of Iceland and the Data Protection Commission of Iceland. Patients were identified from two sources. First, medical notes and, if applicable, the death certificates of 470 patients included in a total-population survey carried out in Iceland from 1953 to 1963<sup>17</sup> were independently reviewed by two neurologists. Patients were considered to have Parkinson's disease if they had at least two of the following signs: tremor, rigidity, bradykinesia, and postural

From the National University Hospital (S.S., G.G.) and deCODE genetics (A.A.H., T.J., H.P., M.L.F., A.K., J.R.G., K.S.), Reykjavik, Iceland; and the Department of Human Genetics, University of Chicago, Chicago (A.K.). Address reprint requests to Dr. Stefánsson at deCODE genetics, Lynghals 1, Reykjavik 110, Iceland, or at kstefans@decode.is, or to Dr. Sveinbjörnsdóttir at the National University Hospital, Reykjavik, Iceland, or at sigurl@rsp.is.

instability.<sup>18</sup> Seventy-six of these patients were excluded because of postencephalitic parkinsonism or because the diagnosis was uncertain. Second, an ongoing, population-based study begun in 1994 identified an additional 420 patients from a variety of sources, including the Icelandic Parkinson's Disease Society, information from neurologists and general practitioners, and records of prescriptions for levodopa and other drugs commonly given to patients with Parkinson's disease. All nursing homes and other homes for elderly persons in Reykjavik and those in approximately 70 percent of the rest of Iceland were visited to examine these 420 patients. Those thought to have multiple-system atrophy, progressive supranuclear palsy, or drug-induced Parkinson's disease were excluded, as were those who had no response to levodopa. After these exclusions, 378 of the 420 patients remained; 292 of them were examined by one of the authors, a specialist in movement disorders, after the patients' written informed consent had been obtained.

The final combined group of 772 patients (394 plus 378) may have included a small number of patients from the early survey<sup>17</sup> who had misdiagnosed parkinson-like syndromes (such as progressive supranuclear palsy, multiple-system atrophy, or corticobasal ganglionic degeneration), because such patients could not be retrospectively excluded without a histopathological examination. Patients who received a diagnosis of Parkinson's disease after the early survey but who died before the initiation of the current survey were not included, whereas those who received a diagnosis after the early survey and remained alive are included. Information about the age at onset was obtained for 693 of the 772 patients; in 560 of them symptoms had begun at an age greater than 50 years. The overall group of 772 patients and this subgroup of 560 patients were studied in separate analyses.

#### Genealogic Data Base

We are electronically registering all available genealogic information for the past 11 centuries in Iceland in a computerized, relational data base that contained 610,920 names at the time of the study, including the names of all 270,000 living Icelanders.<sup>19</sup> Control groups were selected from among these 610,920 members of the population. Data on the 772 study patients, along with the entire genealogic data base, were reversibly encrypted by the Data Protection Commission of Iceland before being sent to our laboratory.<sup>20</sup> We developed algorithms that find all ancestors in the data base who are related to each member of an input group within a given number of generations. Other algorithms identify, for each person in an input group, all relatives of a specific type, such as siblings or aunts or uncles. In this study, these algorithms allowed us to identify pedigrees and to estimate kinship coefficients and risk ratios.

#### Kinship Coefficients

The kinship coefficient is one measure of the genetic relationship between two subjects. For example, with no consanguinity in previous generations, the kinship coefficient is  $\frac{1}{4}$  for siblings and other first-degree pairs of relatives,  $\frac{1}{8}$  for second-degree pairs of relatives,  $\frac{1}{16}$  for third-degree pairs of relatives, and so on, each value being half the expected fraction of the genome shared by these relatives. Formally, the kinship coefficient is defined as the probability that a randomly selected allele from each member of a pair of subjects was inherited from a common ancestor.<sup>21</sup> The average kinship coefficient of the patients in the current study was calculated by averaging the kinship coefficients of every possible pairwise combination of patients. To assess the effect of close relationships on the size of the average kinship coefficient, we also computed the average coefficients of only the pairs of patients who were not first-degree relatives and of only the pairs of patients who were not first- or second-degree relatives. For the latter calculations, all patients were included in the calculation of the average kinship coefficient, but not all possible pairwise combinations. Kinship coefficients for groups of controls were calculated similarly.

Because the pedigrees were extensive, the overall average kinship coefficient could not be calculated exactly. We used Monte

Carlo simulations to approximate the average kinship coefficient for each group (patients or control subjects) and ensured that the Monte Carlo errors had a negligible effect on the reported results.

#### Calculations of the Risk Ratio

The risk ratio for relatives of affected patients was defined as the risk of Parkinson's disease in the relatives divided by the risk in the general population; this ratio is directly related to the power to identify or map susceptibility genes.<sup>22</sup> Obtaining valid estimates of the risk ratio is not straightforward, since many sampling schemes lead to biased or inflated estimates.<sup>23</sup> The use of a population-based group of patients eliminates much of the potential sampling bias. In calculating the estimated risk of Parkinson's disease in relatives, we restricted our analyses to relatives born during the period covering the life span of the group of patients in question. We used the same restriction according to year of birth in estimating the risk in the general population for the given risk ratio.

#### Statistical Analysis

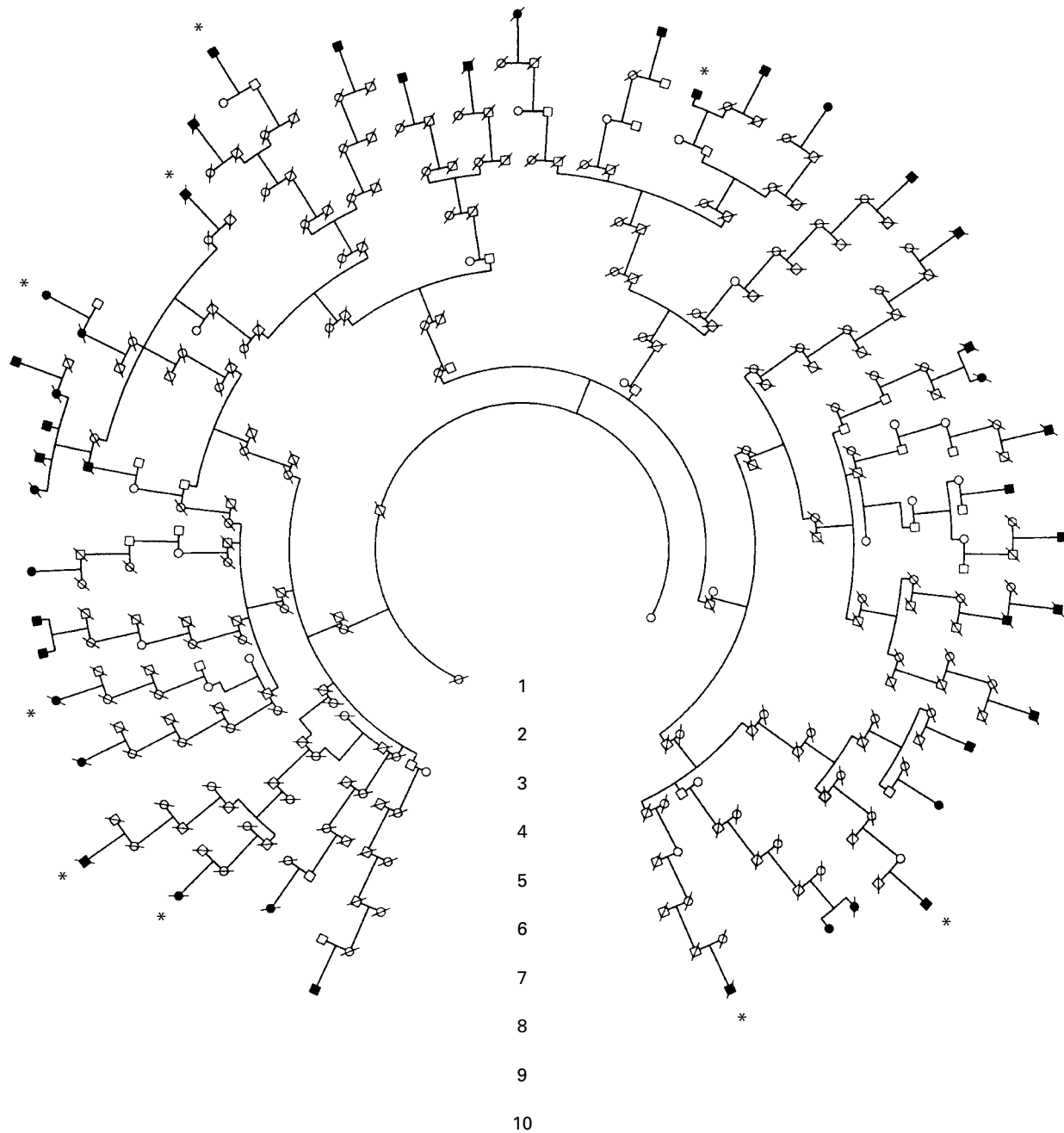
To assess the significance of the kinship coefficients and relative risks obtained for a given group of patients, we compared their observed values with the kinship coefficients and relative risks computed for 1000 independently drawn, matched groups of control subjects. Each patient was matched to a specific control subject in each control group. The control subjects were drawn at random from the genealogic data base, irrespective of their disease status, and had the same year of birth and the same number of ancestors recorded in the data base as did the patients to whom they were matched.

A reported P value of 0.005 for the relative risk would indicate that 5 of the 1000 matched control groups had values as large or larger than that for the patients. When none of the values computed for the control groups were larger than the value for the patients, we reported the P value as less than 0.001. The confidence intervals of the risk ratios for the patients were also calculated by comparing those values to the risk ratios for the control groups. Further details about the selection of the control groups and the construction of the confidence intervals for the risk ratios are provided on our Web site (<http://internotes.decode.is/nejm.nsf>).

## RESULTS

Our investigation of the group of patients with Parkinson's disease using the Icelandic genealogic data base led to the identification of many pedigrees containing two or more related patients with Parkinson's disease. Figure 1 shows a large pedigree containing 44 patients with early- or late-onset Parkinson's disease from a common founder. Some of these patients may therefore share disease allele or mutations that are identical by descent from this common ancestor.

To test whether the relatedness of the patients was significantly different from the background relatedness that occurs in a general population, we compared the average kinship coefficient of the patients with Parkinson's disease with that of the control subjects. For any two relatives, the kinship coefficient is approximately half the proportion of their genome shared as a result of common ancestry. The average kinship coefficient of the patients with Parkinson's disease, in both the entire group of 772 patients and the group of 560 patients with late-onset disease (81 percent of the 693 patients for whom we obtained data about the age at onset) was significantly differ-



**Figure 1.** A Pedigree Showing 44 Patients with Parkinson's Disease.

The patients with Parkinson's disease (solid symbols), previously thought to be largely unrelated, could be traced to a common ancestor, six generations (indicated by the numbers) before the oldest patient, with use of a genealogic data base. To protect the anonymity of the family, most of the unaffected relatives in this pedigree are not shown. The circles denote female family members, and the squares male family members. The asterisks indicate patients with early-onset Parkinson's disease. Slash marks denote family members listed in a local death registry.

ent from their respective control groups (Table 1). The patients were significantly more interrelated than the control subjects. This significance persisted for both the overall group of patients and for the subgroup with late-onset disease, even after first-degree pairs of relatives were excluded from the calculations. When, in addition, second-degree pairs of relatives were excluded from the calculations, the average kinship coefficient of the patients remained larger than the mean for the control subjects.

When we divided the patients according to their identification in the early survey<sup>17</sup> or the ongoing survey, the kinship coefficients of these subgroups

were similar to one another, and there were significant differences between the patients in each of the two subgroups and the control subjects. Estimates of the risk ratios for relatives of patients with Parkinson's disease are presented in Table 2. The risk ratios for siblings, offspring, and nephews and nieces are all significantly larger than 1 for both the entire group of patients with Parkinson's disease and the subgroup of patients with late-onset disease. Although both siblings and offspring are first-degree relatives, the risk ratio was higher for the former than for the latter (6.7 for siblings and 3.2 for offspring of patients with late-onset disease). Nephews and nieces are sec-

**TABLE 1. KINSHIP COEFFICIENTS OF PATIENTS WITH PARKINSON'S DISEASE AND OF MATCHED GROUPS OF CONTROL SUBJECTS.\***

GROUP	AVERAGE KINSHIP COEFFICIENT		P VALUE
	PATIENTS	CONTROL GROUPS	
All patients with Parkinson's disease (n=772)	2.7	2.0±0.1	<0.001
Excluding all pairs of first-degree relatives	2.1	1.9±0.1	<0.001
Excluding all pairs of first- and second-degree relatives	1.9	1.8±0.1	0.03
Patients with late-onset Parkinson's disease (n=560)	2.8	2.0±0.1	<0.001
Excluding all pairs of first-degree relatives	2.1	1.9±0.1	0.007
Excluding all pairs of first- and second-degree relatives	1.9	1.8±0.1	0.10

\*Plus-minus values are means ±SD. The kinship coefficient (shown here multiplied by 10,000) is the probability that a randomly selected allele from each member of a pair of subjects was inherited from a common ancestor.<sup>21</sup> The results remain significant after the contribution of first-degree relatives is removed. Removal of the contribution of first- and second-degree relatives decreases the significance only slightly. These results suggest that there are familial effects, and that they extend beyond the nuclear family. The standard deviation reported for the control groups is that of the average kinship coefficient for the 1000 groups of control subjects.

**TABLE 2. ESTIMATED RISK RATIOS FOR THE RELATIVES OF ALL PATIENTS WITH PARKINSON'S DISEASE AND FOR THE RELATIVES OF THE SUBGROUP OF PATIENTS WITH LATE-ONSET PARKINSON'S DISEASE.\***

GROUP	RELATIVE	RELATIVE'S BIRTH YEAR	TOTAL NO. OF RELATIVES	NO. OF AFFECTED RELATIVES	RISK RATIO (95% CI)	P VALUE
All patients with Parkinson's disease (n=772)	Siblings	1885-1930	2528	95	6.3 (4.4-7.8)	<0.001
	Offspring	1915-1930	737	14	3.0 (1.5-5.0)	0.001
	Nephews and nieces	1915-1930	2815	43	2.4 (1.7-3.2)	<0.001
	First cousins	1885-1930	5048	43	1.4 (0.9-2.0)	0.10
	Spouses	1885-1930	532	6	1.9 (0.5-5.7)	0.16
Patients with late-onset Parkinson's disease (n=560)	Siblings	1885-1930	1911	57	6.7 (4.3-9.6)	<0.001
	Offspring	1915-1930	647	9	3.2 (1.2-7.8)	0.01
	Nephews and nieces	1915-1930	2376	28	2.7 (1.6-3.9)	<0.001
	First cousins	1885-1930	3638	21	1.3 (0.6-1.9)	0.30
	Spouses	1885-1930	396	2	1.1 (0.0-4.3)	0.55

\*Risk ratios and 95 percent confidence intervals (CI) were calculated for the 772 patients with Parkinson's disease and separately for the subgroup of 560 patients with late-onset Parkinson's disease. Confidence intervals for spouses are very wide because of small samples (for example, a person usually has many more nephews and nieces than spouses). The denominator for calculating the risk ratios is the prevalence of Parkinson's disease in the population during that period. For the period from 1885 to 1930, the denominator for the overall group of patients with Parkinson's disease was 0.0059 (638/107,798) and the denominator for the subgroup of patients with late-onset Parkinson's disease was 0.0044 (479/107,798). For the period from 1915 to 1930, the denominator for the overall group was 0.0063 (298/47,507) and the denominator for the late-onset subgroup was 0.0043 (204/47,507).

ond-degree relatives, and their estimated risk ratios were significantly greater than 1. Cousins are third-degree relatives; their estimated risk ratios were larger than 1 but not significantly so (Table 2). The risk ratios for spouses were not significant.

To investigate whether the significant familial association in the patients with late-onset Parkinson's disease might be due entirely or in part to the inheritance of longevity that has been observed in Iceland,<sup>24</sup> we drew an additional 1000 groups of control subjects, which we also matched to the age distribution of the patients. There were no substantial differences between the patients and these additional control groups in the P values for the kinship coefficients or risk ratios.

### DISCUSSION

In this study, we reexamined the issue of genetic and environmental contributions to Parkinson's disease by analyzing computerized genealogic data in relation to information about a population-based group of patients. Although this approach cannot eliminate the possibility of every type of ascertainment bias, it had several benefits. The population-based group allowed us to avoid the sampling bias that might result from proband identification and oversampling of families with several affected members. Using the population-based genealogic data base, we also avoided the customary classification of patients into familial and sporadic cases, because any familial relationship between patients, even when distant, was known. In addition, the use of the Icelandic population, with its single-payer health care system with universal access, may have reduced certain types of diagnostic bias.

Our data are consistent with the possibility that Parkinson's disease has a familial component that may be masked since this complex and multifactorial disease can skip generations. This familial component may arise from a combination of environmental and genetic factors. By demonstrating that the familial clustering of Parkinson's disease extends beyond the nuclear family, we have provided more evidence that the disease has a genetic component. Moreover, we found that the spouses of patients with Parkinson's disease were not at increased risk for the disease. It is therefore unlikely that a shared environmental factor, late in life, accounts for the Parkinson's disease in patients drawn from the entire Icelandic population for many years. However, there is a noteworthy difference between the risk ratios for siblings and those for offspring. This may indicate a role for some shared environmental factor early in life, as has been suggested for Alzheimer's disease,<sup>25</sup> or recessive inheritance of susceptibility.

The results of our study also challenge the concept of etiologic differences between early-onset and late-onset Parkinson's disease.<sup>15</sup> Since several studies in twins revealed no genetic component in late-onset

Parkinson's disease, and since there are rare pedigrees containing many patients with early-onset Parkinson's disease caused by single-gene mutations, it has been proposed that early-onset Parkinson's disease is likely to have a substantial genetic component.<sup>15</sup> Accordingly, the causes of early-onset Parkinson's disease might differ from those of late-onset Parkinson's disease, although clinically and pathologically these disorders are similar. Approximately 20 percent of patients in this study for whom we had data about the age at onset had early-onset Parkinson's disease, but attempts to cluster the patients with early-onset disease into pedigrees revealed no families with a highly penetrant mendelian pattern of inheritance. This suggests that most early-onset cases of Parkinson's disease are not due to single, highly penetrant genes. Rather, just as in the late-onset cases, the early-onset disorder skips generations. In fact, as the pedigree in Figure 1 shows, the early-onset and late-onset cases of Parkinson's disease may even cosegregate within the same family. Although these findings may be specific for patients with Parkinson's disease in Iceland, the disease in this population has the same phenotype, prevalence, and age of onset as that in most other Western countries.

There has been a recent trend to discount the possibility that genetic factors contribute to the late-onset form of the disease, which represents the majority of cases of Parkinson's disease. Although the search for environmental factors contributing to late-onset Parkinson's disease is important and should continue, our data suggest that the search to discover its genetic basis should also continue.

Supported in part by the National University Hospital Research Fund and by the Icelandic Research Council.

*We are indebted to the patients, the control subjects, and family members for their generous participation in this work; to the general practitioners and clinical neurologists who contributed information; and to the Icelandic Parkinson's Disease Society.*

### REFERENCES

1. Polymeropoulos MH, Higgins JJ, Golbe LI, et al. Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. *Science* 1996;274:1197-9.
2. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 1997;276:2045-7.
3. Kruger R, Kuhn W, Muller T, et al. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nat Genet* 1998;18:106-8.
4. Matsumine H, Yamamura Y, Hattori N, et al. A microdeletion of D6S305 in a family of autosomal recessive juvenile parkinsonism (PARK2). *Genomics* 1998;49:143-6.
5. Hattori N, Matsumine H, Asakawa S, et al. Point mutations (Thr240Arg and Ala311Stop) in the Parkin gene. *Biochem Biophys Res Commun* 1998;249:754-8. [Erratum, *Biochem Biophys Res Commun* 1998;251:666.]
6. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* 1998;392:605-8.
7. Abbas N, Lucking CB, Ricard S, et al. A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. *Hum Mol Genet* 1999;8:567-74.

8. Tassin J, Durr A, de Broucker T, et al. Chromosome 6-linked autosomal recessive early-onset parkinsonism: linkage in European and Algerian families, extension of the clinical spectrum, and evidence of a small homozygous deletion in one family. *Am J Hum Genet* 1998;63:88-94.
9. Gasser T. Genetics of Parkinson's disease. *Ann Neurol* 1998;44:Suppl 1: S53-S57.
10. Wood NW. Genetic risk factors in Parkinson's disease. *Ann Neurol* 1998;44:Suppl 1:S58-S62.
11. Duvoisin RC, Eldridge R, Williams A, Nutt J, Calne D. Twin study of Parkinson disease. *Neurology* 1981;31:77-80.
12. Ward CD, Duvoisin RD, Ince SE, Nutt JD, Eldridge R, Calne DB. Parkinson's disease in 65 pairs of twins and in a set of quadruplets. *Neurology* 1983;33:815-24.
13. Bharucha NE, Stokes L, Schoenberg BS, et al. A case-control study of twin pairs discordant for Parkinson's disease: a search for environmental risk factors. *Neurology* 1986;36:284-8.
14. Duvoisin RC. Genetics of Parkinson's disease. *Adv Neurol* 1987;45: 307-12.
15. Tanner CM, Ottman R, Goldman SM, et al. Parkinson disease in twins: an etiologic study. *JAMA* 1999;281:341-6.
16. Piccini P, Burn DJ, Ceravolo R, Maraganore D, Brooks DJ. The role of inheritance in sporadic Parkinson's disease: evidence from a longitudinal study of dopaminergic function in twins. *Ann Neurol* 1999;45:577-82.
17. Gudmundsson KR. A clinical survey of parkinsonism in Iceland. *Acta Neurol Scand* 1967;43:Suppl 33:1-61.
18. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967;17:427-42.
19. Gulcher J, Stefansson K. Population genomics: laying the groundwork for genetic disease modeling and targeting. *Clin Chem Lab Med* 1998;36:523-7.
20. Gulcher JR, Kristjansson K, Gudbjartsson H, Stefansson K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur J Hum Genet* 2000;8:739-42.
21. Genetic identity coefficients. In: Lange K. *Mathematical and statistical methods for genetic analysis*. New York: Springer-Verlag, 1997:70-84.
22. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 1990;46:222-8.
23. Guo S-W. Inflation of sibling recurrence-risk ratio, due to ascertainment bias and/or overreporting. *Am J Hum Genet* 1998;63:252-8.
24. Guðmundsson H, Guðbjartsson DE, Kong A, et al. Inheritance of human longevity in Iceland. *Eur J Hum Genet* 2000;8:743-9.
25. Mocerri VM, Kukull WA, Emanuel I, van Belle G, Larson EB. Early-life risk factors and the development of Alzheimer's disease. *Neurology* 2000;54:415-20.