

Huntington Disease: No Evidence for Locus Heterogeneity

P. MICHAEL CONNEALLY,¹ JONATHAN L. HAINES,² RUDOLF E. TANZI,² NANCY S. WEXLER,³
 GRACIELA K. PENCHASZADEH,³ PETER S. HARPER,⁴ SUSAN E. FOLSTEIN,⁵ JEAN J. CASSIMAN,⁶
 RICHARD H. MYERS,⁷ ANNE B. YOUNG,⁸ MICHAEL R. HAYDEN,⁹ ARTHUR FALEK,¹⁰ EDUARDO S. TOLOSA,¹¹
 SEBASTIAN CRESPI,¹² LUIGI DI MAIO,¹³ GOSTA HOLMGREN,¹⁴ MARIA ANVRET,¹⁵
 ICHIRO KANAZAWA,¹⁶ AND JAMES F. GUSELLA²

¹Department of Medical Genetics, Indiana University Medical Center, Indianapolis, Indiana 46223; ²Neurogenetics Laboratory, Massachusetts General Hospital, Boston, Massachusetts 02114; ³Psychiatric Institute, Columbia University, New York, New York 10032; ⁴Institute of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XN, United Kingdom; ⁵Department of Psychiatry, Johns Hopkins Hospital, Baltimore, Maryland 21205; ⁶Center of Human Genetics, University of Leuven, B 3008 Leuven, Belgium; ⁷Department of Neurology, Boston University Medical Center, Boston, Massachusetts 02118; ⁸Neuroscience Laboratory, University of Michigan, Ann Arbor, Michigan 48104; ⁹Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada V6T 2B5; ¹⁰Georgia Mental Health Institute, Atlanta, Georgia 30306; ¹¹Hospital Clinic 1, Villarroel, Barcelona, Spain; ¹²Departamento de Genetica, Universidad de las Islas Baleares, Palma, Spain; ¹³Clinica Neurologica II, Facolta di Medicina, 80131 Naples, Italy; ¹⁴Department of Pediatrics, University Hospital, Umea, Sweden; ¹⁵Department of Clinical Genetics, Karolinska Hospital, S-10401 Stockholm, Sweden; and ¹⁶Department of Neurology, University of Tsukuba, Ibaraki 300 31, Japan

Received February 27, 1989; revised April 18, 1989

A total of 63 families with Huntington disease (HD) were examined for linkage between HD and G8 (D4S10). The families included 57 Caucasian, four Black American, and two Japanese. The combined maximum lod score was 87.69 at $\Theta = 0.04$ (99% confidence interval 0.018–0.071). The maximum frequency of recombination was 0.03 in males and 0.05 in females. Fifty-seven families gave positive lod scores; five small families gave mildly negative lod scores. The maximum likelihood estimate of α , the proportion of linked loci, was 1.0 with a lower 99% confidence interval of 0.88. These data suggest that there is only one HD locus, although a second rare locus cannot be ruled out. © 1989 Academic Press, Inc.

INTRODUCTION

Huntington disease (HD) is a neurodegenerative disorder caused by an autosomal dominantly inherited defect. It is characterized by programmed premature death of localized nerve cells occurring throughout the brain, but most prominently in the striatum (Martin, 1984).

The three cardinal symptoms of this disorder are changes in personality, movement, and mentation. Temporally, cognitive symptoms may precede the appearance of abnormal movements by many years. Emotional disturbance is variable in HD but usually includes depression, irritability, apathy, sometimes vi-

olent outbursts, and, less commonly, hallucinations and delusions. Progressively more severe chorea, or involuntary movements of all parts of the body, is the most typical and early motoric symptom. Most patients gradually develop dystonia, bradykinesia, and other parkinsonian features as well. Dysarthria impairs speech beyond recognition until patients become mute. Choking is a frequent acute cause of death. Patients also lose memory (particularly short term), organizational abilities, and other cognitive capacities. They usually maintain orientation of time and place and also maintain recognition of their own identities and that of family and friends (Hayden, 1981).

The average age of onset is approximately 38 years, ranging from less than 5 years of age to more than 70 years. The precise age of onset in a given individual is difficult to ascertain due to the insidious and variable nature of the initial stage of the disease. The disease has a mean duration of approximately 19 years. Thus, the actuarial profile of the average HD gene carrier is a shortened life span of about 57 years, with the initial two-thirds of this period in a presymptomatic stage and the final one-third beginning with mild symptoms and progressing to complete impairment, both physical and psychological (Conneally, 1984).

In approximately 10% of cases the disease begins before the age of 20 years. A large proportion of such cases show a rigid parkinsonian form of the disease and have a more rapid clinical decline than the typical adult onset form. One intriguing and unexplained as-

pect of juvenile onset HD is the large excess of paternal transmission. In cases with juvenile onset, the father is approximately four times as likely to be affected than the mother (Merritt *et al.*, 1969).

Although invariably fatal, HD is the underlying rather than the immediate cause of death in patients. The most common primary causes of death are heart disease and aspiration-induced pneumonia (Haines and Conneally, 1986). Approximately 6% of HD patients in the United States commit suicide and 28% attempt suicide at least once (Farrer, 1986).

The basic gene defect in HD is unknown. Extensive neuropathological and neurochemical studies in HD (Martin and Gusella, 1987) reveal localized neuronal loss, particularly in the striatum, and decreases in the levels of a number of neurotransmitters in specific brain areas. Therapeutic strategies aimed at transmitter replacement have not been successful.

The discovery of the probe G8, which identifies the DNA marker *D4S10* linked to the HD gene (Gusella *et al.*, 1983), has introduced the possibility of cloning the mutant gene and developing effective therapeutic interventions based on an understanding of the gene defect. Closely linked DNA markers have also made presymptomatic and prenatal diagnostic testing now available. Given the late onset of the disorder and its severity, many at-risk individuals have expressed interest in a presymptomatic test which has both advantages and disadvantages: it can assist in planning but is an inevitably fatal diagnosis with no treatment available. Since the majority of HD gene carriers are born while their heterozygous parents are still asymptomatic, at-risk individuals would like to use prenatal diagnosis to ensure that, regardless of whether they carry the HD gene, they will have children who are free of the disease (Gusella *et al.*, 1984).

Presymptomatic and prenatal diagnoses rely on observing *D4S10*, the DNA marker closely linked to the HD gene. For the accuracy of this test, it is imperative that only one HD locus exist, or, if there is locus heterogeneity, that it can be recognized. An intensive effort has therefore been made to obtain a large number of informative HD families from as many ethnic and racial backgrounds as possible to test for locus heterogeneity.

MATERIALS AND METHODS

Families with a documented history of HD were ascertained. Those whose structures were deemed suitable for linkage analysis were invited to participate. A total of 63 families were examined for linkage between the HD gene and *D4S10*. Four were American Black, two were Oriental (Japanese), and the remainder were Caucasian from different parts of the world. Procedures were the same in the analyses of all families, although

some families had DNA typing for more RFLPs than others, as new markers were developed. DNA was prepared directly from peripheral blood lymphocytes, from fibroblasts, or from Epstein-Barr virus-transformed lymphoblastoid cell lines. Lymphoblast DNA was prepared using methods described elsewhere (Gusella *et al.*, 1979).

The *D4S10* locus probes G8, R7, pk082, and pk083 and the resulting polymorphisms have been previously described (Gusella *et al.*, 1984). The probes were labeled with ^{32}P following the procedure of Feinberg and Vogelstein (1983). Filters were prehybridized and hybridized as described in Gusella *et al.* (1983). The filters were washed under conditions equal to $0.1\times \text{SSC}$, 0.1% SDS at 65°C ; results were obtained by autoradiography at -70°C for 1–5 days.

Two-point linkage analysis was carried out using the program LIPED version 3 (Ott, 1974), and multipoint linkage analysis using the program package LINKAGE (version 3.5) described by Lathrop *et al.* (1984). At-risk individuals were assigned a probability of carrying the HD gene based on a cumulative age-of-onset distribution from data on 999 affected individuals from the National Huntington's Disease Research Roster at Indiana University. The probability that the offspring of an affected individual is asymptomatic at age i , given that he or she is heterozygous for the HD gene, is $1 - Y_i$ where Y_i is the cumulative probability of being affected by age i (Conneally *et al.*, 1984). Asymptomatic individuals under age 30 were not used in the linkage analysis, since they contribute little information on linkage and the authors did not want to discover genotypes of presymptomatic individuals. Occasionally, genotype data from at-risk persons were needed to reconstruct parental haplotypes; all such information was maintained in the laboratory under code, was accessed by only the absolute minimum number of investigators needed to analyze the data, and was not revealed to the family. The participating families knew in advance that these were research studies only and that they would not be provided with diagnostic feedback since it is important that diagnostic information be provided in the context of intensive counseling and follow-up support. Individuals were told that HD in their families was found to be linked to *D4S10* and whether they were genetically informative, if they requested this information.

The existence of linkage heterogeneity, i.e., the possibility of more than one locus being involved in the etiology of HD, was tested using the program HOMOG (Ott, 1985, using the method of Smith, 1963). This program allows for three alternatives for explaining the data: (i) two loci, one linked and one unlinked; (ii) the same as (i) but including a possible sex difference in recombination frequencies; and (iii) two loci, both linked, but with different recombination frequencies.

TABLE 1
Lod Scores for HD vs G8 $\Theta_M = \Theta_F$

Location (number of families)	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	Max	
												Θ	Lod
Atlanta (1)	3.02	2.98	2.93	2.88	2.83	2.78	2.73	2.67	2.62	2.56	2.51	0.00	3.02
Baltimore (7)	—	17.05	18.73	19.49	19.89	20.07	20.13	20.08	19.98	19.81	19.58	0.06	20.13
Belgium (8)	4.00	3.92	3.84	3.75	3.69	3.59	3.50	3.42	3.34	3.23	3.11	0.00	4.00
Boston (3)	—	0.68	1.24	1.55	1.75	1.89	1.99	2.06	2.11	2.14	2.16	0.11	2.16
Iowa (1)	3.79	3.72	3.65	3.58	3.50	3.43	3.35	3.27	3.19	3.11	3.03	0.00	3.79
Italy (1)	1.93	1.90	1.87	1.84	1.81	1.78	1.75	1.71	1.68	1.64	1.61	0.00	1.93
Japan (2)	1.98	1.94	1.89	1.85	1.80	1.75	1.71	1.66	1.61	1.57	1.52	0.00	1.98
Michigan (2)	—	3.49	3.68	3.75	3.77	3.76	3.72	3.68	3.62	3.56	3.49	0.04	3.77
Minnesota (2)	—	0.42	0.75	0.94	1.06	1.16	1.22	1.27	1.31	1.33	1.34	0.12	1.35
Spain (5)	—	0.46	0.11	0.43	0.64	0.79	0.90	0.98	1.03	1.08	1.11	0.12	1.13
Sweden (2)	3.59	3.55	3.51	3.47	3.42	3.37	3.32	3.26	3.20	3.14	3.08	0.00	3.59
Vancouver (12)	—	8.35	8.49	8.49	8.39	8.35	8.23	8.10	7.94	7.79	7.62	0.03	8.49
Venezuela (1)	—	19.04	19.03	18.90	18.72	18.50	18.26	18.00	17.72	17.43	17.13	0.01	19.04
Wales (16)	—	16.56	16.70	16.61	16.43	16.18	15.89	15.57	15.24	14.88	14.51	0.02	16.70
Total (63)		83.14	86.42	87.53	87.69	87.40	86.70	85.73	84.59	83.27	81.80		

One Lod confidence interval: $0.022 < \Theta < 0.060$.

Two Lod confidence interval: $0.018 < \Theta < 0.071$.

The program output provides confidence intervals for the proportion of linked families.

The 99% confidence interval for the recombination frequency Θ is obtained by drawing a horizontal line 2 lod scores below the peak lod score on the lod score curve, as shown in Fig. 1. The two points of intersection between the straight line and the lod score curve mark the two endpoints of the confidence interval (Conneally *et al.*, 1985).

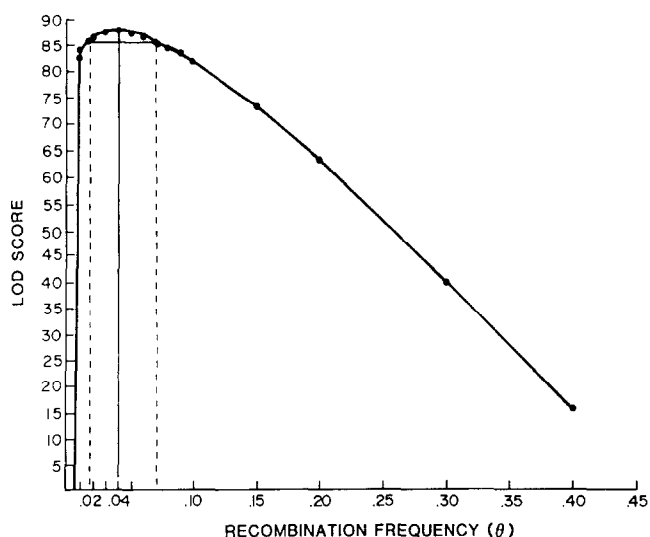


FIG. 1. The lod score for HD and *D4S10* loci. The solid vertical line represents the peak lod score. The dashed lines are the 99% confidence limits for Θ and are obtained by drawing a horizontal line 2 lod scores below the peak on the curve; the limits are obtained from the intersections of the line and the curve.

Evidence for significant disequilibrium between the HD and *D4S10* loci would suggest a very limited number of different mutational events at the HD locus. If the marker locus is very close to the HD locus, one would expect that a specific haplotype(s) would be traveling with the HD gene much more frequently than would be expected if the two loci were independent of one another.

The *Hind*III endonuclease cuts at two polymorphic sites with the *D4S10* locus, giving four haplotypes: A, B, C, and D (Gusella *et al.*, 1983). *Eco*RI cuts at one variable site with two alleles, 1 and 2. In a total of 44 families, the *Hind*III-*Eco*RI combined haplotypes traveling with the HD gene could be determined. Control haplotype frequencies were determined from a group of individuals without the HD allele.

RESULTS

The lod scores for the 63 families for recombination values (Θ) from 0 to 10 cM together with the maximum lod score and its corresponding recombination frequency, are summarized in Table 1. The combined maximum lod score was 87.69 at $\Theta = 0.04$. This is shown graphically in Fig. 1. The 99% confidence interval for Θ is 0.018 to 0.071. Additionally, joint lod scores were obtained where male and female recombination frequencies were allowed to vary from each other. The maximum frequency of recombination in males was 0.03 with 95% confidence intervals of 0.01 and 0.07. The maximum frequency in females was 0.05

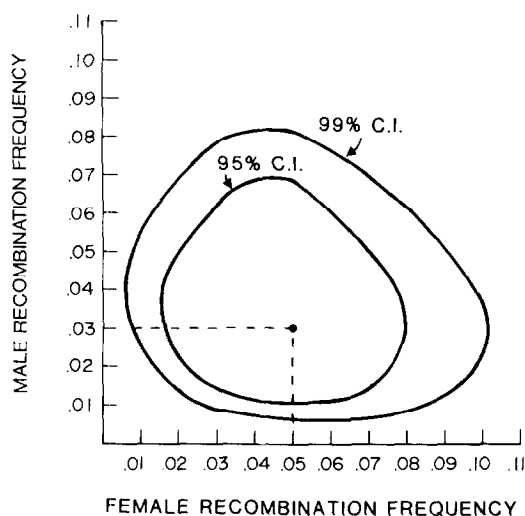


FIG. 2. The 95 and 99% confidence ellipses for male and female recombination values. The peak lod score is at the intersection of the two dashed lines.

with 95% confidence intervals of 0.02 and 0.08. The peak lod score at $\hat{\Theta}_M = 0.03$, $\hat{\Theta}_F = 0.05$ was 87.74. This is shown graphically in Fig. 2.

Five of the 63 families had negative lod scores, all of which were relatively small. The values were never below -2 within the confidence interval for Θ . Therefore, the HD gene in these families could not be assumed to be at an alternate location.

There was no evidence for locus heterogeneity; that is, one linked and one unlinked locus, among these analyzed families. The maximum likelihood estimate of α , the proportion of linked loci, was 1.0, with a lower 99% confidence interval of 0.88. When the possibility of two linked loci was tested, the program HOMOG gave maximum likelihood estimates of 0.02 for $\hat{\Theta}_1$ and 0.04 for $\hat{\Theta}_2$. However, these estimates were not significantly different from the one-locus hypothesis ($\chi^2 = 0.08$, $P > 0.48$).

Table 2 shows the distribution of G8 haplotypes for *HindIII* and *EcoRI* combined among the HD families. There are no significant differences between the haplotypes of those affected with HD and the haplotypes of normal controls ($\chi^2_{(7)} = 6.53$, $P > 0.40$). Thus, there is no evidence for genetic disequilibrium between the *D4S10* and the HD loci.

DISCUSSION

The discovery of a genetic marker linked to the gene for Huntington disease should eventually lead to the cloning of the mutant gene and characterization of its defect. A more immediate outcome of the finding is the capacity for presymptomatic and prenatal identification of gene carriers. These preclinical tests have been

described in detail elsewhere (Meissen *et al.*, 1988). It is imperative that presymptomatic testing for the presence of HD in asymptomatic individuals be delivered as accurately as possible. Catastrophic consequences could result from erroneous test results, particularly a false positive test, where an individual is informed that he or she is at high risk for carrying the HD gene when, in fact, the person is homozygous for the normal non-HD allele. Informed genetic counseling for presymptomatic testing is not possible in the absence of thorough information regarding nonallelic heterogeneity for the HD locus. If locus heterogeneity existed, the predictive test would not be useful in those families where the HD mutation occurred at a locus unlinked to the current markers. Unfortunately, the only way to definitively rule out locus heterogeneity is to perform a full linkage analysis on each family desiring testing, but this is usually not feasible, since large informative families are not common. In fact, only 40% of at-risk individuals listed in the National Huntington's Disease Research Roster have the minimal family structure needed for an informative "predictive" test (as few as two to three relatives sufficient to establish the phase of the marker with respect to the disease allele in the given family).

Fortunately, as the results of these analyses show, it is likely that only one locus for HD exists. The 99% lower confidence interval for the proportion of linked families is 88%. This is a conservative estimate, especially since there is no significant evidence that any of the families are unlinked to *D4S10*. While it is still possible that an HD mutation might occur at another site, at best it would appear to be a rare occurrence which should have minimal effect on the conduct of predictive testing. If a family in which the HD allele does not appear to be near the telomere of chromosome 4 is located, all diagnoses must be scrupulously reconfirmed by clinicians blind to the genotypes, and neuropathological analyses of postmortem brain tissue should be sought.

TABLE 2

Distribution of G8 Haplotypes (*HindIII* and *EcoRI*) in Families with HD versus Controls

G8		Families with HD	Control frequency
<i>HindIII</i>	<i>EcoRI</i>		
A	1	8	0.27
A	2	20	0.33
B	1	0	0.06
B	2	4	0.07
C	1	4	0.09
C	2	6	0.12
D	1	1	0.02
D	2	1	0.03

A precise estimate of the recombination fraction is also important for accurate prediction and for estimating the approximate distance in kilobases of the marker from the HD locus. Although the data suggest that the recombination frequency in females is greater than that in males (0.05 vs 0.03), this is not significant, and it is reasonable to use the overall estimate of 4 cM as the distance between the marker and the gene.

The test for linkage disequilibrium between *D4S10* and the HD locus showed no significant difference from that of controls. This is not surprising, considering that the two loci are 4 cM apart. If, as is suspected, the HD mutation is ancient, there would have been sufficient time for the two loci to have reached approximate equilibrium.

ACKNOWLEDGMENTS

The authors are most grateful to the families who willingly participated in this research. The work was supported by NINCDS Contract NO1NS7-2311, NINCDS Grants 1RO1 NS22031 and 1RO1 NS16367, MRC (Canada), and the Hereditary Disease Foundation. Computing time was provided by Indiana University Computing Services.

REFERENCES

1. CONNEALLY, P. M. (1984). Huntington's disease: Genetics and epidemiology. *Amer. J. Hum. Genet.* **36**: 506-526.
2. CONNEALLY, P. M., WALLACE, M. R., GUSELLA, J. F., AND WEXLER, N. S. (1984). Huntington's disease: Estimation of heterozygote status using linked genetic markers. *Genet. Epidemiol.* **1**: 81-88.
3. CONNEALLY, P. M., EDWARDS, J. H., KIDD, K. K., LALOUEL, J.-M., MORTON, N. E., OTT, J., AND WHITE, R. (1985). Report of the committee on methods of linkage analysis and reporting. *Cytogenet. Cell Genet.* **21**: 356-359.
4. FARRER, L. A. (1986). Suicide and attempted suicide in Huntington disease: Implications for preclinical testing of persons at risk. *Amer. J. Med. Genet.* **24**: 305-311.
5. FEINBERG, A. P., AND VOGELSTEIN, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.
6. GUSELLA, J. F., VARSANYI-BREINER, D., KAO, F. T., JONES, C., PUCK, T. T., KEYS, C., ORKIN, S., AND HOUSMAN, D. (1979). Precise localization of the human β -globin gene complex on chromosome 11. *Proc. Natl. Acad. Sci. USA* **76**: 5239-5243.
7. GUSELLA, J. F., WEXLER, N. S., CONNEALLY, P. M., NAYLOR, S. L., ANDERSON, M. A., TANZI, R. E., WATKINS, P. C., OTTINA, K., WALLACE, M. R., SAKAGUCHI, A. Y., YOUNG, A. B., SHOULSON, I., BONILLA, E., AND MARTIN, J. B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature (London)* **306**: 234-239.
8. GUSELLA, J. F., TANZI, R. E., ANDERSON, M. A., HOBBS, W., GIBBONS, K., RASCHTCHIAN, R., GILLIAM, T. C., WALLACE, M. R., WEXLER, N. S., AND CONNEALLY, P. M. (1984). DNA markers for nervous system diseases. *Science* **225**: 1320-1326.
9. HAINES, J. S., AND CONNEALLY, P. M. (1986). Causes of death in Huntington disease as reported on death certificates. *Genet. Epidemiol.* **3**: 417-423.
10. HAYDEN, M. R. (1981). "Huntington's Chorea," Springer-Verlag, New York.
11. LATHROP, G. M., LALOUEL, J.-M., JULIER, C., AND OTT, J. (1984). Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci. USA* **81**: 3443-3446.
12. MARTIN, J. B. (1984). Huntington's disease: New approaches to an old problem. *Neurology (NY)* **34**: 1059-1072.
13. MARTIN, J. B., AND GUSELLA, J. F. (1987). Huntington's disease: Pathogenesis and management. *N. Engl. J. Med.* **315**: 1267-1276.
14. MEISSEN, G. J., MYERS, R. H., MASTROMAURO, C. A., KOROSHETZ, W. J., KLINGER, K. W., FARRER, L. A., WATKINS, P. A., GUSELLA, J. F., BIRD, E. D., AND MARTIN, J. B. (1988). *N. Engl. J. Med.* **318**: 535-542.
15. MERRITT, A. D., CONNEALLY, P. M., RAHMAN, N. F., AND DREW, A. L. (1969). Juvenile Huntington's chorea. In "Progress in Neurogenetics" (A. Barbeau and T. R. Brunette, Eds.), pp. 645-650, Exerpa Medica Foundation, Amsterdam.
16. OTT, J. (1974). Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. *Amer. J. Hum. Genet.* **26**: 588-597.
17. OTT, J. (1985). "Analysis of Human Genetic Linkage," Johns Hopkins Press, Baltimore.
18. SMITH, C. A. B. (1963). Testing for heterogeneity of recombination fraction values in human genetics. *Ann. Hum. Genet.* **27**: 175-182.