Am. J. Hum. Genet. 60:625–629, 1997

Hereditary Spastic Paraplegia: LOD-Score Considerations for Confirmation of Linkage in a Heterogeneous Trait

Marie-Pierre Dubé,[1] Melinda A. Mlodzienski,[2] Zoha Kibar,[1] Martin R. Farlow,[3] George Ebers,[4] Peter Harper,[5] Edwin H. Kolodny,[6] Guy A. Rouleau,[1] and Denise A. Figlewicz[2]

[1]Montreal General Hospital Research Institute, McGill University, and Centre for Research in Neuroscience, Montreal; [2]University of Rochester, Rochester, NY; [3]Indiana University, Indianapolis; [4]University of Western Ontario, London, Ontario; [5]University Hospital of Wales, Cardiff; and [6]New York University School of Medicine, New York

Summary

Hereditary spastic paraplegia (HSP) is a degenerative disorder of the motor system, defined by progressive weakness and spasticity of the lower limbs. HSP may be inherited as an autosomal dominant (AD), autosomal recessive, or an X-linked trait. AD HSP is genetically heterogeneous, and three loci have been identified so far: SPG3 maps to chromosome 14q, SPG4 to 2p, and SPG4a to 15q. We have undertaken linkage analysis with 21 uncomplicated AD families to the three AD HSP loci. We report significant linkage for three of our families to the SPG4 locus and exclude several families by multipoint linkage. We used linkage information from several different research teams to evaluate the statistical probability of linkage to the SPG4 locus for uncomplicated AD HSP families and established the critical LOD-score value necessary for confirmation of linkage to the SPG4 locus from Bayesian statistics. In addition, we calculated the empirical *P*-values for the LOD scores obtained with all families with computer simulation methods. Power to detect significant linkage, as well as type I error probabilities, were evaluated. This combined analytical approach permitted conclusive linkage analyses on small to medium-size families, under the restrictions of genetic heterogeneity.

Introduction

Hereditary spastic paraplegia (HSP) (also known as familial spastic paraparesis and Strumpell-Lorrain syndrome) is defined by progressive weakness and spasticity of the lower extremities (Sutherland 1975; Harding 1993). The disease comprises clinically and genetically diverse disorders, classified according to the mode of inheritance and based on whether progressive spasticity occurs in isolation, referred to as "pure" or "uncomplicated," or with other neurologic abnormalities, referred to as "complicated," including optic neuropathy, retinopathy, extrapyramidal disturbance, dementia, ataxia, ichthyosis, mental retardation, and deafness. The age at onset of the disorder is highly variable between and within families and ranges from early childhood to the 9th decade of life. The rate of symptom progression and the extent of disability are also variable. No treatment to prevent, retard, or reverse the progressive disability is available. The more common form of uncomplicated HSP is inherited as an autosomal dominant (AD) trait; however, there are cases of autosomal recessive (AR) and X-linked inheritance as well. There are two distinct forms of X-linked HSP, located at Xq28 (SPG1) (Jouet et al. 1994) and Xq22 (SPG2) (Saugier-Veber et al. 1994). Uncomplicated AR HSP maps to the pericentric region of chromosome 8 (SPG5) and was shown to be genetically heterogeneous (Hentati et al. 1994*a*). Uncomplicated AD HSP is genetically heterogeneous, and linkage was established to loci on chromosomes 14q (SPG3) (Hazan et al. 1993), 2p (SPG4) (Hazan et al. 1994; Hentati et al. 1994*b*), and 15q (SPG4a) (Fink et al. 1995). The existence of additional loci is indicated by families for whom the known loci are excluded. The high degree of similarity between HSP kindreds with linkage to different genetic loci suggests a common pathway in which the gene products would act together. Clinical features and genetics of HSP have been reviewed recently (Fink et al. 1996).

We carried out linkage analysis of 21 uncomplicated AD HSP families by use of microsatellite markers at the SPG3, SPG4, and SPG4a loci. We used linkage information from several different research teams provided by the Hereditary Spastic Paraplegia Working Group (Fink et al. 1996) to evaluate the statistical probability of linkage to the SPG4 locus for uncomplicated AD HSP families and established the critical LOD-score value necessary for confirmation of linkage to the SPG4 locus from Bayesian statistics. Using computer simulation methods, we also calculated the empirical *P*-values for the LOD

Received October 16, 1996; accepted for publication December 23, 1996.

Address for correspondence and reprints: Dr. Guy A. Rouleau, Montreal General Hospital Research Institute, 1650 Cedar Avenue, Room L7-126, Montreal H3G 1A4, Quebec, Canada.

0002-9297/97/6003-0020$02.00

scores obtained with all families. Power to detect significant linkage, as well as type I error probabilities, were evaluated. We report significant linkage of three AD HSP families to markers at the SPG4 locus. Exclusion of linkage in several families provides additional evidence for the genetic heterogeneity of this disorder.

Subjects, Material, and Methods

Families

The identification, clinical characterization, and collection of blood samples of HSP families was achieved in collaboration with neurologists working in the United States and Canada. Twenty-one different families were used in this study: a total of 357 individuals, including 122 affected. Pedigree information was obtained from many relatives in each family. The diagnosis was made by certified neurologists, and patients with early signs of the disease were carefully examined and independently judged to assess the severity of the spasticity on the basis of the Ashworth scale, muscle stretch reflexes, presence or absence of Babinski signs, difficulty with normal, heel, toe, or tandem gaits, heel-to-shin tests of cerebellar function, and lower extremity position sense tests.

Linkage Studies

We genotyped 15 microsatellite loci by use of lymphocyte DNA from 21 pure AD HSP families. Genetic linkage to the three HSP loci SPG3, SPG4, and SPG4a was tested. DNA was extracted from lymphocytes by standard procedures. All markers reported are available from Research Genetics, Inc., as human map pairs, except for D2S266 (from Généthon). PCR amplifications were carried out using the conditions recommended by Research Genetics.

Programs from the FASTLINK package, version 3.0P for UNIX compiled on the Solaris 2.4 operating system, were used (Cottingham et al. 1993; Schäffer et al. 1994) as well as utility programs of the LINKAGE package version 5.1 for UNIX and DOS operating systems (Lathrop and Lalouel 1984; Lathrop et al. 1984, 1986). Two-point linkage analysis was performed with the MLINK subroutine, and maximum-likelihood estimates were calculated with ILINK for families showing positive two-point linkage results. Exclusion analysis was performed on all families individually, from four-point linkage analyses using the FASTLINK LINKMAP program. We declared exclusion when LOD scores of ≤ -2.0 were obtained for the entire candidate regions.

All analyses were done using an autosomal dominant HSP gene model, with a frequency of 10^{-5} and equal female-to-male recombination rates. For each marker, allele frequencies were assumed to be equal. Although this can lead to bias for positive linkage in criterion-based analyses, it should not affect linkage based on empirical P-values. For families with known ages at onset, a probability density function was used in the construction of age-dependent penetrance. For the other families, a cumulative distribution of affected individuals' age was used.

The map distance between the markers was determined from the two independently constructed linkage maps of Généthon and CHLC. It was necessary to integrate information from the two maps in order to establish the location of the CHLC markers D2S405 and D15S217. Because of discrepancies, the genetic distance was estimated from the average value of the two maps, when necessary.

Proportion of Linked Families

The report from the Hereditary Spastic Paraplegia Working Group establishes the proportion of linked and excluded families to each of the HSP loci (Fink et al. 1996). They report a total of 16 families (our data not included) tested for linkage to the AD HSP loci, from which 10 link to SPG4, 2 to SPG3, and 1 to SPG4a, and 3 families showed exclusion to all three loci. The prior probability of linkage to the three HSP loci was determined using the lower confidence limit of the 95% confidence interval from the binomial sampling distribution, on the proportion of families linked to each loci. The proportion of families linked to SPG4 is .625, and the lower confidence bound was calculated to be .35. The probability $P = .35$ provides the prior probability of linkage to the SPG4 locus for uncomplicated AD HSP families, on the assumption of consistent clinical diagnosis. The small number of families linked to SPG3 and SPG4a were not sufficient to exclude 0 as a possible proportion of linked families from the 95% confidence interval.

Calculation of the Critical LOD Score and Significance Level

The Bayesian posterior probability of linkage can be expressed as

$$P(H_1|F) = \frac{P(H_1)R}{P(H_1)R + P(H_0)}, \qquad (1)$$

where $P(H_1)$ is the probability of linkage to a given locus, $P(H_0)$ of independent assortment of the marker and the disease locus, $P(F)$ of the data, and $P(H_1|F)$ is the posterior probability of linkage defined to be .95. The likelihood ratio R is

$$R = \frac{P(F|H_1)}{P(F|H_0)}.$$

By use of the prior probability of linkage of uncomplicated AD HSP families to the SPG4 locus, R can be

calculated from equation (1), where $P(H_1) = .35$, $P(H_0) = .65$, with a posterior probability of linkage $P(H_1|F) = .95$, gives $R = 35.29$. The corresponding critical LOD-score value is $Z_0 = \log_{10}(35.29) = 1.55$. The associated significance level for this Z_0 is $\alpha \leq .028$, as defined by the inequality $\alpha = P(Z_{max} \geq Z_0|H_0) \leq 10^{-Z_0}$. This significance level will be used as the critical P-value necessary to declare linkage to SPG4.

Simulation and Empirical P-Values

We used the computer simulation programs SIMULATE and ISIM (Ott 1989; Weeks et al. 1990) to determine the empirical P-value associated with the observed maximum LOD score (Z_{max}) for families showing positive LOD scores. The typed markers were simulated 1,000 times, independent of the disease, and analyzed for exactly the same individuals and mode of inheritance as used in the linkage analysis. The empirical P-value was calculated to be the proportion of replicates that generated maximum LOD scores equal to or higher than the observed Z_{max} by use of the proportion's upper bound of the binomial 95% confidence interval.

Linkage was declared significant when a family's empirical P-value was less or equal to the critical P-value ($P \leq .028$), with the exception of family 28, which was tested under the criterion of $P \leq .014$, corrected for multiple testing (see below). The critical P-value remains unchanged ($P \leq .001$) for SPG3 and SPG4a. Linkage was declared significant when a family's empirical P-value was smaller or equal to the critical P-value.

The increase in type I error caused by multiple testing at the SPG4 locus is assumed to be counterbalanced by an increase in prior probability of linkage when exclusion is reported at the other AD HSP loci. To control for the bias in family 28, where there is no exclusion at SPG3 or SPG4a, the critical P-value .028 is divided by 2 (for the two flanking markers), giving a critical P-value of .014, as described by the equation $\alpha_1 = \alpha_g/g$, where α_1 is the significance for an individual comparison, α_g is the overall significance level, and g is the number of comparisons (Ott 1991).

Calculation of the Average Z_{max}

We used the computer simulation package SLINK and ISIM (Ott 1989; Weeks et al. 1990) to compute the expected average maximum LOD scores of each family. This value represents the average maximum LOD score obtained from the simulation of 500 replicates of four linked alleles at a recombination fraction of .01, for exactly the same individuals and mode of inheritance as used in our linkage analyses.

Results

We genotyped 357 individuals, including 122 affected individuals from 21 pure AD HSP families, at 15 microsatellite markers. The chromosome 14q markers cen–D14S306–(2.0 cM)–D14S79–(2.8 cM)–D14S266–(3.0 cM)–D14S269–(7.1 cM)–D14S66–tel contain the SPG3 locus localized to a region flanked by D14S288 and D14S281 (Gispert et al. 1995) located 0.1 cM telomeric to D14S266 and 3.2 cM centromeric to D14S66, respectively. Probable linkage in families 44 and 45 is observed. The observed maximum LOD scores are reported in table 1 with empirical P-values obtained from simulation analyses. Family 44 had a maximum LOD score of 1.60 with marker D14S79 at $\theta = .0$, with an empirical P-value of .0157. Family 45 had a maximum LOD score of 1.77 with marker D14S306, with an empirical P-value of .0073. There is exclusion of linkage to SPG3 in eight of our families.

The chromosome 2p markers tel–D2S405–(4.5 cM)–D2S170–(0.7 cM)–D2S146–(2.2 cM)–D2S352–(0 cM)–D2S400–(4.2 cM)–D2S367–cen contain the SPG4 locus flanked by D2S352 and D2S367. Family 7 is linked with a maximum LOD score of 2.17 for marker D2S367 at $\theta = .0$. Family 21 is linked with a maximum LOD score of 1.77 for both D2S352 and D2S367 at $\theta = .0$. We report linkage in family 28, as well, according to the significance of the empirical P-value. This family had a maximum LOD score of .77 for all three markers D2S405, D2S170, and D2S367, at $\theta = .0$. Figure 1 represents the range of LOD scores (95%) obtained from simulation of 1,000 unlinked replicates of family 28. It is interesting to note that the P-values associated with the critical LOD score of 1.55, based on the simulation analysis, are .0184, .0208, and .0037, for families 7, 21, and 28, respectively. Eleven families were excluded for SPG4. Multipoint linkage analysis performed on these three families peaked with a LOD score of 4.80 at $\theta = .0$ from marker D2S352 for the order D2S400–D2S352–SPG4–D2S367.

The chromosome 15q markers cen–D15S128–(8.3 cM)–D15S156–(0.5 cM)–D15S217–(5.3 cM)–D15S165–tel contain the SPG4a locus flanked by D15S128 and D15S156. None of our families linked to this locus, and exclusion is reported for seven families.

Discussion

The information provided by the Hereditary Spastic Paraplegia Working Group has allowed us to redefine the probability of finding linkage with AD HSP families to one of the three AD HSP loci known to date. This is the first report of an extensive linkage study with AD HSP families to use this information to define significant linkage. We have shown that this approach is a very powerful tool to handle problems of linkage analysis under genetic heterogeneity, especially for smaller families.

From our analysis of 21 families, we report three sig-

Table 1

Summary of Linkage and Simulation Results

Family	Average Maximum LOD Score E [Z(θ = .01)]	Z_{max} SPG3	Z_{max} SPG4	Z_{max} SPG4a
1	3.186	Exclusion	Exclusion	Exclusion
2	.417	Inconclusive	Exclusion	Inconclusive
3	.930	Exclusion	Exclusion	Inconclusive
5	5.791	Exclusion	Exclusion	Exclusion
6	.768	Inconclusive	Exclusion	Inconclusive
7	1.249	Exclusion	2.17 $P \leq .0056$	Exclusion
9	.354	Inconclusive	Inconclusive	Inconclusive
10	1.520	Inconclusive	Exclusion	Inconclusive
19	.126	Inconclusive	Inconclusive	Inconclusive
21	.767	Exclusion	1.77 $P \leq .0056$	Exclusion
28	.258	Inconclusive	.77 $P \leq .0037$	Inconclusive
29	3.360	Exclusion	Exclusion	Inconclusive
30	.297	Exclusion	Inconclusive	Inconclusive
31	.223	Inconclusive	Inconclusive	Inconclusive
32	.180	Inconclusive	Inconclusive	Inconclusive
33	.184	Exclusion	Inconclusive	Inconclusive
37	.499	Not evaluated	Exclusion	Exclusion
42	.013	Inconclusive	Inconclusive	Inconclusive
43	1.037	Inconclusive	Exclusion	Exclusion
44	1.127	1.60 $P \leq .0157$	Exclusion	Inconclusive
45	1.090	1.71 $P \leq .0073$	Exclusion	Exclusion

NOTE.—P represents the empirical P-value.

nificant linkages to the SPG4 locus on chromosome 2p by use of a combined analysis of prior probability of linkage to SPG4 and empirical P-values generated from simulation method. Suggestive linkage for two families to SPG3 on chromosome 14q was also found. A total of 26 exclusions and 31 inconclusive LOD-score results were obtained. As additional families are assigned to the SPG3 and SPG4a loci, the power of detection of linkage to these loci will increase as is seen in SPG4. The relatively low empirical P-values for linkage of families 44 and 45 to SPG3 suggests that their posterior probability of linkage may prove significant once additional information of prior probability of linkage to this locus is found. This report, which is based on individual family results, provides essential linkage information to perform more precise evaluations when testing linkage to known loci and for genomewide searches for the other HSP loci. Testing families individually has the disadvantage of generating more inconclusive LOD scores, however, especially with smaller families.

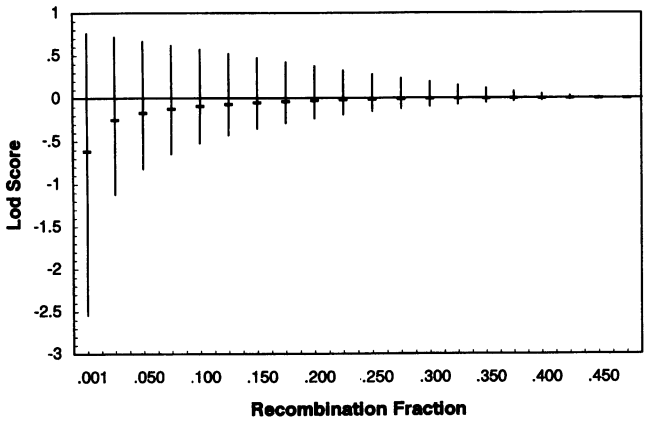

Figure 1 Distribution of LOD-score results from 1,000 simulated replicates of unlinked allele data with family 28. The average LOD score is indicated by the horizontal marking. The LOD-score range represents the 95% confidence interval from binomial sampling distribution, obtained from the simulation analysis, analyzed for linkage over a series of recombination fractions.

Small families are of limited use in genomewide searches for new disease loci or positional cloning when there is significant genetic heterogeneity. One can wait for genes to be cloned and then look for the mutation responsible for the disorder in a small family. However, not all genes will be cloned at the same time, there may be many loci, and in some cases mutation screening may be very difficult. Evidence of linkage to the region will be a determining factor for investment of further work on any locus and kindred.

The change in criteria for linkage analysis presented in this report is applicable for testing of linkage in heterogeneous disorders to loci that have already been

mapped and from which there is sufficient information available to determine with statistical significance the proportion of families, with the same clinical phenotype, linked to a locus under study. As shown here, this has the effect of reducing the critical LOD score required to declare linkage. However, since not all families have the same power to detect linkage, the empirical *P*-value associated with an observed maximum LOD score is a better determinant of significant linkage than is the sole criterion of exceeding a critical LOD-score value. This is particularly true for a reduced critical LOD score. Simulation of linked and unlinked alleles at markers is a powerful method for evaluating the type I error specific to individual families, and analysis of linkage with reduced LOD-score criteria should not be performed without it.

The sample of families used to evaluate the prior probability of linkage with AD HSP families was not the result of an extensive literature search but was based on information pooled at the Spastic Paraplegia Workshop and presented in a report by Fink et al. (1996). This sampling ensures equal input of information for tested, untested, excluded, and linked families, from each research team present at the workshop, and is less prone to bias for positive results. We feel that workshops such as the one reported by Fink et al. provide an excellent platform for the exchange and integration of information otherwise not necessarily publicly available.

Genetic heterogeneity is high for AD HSP. Larger families with linkage exclusion under one-locus and two-loci analyses will provide the most powerful tool for genomewide searches of the yet unknown HSP gene(s).

Acknowledgments

The authors wish to thank all the patients and their families and the physicians who have participated in this study. Financial support was provided by the Muscular Dystrophy Association—USA (to D.A.F.). M.-P.D. is supported by the Network of Centers of Excellence of Canada. G.A.R. is supported by the Medical Research Council of Canada.

References

Cottingham RW Jr, RM Idury, AA Schäffer (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53:252–263

Fink JK, Brocade Wu C-t, Jones SM, Sharp GB, Lange BM, Lesicki A, Reinglass T, et al (1995) Autosomal dominant familial spastic paraplegia: tight linkage to chromosome 15q. Am J Hum Genet 56:188–192

Fink JK, Heiman-Patterson T, Bird T, Cambi F, Dubé M-P, Figlewicz DA, Haines JL, et al (1996) Hereditary spastic paraplegia: advances in genetic research. Neurology 46: 1507–1514

Gispert S, Santos N, Damen R, Voit T, Schulz J, Klockgether T, Orozco G, et al (1995) Autosomal dominant familial spastic paraplegia: reduction of the FSP1 candidate region on chromosome 14q to 7 cM and locus heterogeneity. Am J Hum Genet 56:183–187

Harding AE (1993) Hereditary spastic paraplegia. Semin Neurol 13:333–336

Hazan J, Fontaine B, Bruyn RPM, Lamy C, van Deutekom JCT, Rime C-S, Dürr A, et al (1994) Linkage of a new locus for autosomal dominant familial spastic paraplegia to chromosome 2p. Hum Genet 3:1569–1573

Hazan J, Lamy C, Melki J, Munnich A, de Recondo J, Weissenbach J (1993) Autosomal dominant familial spastic paraplegia is genetically heterogeneous and one locus maps to chromosome 14q. Nat Genet 5:163–167

Hentati A, Pericak-Vance MA, Hung W-Y, Laing N, Boustany R-M, Hentati F, Ben Hamida M, et al (1994*a*) Linkage of "pure" autosomal recessive familial spastic paraplegia to chromosome 8 markers and evidence of genetic locus heterogeneity. Hum Genet 3:1263–1267

Hentati A, Pericak-Vance MA, Lennon F, Wasserman B, Hentati F, Juneja T, Angrist MH, et al (1994*b*) Linkage of a locus for autosomal dominant familial spastic paraplegia to chromosome 2p markers. Hum Genet 3:1867–1871

Jouet M, Rosenthal A, Armstrong G, McFarlane J, Stevenson R, Paterson J, Metzenberg, et al (1994) X-linked spastic paraplegia (SPG1), MASA syndrome, and X-linked hydrocephalus result from mutations in the L1 gene. Nat Genet 7:402–407

Lathrop GM, Lalouel J-M (1984) Easy calculations of LOD scores and genetic risks on small computers. Am J Hum Genet 36:460–465

Lathrop GM, Lalouel J-M, Julier C, Ott J (1984) Strategies for multilocus analysis in humans. Proc Natl Acad Sci USA 81:3443–3446

Lathrop GM, Lalouel J-M, White RL (1986) Construction of human genetic linkage maps: likelihood calculations for multilocus analysis. Genet Epidemiol 3:39–52

Ott J (1989) Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci USA 86:4175–4178

——— (1991) Analysis of human genetic linkage, 2d ed. Johns Hopkins University Press, Baltimore

Saugier-Veber P, Munnich A, Bonneau D, Rozet JM, Le Merrer M, Gil R, Boesplfug-Tanguy O (1994) X-linked spastic paraplegia and Pelizaeus-Merzbacher disease are allelic disorders at the proteolipid protein locus. Nat Genet 6:257–262

Schäffer AA, Gupta SK, Shriram K, Cottingham RW Jr (1994) Avoiding recomputation in linkage analysis. Hum Hered 44:225–237

Sutherland JM (1975). Familial spastic paraplegia. In: Vinken PJ, Bruyn GW (eds) Handbook of clinical neurology. Vol 22: system disorders and atrophies, part II. North Holland, Amsterdam, pp 420–431

Weeks DE, Ott J, Lathrop GM (1990) SLINK: a general simulation program for linkage analysis. Am J Hum Genet Suppl 47:A204