Gene for von Recklinghausen Neurofibromatosis Is in the Pericentromeric Region of Chromosome 17

D. BARKER, E. WRIGHT, K. NGUYEN, L. CANNON, P. FAIN, D. GOLDGAR, D. T. BISHOP, J. CAREY, B. BATY, J. KIVLIN, H. WILLARD, J. S. WAYE, G. GREIG, L. LEINWAND, Y. NAKAMURA, P. O'CONNELL, M. LEPPERT, J.-M. LALOUEL, R. WHITE, M. SKOLNICK

Linkage analysis of 15 Utah kindreds demonstrated that a gene responsible for von Recklinghausen neurofibromatosis (NF) is located near the centromere on chromosome 17. The families also gave no evidence for heterogeneity, indicating that a significant proportion of NF cases are due to mutations at a single locus. Further genetic analysis can now refine this localization and may lead to the eventual identification and cloning of the defective gene responsible for this disorder.

VON RECKLINGHAUSEN NEUROFIBROMATOSIS (NF) is one of the most common autosomal dominant conditions in humans. Crowe *et al.* estimated the prevalence to be 1 in 2500 to 3300 (*1*). From this figure and the assumption of a single autosomal dominant disorder, the mutation rate was estimated to be approximately 1 in 10,000 gametes per generation. This is one of the highest mutation rates in humans (*2*). Characteristic manifestations of NF include cutaneous pigmentary changes, multiple benign neurofibromas, and iris hamartomas or Lisch nodules (*3*). Affected individuals are at risk for a diverse array of complications including osseous and central nervous system disorders and neoplasias, which affect approximately 15% of children and adults with NF (*4*). Presentation can range from a child with multiple congenital anomalies to an adult with a solid tumor malignancy. Linkage studies in large families are especially important for understanding NF since the high mutation rate and variable expressivity are suggestive of a heterogeneous genetic etiology.

If a genetic map including NF could be established, it would be useful in genetic counseling, for resolving the issue of genetic heterogeneity and for the eventual understanding of the molecular defect. Therefore, genetic linkage studies of NF have been initiated in a number of laboratories (*5*). An exclusion map has been derived from a compilation of their results (*6*). As part of that report we presented preliminary data for a battery of 39 markers. The only suggestion of linkage was between NF and probe p3-6, a clone containing alpha satellite DNA from the centromere of chromosome 17. Odds favoring linkage were 15 to 1 at a recombination fraction of 0.1. Although this ratio is far from significant, the combination of this information with the large number of chromosomes excluded by the collective results led us to focus our attention on completing the typing and analysis of the centromeric marker and other chromosome 17 probes on our NF families.

Families segregating for NF were identified by J. Carey through the Division of Medical Genetics of the Department of Pediatrics, University of Utah School of Medicine and the National Neurofibromatosis Foundation (Utah Chapter). Additional families were identified by G. Meyers, through the Primary Children's Hospital, Salt Lake City, Utah. From each patient, we obtained (with informed consent) the medical history, physical exam, and slit-lamp eye exam; we also collected saliva and approximately 40 ml (up to 20 ml in young children) of peripheral blood for DNA isolation and separation of serum, plasma, and red cells. In order to make a positive diagnosis, at least two of the following conditions had to be present: (i) five or more café-au-lait spots greater than 0.5 cm in diameter, (ii) one or more biopsy-proven neurofibromas, (iii) multiple iris Lisch nodules, (iv) axillary or groin freckling, (v) the presence of distinctive NF manifestations—either optic glioma or tibial pseudoarthrosis, and (vi) presence of NF in a first degree relative. A total of 15 families (out of 17 examined) met the clinical criteria for von Recklinghausen NF and were informative for linkage.

Blood samples were obtained from 128 of 157 affected and at-risk family members and spouses. An example of the hybridization analyses used is shown in Fig. 1 (*7, 8*). For six kindreds, marker information was available from or inferrable for three or more generations of kindred members affected with NF. For two of the remaining nine families, both grandparents were unaffected;

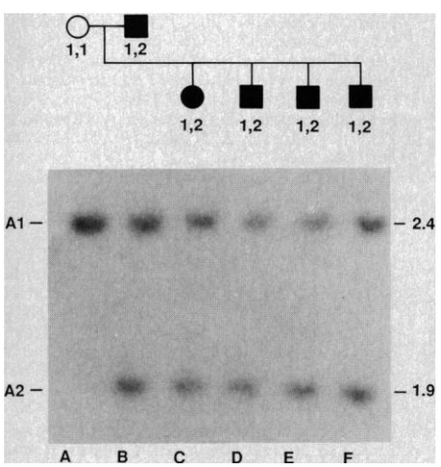

Fig. 1. Hybridization of probe pA10-41 to Msp I–digested DNAs from individuals in NF kindred 1655. The Msp I alleles of pA10-41 are a 2.4-kb (A1) or 1.9-kb (A2) band. The Pvu II alleles (not shown) are a 3.2-kb (B1) or 3.0-kb (B2) band. The two polymorphisms show apparent linkage equilibrium with each other, with an estimated combined heterozygosity of over 50%. The figure shows cosegregation of NF with the A2 allele of the probe in each of four affected offspring. DNAs were prepared from lymphocyte nuclei by standard methods and digested with a fivefold excess of Msp I as suggested by the suppliers except that reactions were incubated at room temperature for 12 hours. Agarose gel electrophoresis, transfers to nylon membranes, nick translations, hybridizations, and autoradiography were performed as described (*7, 8*).

Table 1. Pairwise analyses of NF and chromosome 17 markers. θ, best estimate of θ; $z(\hat{\theta})$, LOD score for that recombination fraction.

Probe	Recombination fraction (θ)						$\hat{\theta}$	$z(\hat{\theta})$
	0.0	0.05	0.10	0.20	0.30	0.40		
p3-6	−0.08	4.21	3.93	2.85	1.54	0.49	0.04	4.21
p10-5	−7.82	−0.97	−0.18	0.36	0.43	0.28	0.27	0.44
pHF12-2	−9.84	−3.76	−2.02	−0.69	−0.27	−0.12	0.50	0.00
pYNZ22	−25.02	−5.31	−2.56	−0.54	0.12	0.18	0.38	0.19
pTHH59	−13.78	−5.19	−3.17	−1.34	−0.56	−0.20	0.50	0.00

D. Barker, E. Wright, K. Nguyen, L. Cannon, P. Fain, D. Goldgar, D. T. Bishop, M. Skolnick, Department of Medical Informatics, University of Utah Medical Center, Salt Lake City, UT 84132.
J. Carey and B. Baty, Department of Pediatrics, University of Utah Medical Center, Salt Lake City, UT 84132.
J. Kivlin, Department of Ophthalmology, University of Utah Medical Center, Salt Lake City, UT 84132.
H. Willard, J. S. Waye, G. Greig, Department of Medical Genetics, University of Toronto, Toronto, Canada, M5S 1AB.
L. Leinwand, Departments of Microbiology and Immunology and of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461.
Y. Nakamura, P. O'Connell, M. Leppert, J.-M. Lalouel, R. White, Howard Hughes Medical Institute and Department of Human Genetics, University of Utah Medical Center, Salt Lake City, UT 84132.

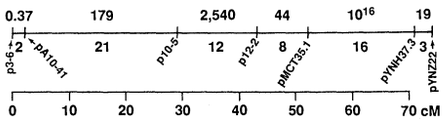

Fig. 2. Genetic map for the short arm and centromere region of chromosome 17. Numbers above each interval (top line) reflect the odds favoring the given order compared to the order obtained by reversing markers that bound the interval. Numbers below each interval give the estimated recombination fraction between the adjacent markers. Genetic distance from the centromere is shown on the bottom line.

however, their marker types contributed information for determining marker phase in multilocus analyses.

The LOD score (z) for a given value of the recombination fraction, θ_i, is the decimal logarithm of the odds in favor of the hypothesis of linkage at θ_i compared to the unlinked case, $\theta = 0.5$. A LOD score of 3.0, representing odds for linkage of 1000:1, is conventionally accepted as significant. For our initial analyses, equal recombination rates were assumed to occur in males and females ($\theta_m = \theta_f$). The gene frequency for NF was fixed at 0.0002 and a penetrance of 95% was assumed, as a few cases of transmission through unaffected individuals have been reported [for example, see (*5*)]. This genetic model is consistent with a population prevalence of about 1 per 2500.

Results of pairwise analyses between NF and five chromosome 17 markers for our current data set are given in Table 1. Analyses were performed with the LOD score method of Morton (*9*) as implemented in the LINKAGE programs (*10*). Characteristics of the DNA markers are provided in Table 2 (*11–15*). A maximum LOD score of 4.2 was observed for the centromeric marker, probe p3-6, at 4% recombination. There was no evidence of close linkage to any of the other chromosome 17 markers.

To identify candidate markers which could confirm linkage of NF to chromosome 17, a linkage map of markers near the centromere and on the short arm of chromosome 17 (17p) was constructed from genotypic analysis of 59 reference families. Forty-six three-generation families with large sibships were ascertained in Utah, including 27 families which have been submitted to the Centre d'Etude du Polymorphisme Humain (CEPH) reference panel (*16*). Thirteen additional families from the CEPH panel (*17*) included a mixture of two- and three-generation families, also with large sibships.

A database of genotypes for markers on chromosome 17 had been previously constructed on these families and linkage to chromosome 17 was detected for two previously unassigned markers, pA10-41 and pMCT35.1. Thus, seven marker loci on chromosome 17 were found to form a linkage group that included the centromere (Fig. 2).

Estimates of recombination fractions and likelihoods for different orders were determined by multilocus linkage analysis with the LINKAGE program (*10*). Results are reported in Fig. 2. While the order of most of the markers is clear from the analyses, the orientation of pA10-41 with respect to the centromere remains unresolved. The recombination fraction in females was significantly higher than in males ($P < 0.01$).

The map led us to expect close linkage between NF and a previously unassigned marker, probe pA10-41. Pairwise analysis of NF and pA10-41 for the complete data set confirmed our expectation. Detailed characteristics of the two significant NF linkages (p3-6 and pA10-41) are reported in Table 3 in accordance with recommended guidelines (*18*). Probe pA10-41 gave a maximum LOD score of 4.4 at 5% recombination ($\theta_m = \theta_f$). LOD scores for individual families are given in Table 4. All but one family were informative for at least one of the two linked markers. No evidence for heterogeneity among families was revealed by a statistical test for admixture ($P = 0.5$) (*19*).

Because the relative order of the centromere and pA10-41 was unresolved, different locations for the NF gene were compared for two map orders and the overall odds in favor of the NF gene being located on each map were calculated. Fig. 3 (A and B) present these results. The overall odds in favor of NF belonging to this linkage group on this map are over 80 million to 1. The order of the centromere and pA10-41 loci within the linkage map remains unresolved. Furthermore the NF gene has nearly equal probability of falling into two intervals: between the centromere and p10-5 (Fig. 3A) or distal to the centromere (Fig. 3B). When both the NF and reference families were used to refine the genetic map, the probability that NF is located between the

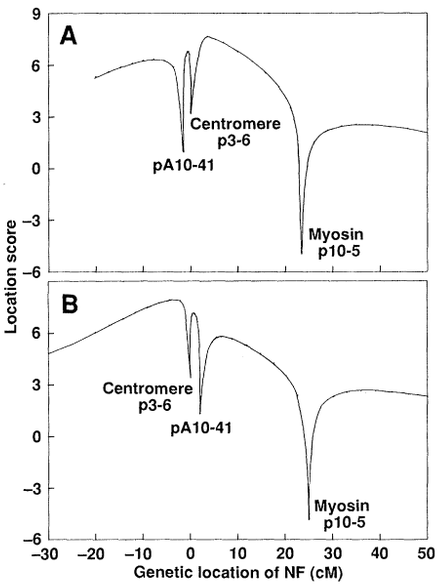

Fig. 3. Location scores for the NF locus relative to the best regional map of chromosome 17 (**A**) and the nearly equally likely alternative (**B**). The LINKMAP program of the LINKAGE package was used to determine the location of NF in relation to the two most likely maps of the region of interest, which included the p3-6, pA10-41, and p10-5 marker loci. The NF locus was moved along each fixed genetic map of marker loci, and the position of the NF gene within each map that maximized the likelihood of the observed family data was obtained.

Table 2. Probes used for mapping of NF. Probe pHF12-2 recognizes locus D17S1; probe p10-5 recognizes gene MYH2; probe p3-6 recognizes locus D17Z1.

Probe*	Enzyme	Heterozygosity
pYNZ22 (*11*)	Pst I	0.86
pYNH37-3 (*11*)	Taq I	0.63
pMCT35.1 (*11*)	Msp I	0.43
pHF12-2 (*12*)	Msp I	0.43
p10-5 (*13*)	Hind III	0.38
	Msp I	0.51
pA10-41 (*14*)	Pvu II	0.29
	Msp I	0.49
p3-6 (*15*)	Hind III	0.42
pTHH59 (*11*)	Taq I	0.71

*References for the probes described are shown in parentheses.

Table 3. Detailed results of pairwise analyses for two DNA probes that show significant linkage to NF. Global maxima were 4.7 and 5.0 for p3-6 and pA10-41, respectively. Descriptions of the theoretical bases for considering different recombination fractions (θ) for males (m) and females (f) have been described (*18*). Symbols: $\hat{\theta}$, best estimate of θ; $z(\hat{\theta})$, LOD score for that recombination fraction.

Probe	Model	Recombination fraction (θ)						$\hat{\theta}$	$z(\hat{\theta})$
		0.0	0.05	0.10	0.20	0.30	0.40		
p3-6	$\theta_m = \theta_f$	−0.08	4.21	3.93	2.85	1.54	0.49	0.04	4.21
	$\theta_m, \theta_f = 0.13$	4.72	4.38	3.94	2.92	1.84	1.03		
	$\theta_f, \theta_m = 0.00$	−0.08	4.55	4.70	4.64	4.41	4.13		
pA10-41	$\theta_m = \theta_f$	−0.45	4.37	4.14	3.03	1.71	0.56	0.05	4.37
	$\theta_m, \theta_f = 0.11$	4.99	4.57	4.14	3.24	2.32	1.48		
	$\theta_f, \theta_m = 0.00$	−0.45	4.78	4.99	4.75	4.23	3.70		

Table 4. Pairwise LOD scores by family between NF and the two closest markers at the most likely recombination fraction for the entire set of families. Ni, not informative.

Kindred	LOD	
	p3-6 ($\hat{\theta}$ = 0.04)	pA10-41 ($\hat{\theta}$ = 0.05)
1388	Ni	Ni
1391	0.26	0.75
1393	1.10	0.50
1611	Ni	0.24
1618	−0.38	0.60
1620	0.81	0.52
1621	0.25	Ni
1623	0.58	Ni
1624	Ni	0.24
1626	0.55	0.54
1629	0.28	Ni
1647	0.79	0.77
1650	Ni	−0.48
1655	Ni	0.81
1765	Ni	−0.12

centromere and pA10-41 increased (see Table 5).

Two other possible chromosomal locations for an NF gene have been previously suggested. Spence *et al.* (*20*) found some evidence indicating a locus linked to Gc on chromsome 4. Ichikawa *et al.* (*21*) described a kindred with cosegregation of NF and myotonic muscular dystrophy which is located on chromosome 19. Several studies attempted to map other families to chromosomes 4 or 19 with negative results (*5*).

The present study provides conclusive evidence that a locus for NF is located on chromosome 17 in the centromeric region. Although our study does not eliminate the possibility of more than one NF locus, it does provide evidence that a significant proportion of NF is caused by a locus on chromosome 17. The issue of heterogeneity can best be resolved by studying a large set of families with a battery of highly informative ordered markers in the centromeric region of chromosome 17. The availability of a chromosome 17 map constructed from the CEPH and Utah control families was central to the mapping of NF. Once the initial suggestion of linkage to the chromosome 17 centromere was obtained, a regional map was constructed. This map explained the negative LOD scores of the more distal probes. It also indicated that a new probe, pA10-41, was in the region of the NF gene.

The best two-locus LOD score in this study (NF to pA10-41) provided odds of approximately 23,000 to 1 in favor of linkage. By incorporating the chromosome 17 map in the analysis we were able to demonstrate linkage to the map with odds of approximately 80 million to 1. Tightly linked flanking markers are essential for both genetic counseling and for further molecular studies of NF. However, this study has not provided unambiguous flanking markers and there is still uncertainty whether the centromeric probe (p3-6) or pA10-41 is the nearest marker studied to date. Resolving marker order and finding closer flanking markers is a high priority for future studies.

A number of important diseases have been mapped to DNA markers in the last several years (*8, 22*), and several significant examples of gene mapping leading to further understanding of an underlying defect have been reported recently (*23*). Most notably Duchenne muscular dystrophy (DMD) has progressed from gene mapping to identification and cloning of the gene which causes the disease; the cloned gene has in turn been used to identify messenger RNA in fetal tissue (*24*). In a similar set of experiments, the gene for chronic granulomatous disease has been identified (*25*). Furthermore, evidence exists that the retinoblastoma gene may have been cloned (*26*). In all three cases deletions were used to narrow down the range of candidate sequences (*23*). If the process of tumor formation in NF-derived malignancies occurs by similar genetic mechanisms as for retinoblastoma, then the techniques used to isolate the retinoblastoma gene could be repeated to isolate the NF gene. DMD and neurofibromatosis have the highest mutation rates known in humans. In both cases, heterogeneity was a possible explanation until the linkage studies proved otherwise; this explanation is clearly an insufficient explanation for the mutation rate in NF given the consistency of our linkage results. In DMD, an unusually large gene size appears to be responsible for the high mutation rate. Therefore, a similarly large gene is a possible explanation for the high mutation rate in NF.

It is thought that recombination is reduced near the centromeres (*27*). Therefore, the approximately 4 centimorgans (cM) between NF and the centromere may represent a longer sequence of DNA than one would predict from the usual equation of 1 million bases per centimorgan. Our results still leave a region of at least 10 cM on either side of the centromere for the physical location of NF.

Table 5. Results of joint four-locus analyses of NF and reference families. LOD scores are for the alternate hypothesis that $\theta_{12} = \theta_{23} = \theta_{34} = 0.5$. Odds are relative to the best order.

Locus 1	$\hat{\theta}_{12}$	Locus 2	$\hat{\theta}_{23}$	Locus 3	$\hat{\theta}_{34}$	Locus 4	LOD	Odds
NF	2.8	p3-6	4.4	pA10-41	20.8	p10-5	27.2	—
p3-6	2.3	NF	3.7	pA10-41	20.8	p10-5	27.1	1:1.2
p3-6	4.8	pA10-41	4.2	NF	20.3	p10-5	25.3	1:72
NF	5.0	pA10-41	4.4	p3-6	20.5	p10-5	25.2	1:95
pA10-41	3.3	NF	2.1	p3-6	20.6	p10-5	26.3	1:7
pA10-41	3.8	p3-6	2.6	NF	20.1	p10-5	26.4	1:6

REFERENCES AND NOTES

1. F. W. Crowe *et al.*, *A Clinical, Pathological, and Genetic Study of Multiple Neurofibromatosis* (Thomas, Springfield, IL, 1956).
2. L. L. Cavalli-Sforza and W. F. Bodmer, *The Genetics of Human Populations* (Freeman, San Francisco, 1971).
3. V. M. Riccardi and J. E. Eichner, *Neurofibromatosis: Phenotype, Natural History, and Pathogenesis* (Johns Hopkins Univ. Press, Baltimore, MD, 1986).
4. J. L. Bader, *Ann. N.Y. Acad. Sci.* **486**, 57 (1986).
5. J. K. Darby *et al.*, *ibid.*, p. 311; B. R. Seizinger *et al.*, *ibid.*, p. 304; S. M. Huson *et al.*, *ibid.*, p. 349; J. N. Dietz *et al.*, *Gen. Epid.* **3**, 313 (1986); B. G. Dunn *et al.*, *Am. J. Med. Genet.* **22**, 403 (1985); S. M. Huson *et al.*, *ibid.* **23**, 55 (1986); M. A. Pericak-Vance, *Cytogenet. Cell Genet.* **40**, 721 (1985); M. L. Rivas and J. H. DiLiberti *ibid.* **37**, 570 (1984); M. A. Spence *et al.*, *J. Med. Genet.* **20**, 334 (1983).
6. A series of papers summarizing linkage studies with NF by a number of groups was presented at the European Symposium on NF. A resulting map of all the chromosomal locations where NF is likely to be found excluded will be available (*J. Med. Genet*, in press).
7. D. Barker, T. Holm, R. White, *Am. J. Hum. Genet.* **36**, 1159 (1984).
8. L. C. Tsui *et al.*, *Science* **230**, 1054 (1985).
9. N. E. Morton, *Am. J. Hum. Genet.* **7**, 277 (1955).
10. G. M. Lathrop, J. M. Lalouel, C. Julier, J. OH, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3443 (1984); G. M. Lathrop *et al.*, *Am. J. Hum. Genet.* **37**, 482 (1985).
11. Y. Nakamura *et al.*, *Science* **235**, 1616 (1987).
12. D. Barker, M. Schafer, R. White, *Cell* **36**, 131 (1984); S. L. Naylor, A. Y. Sakaguchi, D. Barker, R. White, T. B. Shows, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2447 (1984).
13. C. E. Schwartz *et al.*, *Cytogenet. Cell Genet.* **43**, 117 (1986).
14. D. Barker and R. White, *ibid.* **37**, 413 (1984); M. H. Skolnick *et al.*, *ibid.*, p. 210.
15. H. Willard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5611 (1986).
16. R. White *et al.*, *Nature (London)* **313**, 101 (1985).
17. J. Dausset, *La Presse Medicale* **18**, 1801 (1986).
18. P. M. Conneally *et al.*, *Cytogenet. Cell Genet.* **40**, 356 (1985).
19. C. A. B. Smith, *Ann. Hum. Genet.* **27**, 175 (1963).
20. M. A. Spence *et al.*, *J. Med. Genet.* **20**, 334 (1983).
21. K. Ichikawa *et al.*, *ibid.* **18**, 134 (1981).
22. S. T. Reeders *et al.*, *Nature (London)* **317**, 542 (1985).
23. S. H. Orkin, *Cell* **47**, 845 (1986).
24. A. P. Monaco *et al.*, *Nature (London)* **323**, 646 (1986).
25. B. Royer-Pokora *et al.*, *ibid.* **322**, 32 (1986).
26. T. P. Dryja, J. M. Rapaport, J. M. Joyce, R. A. Petersen, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7391 (1986); S. H. Friend *et al.*, *Nature (London)* **323**, 643 (1986); W.-H. Lee *et al.*, *Science* **235**, 1394 (1987).
27. R. White and J.-M. Lalouel, *Adv. Hum. Genet.* **16**, 121 (1987).

28. Supported by NIH grants CA 28854, GM 29090, and CA 36362, March of Dimes service grant C-310 and research grant 1-995, and the Medical Research Council of Canada. L.L. is an established investigator of the American Heart Association and R.W. and J.-M.L. are Investigators of the Howard Hughes Medical Institute.

3 April 1987; accepted 5 May 1987