J 7, 345 (1988).
5. S. J. Weiss, M. B. Lampert, S. T. Test, *Science* **222**, 625 (1983); P. J. Ossanna, S. T. Test, N. R. Matheson, S. Regiani, S. J. Weiss, *J. Clin Invest* 77, 1939 (1986); P. E Desrochers and S. J. Weiss, *ibid* **81**, 1646 (1988); V. Y. Reddy, S. V. Pizzo, S. J. Weiss, *J Biol Chem* **264**, 13801 (1989).
6. S. D. Wright *et al*, *Proc Natl Acad. Sci U.S.A* **80**, 5699 (1983); A. J. McMichael *et al*, Eds., *Leukocyte Typing III, White Cell Differentiation Antigens* (Oxford Univ. Press, Oxford, 1987).
7. W. G. Rice and S. J. Weiss, unpublished observation.
8. P. M. Henson *et al*, in *Inflammation Basic Principles and Clinical Correlates*, J. I. Gallin, I M. Goldstein, R. Synderman, Eds. (Raven, New York, 1988), pp. 363–390; M. C. M. Vissers and C. C. Winterbourn, *Biochim Biophys Acta* **889**, 277 (1986); T. Vartio, *J. Biol Chem* **264**, 4471 (1989).
9. S. Fairbairn, R. Gilbert, G. Ojakian, R. Schwimmer, J. P. Quigley, *J Cell Biol* **101**, 1790 (1985); J-M. Chen and W-T. Chen, *Cell* **48**, 193 (1987).
10. P. C. Amrein and T. P. Stossel, *Blood* **56**, 442 (1980); A. J. Jesaitis, J. O. Tolley, R. A. Allen, *J Biol Chem* **261**, 13662 (1986); K. A. Joiner and J. Schweinle, *J Immunol* **142**, 3164 (1989)
11. S. Malawista and G. Van Blaricom, *Proc. Natl. Acad. Sci U S A* **84**, 454 (1987).
12. K. L. Lomax *et al*, *J Clin Invest* **83**, 514 (1989).
13. T. Ganz, J. Metcalf, J. Gallin, L. Boxer, R. Lehrer, *ibid* **82**, 552 (1988). The amounts of elastase and cathepsin G in lysates prepared from the CHS neutrophils used in Table 1 were 13 and 9%, respectively, of normal controls as assayed according to Ganz *et al*
14. A. Segal, *ibid* **83**, 1785 (1989).
15. R. Kao, N. Wehner, K. Skubitz, B. Gray, J. Hoidal, *ibid* **82**, 1963 (1988).
16. J. Powers, G. Gupton, A Harley, N. Nishino, R. Whitley, *Biochim Biophys Acta* **485**, 166 (1977); J. B. Doherty *et al*., *Nature* **322**, 192 (1986).
17. Purified proteinase 3 (*15*) was inhibited by either Ala-Ala-Pro-Val-chloromethylketone or L-659,166, but was unaffected by rSLPI at the highest concentration tested (100 μg/ml) (B. Gray, S. Regiani, S. Weiss, unpublished observation).
18. The more commonly used cathepsin G inhibitor, *N*-benzyloxycarbonyl-Gly-Leu-Phe-chloromethylketone, is toxic to neutrophils [Y. Sibille, N. Merrill, J. Cooper, Jr., L. Polomski, J. Gee, *Am. Rev Resp Dis.* **130**, 110 (1984)]. Methoxysuccinyl-Val-Pro-Phe-trifluoromethylketone is a nontoxic cathepsin G inhibitor characterized and synthesized by Merrell Dow Research Institute.
19. Although synthetic elastase inhibitors act irreversibly, their association rate constants ($M^{-1}s^{-1}$) are two to three orders of magnitude slower than that of α_1PI or SLPI [M. Padrines and J. Bieth, *Am Rev Resp. Dis* **139**, 199A (1989); (*4*)].
20. K W. Miller, R. J. Evans, S. P. Eisenberg, R. C. Thompson, *J Bacteriol* **171**, 2166 (1989).
21. The ability of rSLPI to inhibit proteolysis was confirmed with insoluble elastin. When 1×10^6 neutrophils were stimulated atop ^{3}H-labeled insoluble elastin [specific activity of 10^3 cpm/μg prepared as described in H A. Chapman, O. L. Stone, Z. Varin, *J Clin Invest* **73**, 806 (1984)] in the presence of 20% heat-inactivated plasma, 7630 ± 641 and 18,483 ± 980 cpm were solubilized at 3 and 18 hours, respectively. In the presence of rSLPI (100 μg/ml), solubilization was reduced to 867 ± 6 cpm at 3 hours and 2193 ± 207 cpm at 18 hours (n = 3). If the remaining elastin was treated with hot alkali to expose more subtle proteolytic damage [P. J. Stone *et al*, *J. Clin. Invest* **82**, 1644 (1988)], no increase in elastin solubility was detected.
22. In the absence or presence of rSLPI (100 μg/ml), PMA-stimulated neutrophils generated 49.9 ± 3.7 and 56 1 ± 7.8 nmol of superoxide per 10^5 cells per hour, respectively [performed as described in S. J. Weiss, S. T Test, C. M. Eckmann, D. Roos, S. Regiani, *Science* **234**, 200 (1986)], released 1.07 ± 0.04 and 1.14 ± 0.18 μg of lactoferrin per 10^6 cells per hour, respectively [performed as described in E. Wilson *et al*, *Arch Biochem Biophys* **259**, 204 (1987)], and cleaved 0.30 ± 0 03 and 0.34 ± 0.03 nmol (per 10^6 cells per hour) of the β-glucuronidase substrate, phenolphthalein, respectively (*ibid*). Similarly, the ability of 10^6 neutrophils to migrate across a type I collagen–coated filter in response to zymosan-activated plasma [performed as described in A. Huber and S. Weiss, *J. Clin Invest.* **83**, 1122 (1989)] was unaffected by the addition of rSLPI (that is, 60.5 ± 1.2 and 60.1 ± 2.3% of the neutrophils migrated through the collagen barrier and accumulated in the lower compartment after an 8-hour incubation in the absence and presence of SLPI, respectively). All values are reported as the means ± SD of three experiments.
23. Although SLPI is sensitive to oxidative damage [J. A. Kramps, Ch. van Twisk, E. C. Klasen, J. H. Dijkman, *Clin Sci* **75**, 53 (1988)], the protective effect exerted by the inhibitor was not enhanced by the addition of antioxidants to the plasma system.
24 S. P. Eisenberg, K. K. Hale, P. Heimdal, R. C. Thompson, *J Biol Chem* **265**, 7976 (1990).
25. W. G. Rice, R. C. Thompson, S. J. Weiss, unpublished observation.
26. C. Wilde, J. Snable, J. Griffith, R. Scott, *J Biol Chem* **265**, 2038 (1989).
27. B. U. Wintroub, J. S. Coblyn, C. E. Kaempfer, K. F. Austen, *Proc Natl Acad Sci U S A* 77, 5448 (1980); S. Pontremoli *et al*, *ibid.* **83**, 1685 (1986); P M. Barling, P. M. Bowles, W. I. Colquhoun, P. Grandison, G. K Scott, *Biochim Biophys Acta* **931**, 1 (1987); E. G. Shephard *et al*, *J Immunol* **143**, 2974 (1989).
28. F. Grinnell, *J. Cell Biol.* **103**, 2697 (1986).
29. After submission of the manuscript, we found that neutrophils express greater proteolytic activity when suspended in Earle's balanced salt solution and incubated in a 5% CO_2/95% air atmosphere. Under these conditions, 1×10^6 PMA-stimulated neutrophils degraded 11,550 ± 1617 cpm in 20% heat-inactivated plasma and 12,260 ± 2028 cpm in 20% heat-inactivated serum after a 3-hour incubation (n = 5). In the presence of rSLPI (100 μg/ml), proteolysis was inhibited by ~80% (for example, 2395 ± 394 cpm and 2483 ± 477 cpm in heat-inactivated plasma and serum, respectively (n = 4).
30. Supported by NIH grants from the National Heart, Lung, and Blood Institute, HL 28024, and from the National Institute of Allergy and Infectious Diseases, AI 21301. We thank R. C. Thompson (Synergen, Inc.) for recombinant SLPI and tissue inhibitor of metalloproteinases, P. Davies (Merck Sharp & Dohme Research Laboratories) for the elastase inhibitor, N. Doherty (Merrell Dow Research Institute) for the cathepsin G inhibitor, R. F. Todd for IB4, L. Boxer for providing samples of CHS and SGD blood, and S. Regiani and T. Lockhart for technical and secretarial assistance, respectively.

22 December 1989; accepted 2 May 1990

Type 1 Neurofibromatosis Gene: Identification of a Large Transcript Disrupted in Three NF1 Patients

MARGARET R. WALLACE, DOUGLAS A. MARCHUK, LONE B. ANDERSEN, ROXANNE LETCHER, HANA M. ODEH, ANN M. SAULINO, JANE W. FOUNTAIN, ANNE BRERETON, JANE NICHOLSON, ANNA L. MITCHELL, BERNARD H. BROWNSTEIN, FRANCIS S. COLLINS

Von Recklinghausen neurofibromatosis (NF1) is a common autosomal dominant disorder characterized by abnormalities in multiple tissues derived from the neural crest. No reliable cellular phenotypic marker has been identified, which has hampered direct efforts to identify the gene. The chromosome location of the *NF1* gene has been previously mapped genetically to 17q11.2, and data from two NF1 patients with balanced translocations in this region have further narrowed the candidate interval. The use of chromosome jumping and yeast artificial chromosome technology has now led to the identification of a large (~13 kilobases) ubiquitously expressed transcript (denoted *NF1LT*) from this region that is definitely interrupted by one and most likely by both translocations. Previously identified candidate genes, which failed to show abnormalities in NF1 patients, are apparently located within introns of *NF1LT*, on the antisense strand. A new mutation patient with NF1 has been identified with a de novo 0.5-kilobase insertion in the *NF1LT* gene. These observations, together with the high spontaneous mutation rate of NF1 (which is consistent with a large locus), suggest that *NF1LT* represents the elusive *NF1* gene.

THE VARIABLE AND DIVERSE MANIFESTATIONS of neurofibromatosis (NF1) have puzzled clinicians and neurobiologists alike since the condition was first described by von Recklinghausen in 1882 (*1*). With an incidence of about 1 in 3000 in all ethnic groups (*2*), it is one of the most common autosomal dominant disorders of man.

Despite considerable efforts, it has not been possible to define a consistent abnormality in NF1 tissues that would provide sufficient information on the gene product to allow direct cloning of the gene. The remaining alternative has been to identify the *NF1* gene by positional cloning [formerly referred to as "reverse genetics" (*3, 4*), although this may be a misleading designator]. In this process, a gene is identified on

M. R. Wallace, D. A. Marchuk, L. B. Andersen, R. Letcher, H. M. Odeh, A. M. Saulino, J. W. Fountain, A. Brereton, J. Nicholson, A. L. Mitchell, F. S. Collins, Howard Hughes Medical Institute and the Departments of Internal Medicine and Human Genetics, University of Michigan, Ann Arbor, MI 48109.
B. H. Brownstein, Center for Genetics in Medicine, Washington University School of Medicine, St. Louis, MO 63110.

the basis of its (known) chromosomal map position rather than by its functional properties; this strategy, although laborious, has led to the identification of the genes for chronic granulomatous disease (*5*), Duchenne muscular dystrophy (*6*), retinoblastoma (*7*), a gene (*DCC*) involved in colon cancer progression (*8*), Wilms tumor (*9*), and the gene for cystic fibrosis (*10*). In all of these successes except the last, the identification of gross chromosomal rearrangements (either germline or somatic) involving the gene greatly aided the search process.

The first step in the identification of the *NF1* gene, genetic linkage analysis, led to the assignment of *NF1* to chromosome 17 in 1987 (*11*). A subsequent collaborative multipoint mapping effort narrowed its genetic location to about 3 centiMorgans of 17q11.2 (*11*). Given the substantial physical size (2 to 6 × 10^6 bp) of this interval, the identification of two patients with NF1 and apparently balanced translocations provided crucial landmarks for finer localization of *NF1*. In one patient, the translocation involves chromosomes 1 and 17 and in the other chromosomes 17 and 22; in both the breakpoint on chromosome 17 is in 17q11.2 (*12*). The tentative conclusion that the *NF1* gene itself is altered by the rearrangements was given additional credence when the translocation breakpoints were shown, by pulsed-field gel electrophoresis and the probes 17L1 and 1F10 (Fig. 1) (*13*), as well as cosmids (*14*) and yeast artificial chromosomes (YAC's) (*15, 16*) spanning this region, to be about 60 kb apart.

Subsequent cloning efforts have been focused on the region between the breakpoints. The first *NF1* candidate gene was identified in mice as a site of retroviral integration in murine leukemia (*17*); the human homolog (*EVI2*) mapped between the *NF1* breakpoints (Fig. 1) (*14*). An adjacent gene of similar genomic structure but unknown function was subsequently identified by walking and jumping experiments, and named *NF1-c2* (for *NF1* candidate gene 2) (Fig. 1) (*18*). Neither of these relatively small genes is interrupted by the translocations, nor have any abnormalities in these genes been identified in NF1 patients (*14, 18*), casting some doubt on their potential role as the cause of NF1. The discovery of more than one gene in this interval forces a reevaluation of simple hypotheses, and suggests that one or more of the following situations must apply: (i) these small genes are embedded within the true *NF1* gene; (ii) the translocations are creating the NF1 phenotype by a position effect rather than direct interruption, in which case *EVI2* or *NF1-c2* could still be involved; (iii) the NF1 phenotype is a complex result of the altered expression of more than one gene. We now present evidence that the first of these hypotheses is likely to be correct: a large, ubiquitously expressed and highly conserved gene in this region, *NF1LT*, is apparently interrupted by both translocations, contains the previous candidate genes within it, is altered in a new mutation NF1 patient, and thus most likely represents the authentic *NF1* gene.

Two different strategies were used to derive cDNA clones that define *NF1LT*. Initial experiments with the end-of-jump of clone EH1, obtained by chromosome jumping (Fig. 1), showed that a single-copy 1.4-kb Eco RI–Hind III subfragment, which lies just telomeric to the t(17;22) breakpoint, is conserved across species and is therefore a potentially useful probe in searching for transcripts in the region (*19*). This probe was used to screen a human peripheral nerve cDNA library (*20*) resulting in the isolation of clone P5, which has an insert of 1.7 kb (Fig. 2).

Transcripts were also sought with the YAC clone A113D7, part of an overlapping contig of clones from this region (*16*). As this YAC contains the entire breakpoint region, direct screening of cDNA libraries with this probe, although technically difficult, would be expected to yield an entire set of expressed transcripts. Clone B3A was isolated by means of this procedure (*21, 22*) from a B lymphoblast cDNA library (*23*), and it contained a 0.8-kb insert. Subsequent analysis revealed that P5 and B3A overlap (Fig. 2B).

To determine whether the *NF1LT* locus is interrupted by one or both translocation breakpoints, the 5′ end of P5 was used as a probe against a Southern (DNA) blot of the translocation hybrids (*12*). These hybrids contain chromosome 17 sequences telomeric to the breakpoints, with the t(1;17) break (hybrid DCR1) occurring 60 kb centromeric to the t(17;22) break (hybrid NF13). This probe detects two human Eco RI fragments of 15 and 4.0 kb (Fig. 3). Two bands of 8.0 and 2.5 kb are seen in mouse DNA, indicating that this transcript is strongly conserved. In the translocation hy-

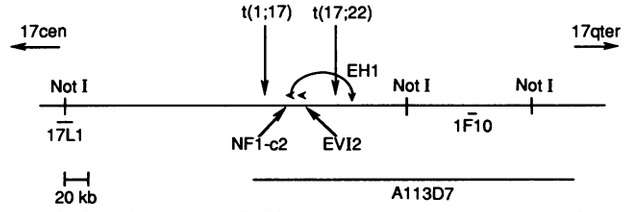

Fig. 1. Schematic of the *NF1* region. The orientation on chromosome 17 is shown and the translocation breakpoints are indicated by arrows. The schematic is drawn to scale (in kilobases). The two anonymous probes shown are 17L1 and 1F10 (*13*), while the previously described candidate genes *EVI2* (*14*) and *NF1-c2* (*18*) are shown above the map line, with the arrows indicating the direction of transcription. The jump clone EH1 (*19*) is indicated by the arc. The 270-kb YAC A113D7 is part of a contig covering this region (*16*) and was obtained from the Center for Genetics in Medicine at Washington University by screening their YAC library with a probe derived from cosmid 1F10.

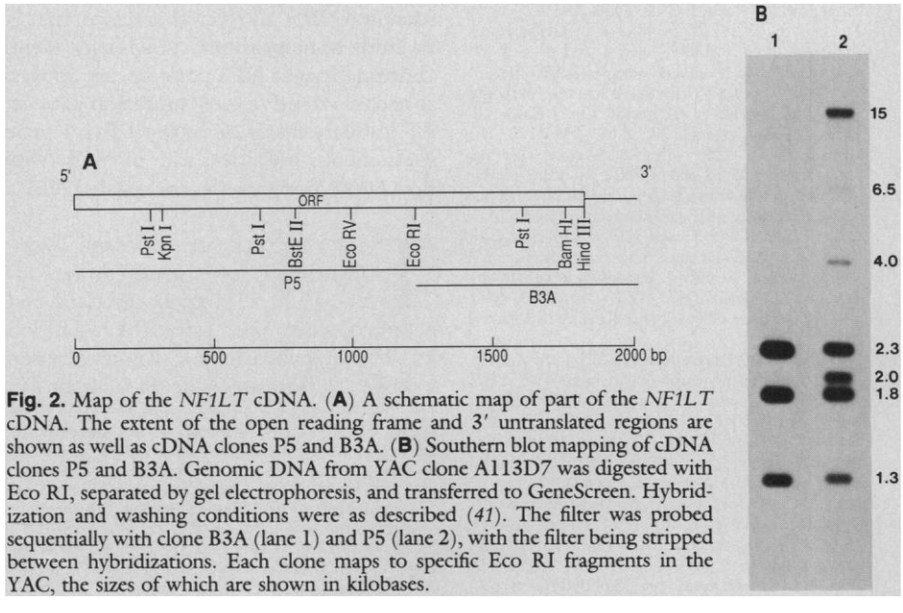

Fig. 2. Map of the *NF1LT* cDNA. (**A**) A schematic map of part of the *NF1LT* cDNA. The extent of the open reading frame and 3′ untranslated regions are shown as well as cDNA clones P5 and B3A. (**B**) Southern blot mapping of cDNA clones P5 and B3A. Genomic DNA from YAC clone A113D7 was digested with Eco RI, separated by gel electrophoresis, and transferred to GeneScreen. Hybridization and washing conditions were as described (*41*). The filter was probed sequentially with clone B3A (lane 1) and P5 (lane 2), with the filter being stripped between hybridizations. Each clone maps to specific Eco RI fragments in the YAC, the sizes of which are shown in kilobases.

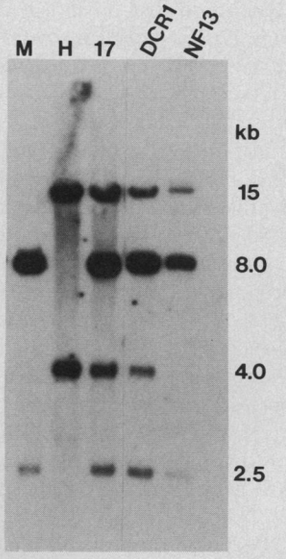

Fig. 3. *NF1LT* spans the t(17;22) breakpoint. For this Southern blot, DNA's were digested with Eco RI, transferred to Hybond N, and hybridized as previously described (*35*). The final washing was 1× SSC with 0.1% SDS at 65°C for 20 minutes. The probe was an 0.8-kb 5′ end fragment from the P5 Bluescript subclone, extending from the vector polylinker to the Bst EII site (Fig. 2). M, mouse DNA; H, normal human DNA; 17, DNA from hybrid MH22-6, a mouse cell line containing human chromosome 17 as its only human material (*42*); DCR1, the mouse hybrid containing the der(1) of t(1;17) (*12*); and NF13, the hybrid containing der(22) of t(17;22) (*12*).

brids, DCR1 contains both human bands, but NF13 lacks the 4.0-kb band, indicating that part of *NF1LT* lies between these breakpoints.

To determine the transcript size of *NF1LT*, cDNA clone P5 was used to probe Northern blots of RNA from various tissues. An approximately 13-kb transcript could be visualized in brain, neuroblastoma, kidney, and melanoma (Fig. 4, A and B). A hybridization signal of similar size was visible in RNA from several other tissues, although the bands were less discrete, probably due to degradation of the large transcript. Because of differences in degradation between RNA samples, it was not possible to judge from the band intensities the relative level of expression in different tissues.

In order to survey the pattern of expression of *NF1LT* in different normal and pathologic tissues we used the RNA polymerase chain reaction (PCR), with primers from the translated region, to analyze (*24*) a number of tissues. Expression of *NF1LT* was apparent in many human tissues, including those giving signals on Northern (RNA) blots, as well as immortalized B lymphoblasts (WBC) (both NF1 and non-NF1), NF1 skin fibroblasts, spleen, lung, muscle, thymoma, neuroblastoma, an NF1 neurofibrosarcoma cell line, a colon carcinoma cell line, and breast cancer (Fig. 4C). In other experiments expression was also detected in colon, thyroid, parathyroid adenoma, lymphoma, endometrial carcinoma, K562 erythroleukemia cells, and normal skin fibroblasts. Leukocyte contamination of the solid tissue samples could potentially account for the PCR signals, but since RNA's from various cell lines show expression, it is likely that *NF1LT* is widely expressed.

The primers used in the above experiments amplified a band of similar size in mouse RNA showing that this translated region is conserved. In order to determine whether the *NF1LT* gene is inactivated in three hybrid cell lines from NF1 patients with cytogenetic rearrangements, we included one PCR primer from the 3′ untranslated region, expecting that this area would be less conserved across species. The three hybrid cell lines included in this experiment were DCR1 and NF13, described above, and del(17). This last hybrid contains the deleted chromosome 17 from an NF1 patient with a deletion of part of the proximal long arm of chromosome 17 (*25*). The expected product is seen in B lymphoblasts (WBC) and human brain, but not in mouse RNA (Fig. 4D). A hybrid containing the normal human chromosome 17 does express human *NF1LT*, but no product is visible with either NF1 translocation hybrid or with the del(17) hybrid, indicating that these rearrangements abolish expression.

Identifying mutations in NF1 patients is crucial to identifying the gene from among candidate genes. New mutation NF1 patients are helpful, since comparison of their DNA with that of their parents allows the

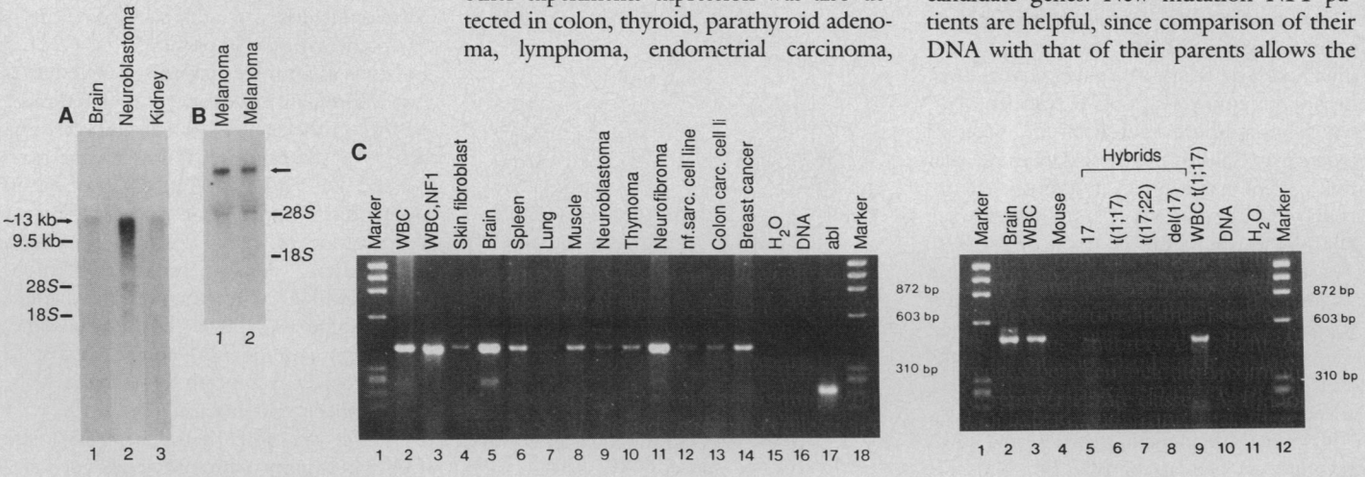

Fig. 4. Analysis of *NF1LT* expression. Northern blot results are shown for probe P5 in the following human tissues: (**A**) lane 1, brain frontal lobe; lane 2, neuroblastoma; lane 3, kidney; and (**B**) lanes 1 and 2, two independent melanoma cell lines. RNA samples were prepared according to standard methods (*43*). Approximately 15 μg of total RNA from each tissue was separated on 1% formaldehyde gels and transferred onto Hybond N. The P5 insert was labeled by the random priming method (*22*). Exposures were 24 to 96 hours. The position of the 28*S* and 18*S* ribosomal RNA bands as well as the 9.5-kb RNA ladder band (BRL) are indicated. The arrows point to the ~13 kb transcript, the size of which was estimated by extrapolation from the 0.24- to 9.49-kb RNA ladder (BRL). Some cross-hybridization with the ribosomal RNA bands (28*S* and 18*S*) is apparent. (**C**) PCR analysis of RNA expression. An ethidium bromide–stained gel is shown, demonstrating the presence of a 401-bp amplified fragment in reverse transcribed RNA from a variety of tissues as indicated. A control RNA-PCR amplification with neurofibroma RNA and primers from the *abl* locus (*44*) gives the expected product. Primers A and B from the coding region of *NF1LT* (Fig. 6) were used. Total RNA (0.2 μg) from each tissue was reverse-transcribed with oligo(dT) (*45*) and then amplified by 35 cycles of PCR with 1 minute at 94°C, 1 minute at 60°C, and 2 minutes at 72°C. The absence of a band in the human DNA control indicates that the primers lie in different distant exons, and that signals derived from RNA-PCR therefore reflect mRNA expression. (**D**) PCR results from RNA from mouse-human hybrid and parental cell lines using primers which do not amplify mouse *NF1LT* RNA. Primers C and D (Fig. 6) were used under the same amplification conditions as mentioned above. An amplified fragment of 500 bp is visible in lanes 2, 3, 5, and 9. Southern blots of this gel confirmed the presence of a PCR product in lane 5 and its absence in lanes 6 to 8. In a control experiment, amplification of these hybrid cell line RNA samples with the conserved *NF1LT* primers [described in (C)] produced the expected mouse product, indicating that the RNA was intact.

distinction between causative mutation and polymorphism to be more readily made. Since pulsed-field gel studies by several groups with probes in the *NF1* region have failed to reveal large rearrangements (*13*), we used Southern blots to examine patient DNA with *NF1LT*. Accordingly, DNA from 35 individuals with NF1 were analyzed with the probe P5, and the autoradiograms were examined for abnormal bands or obvious differences in band intensity.

A single patient showed a difference with this assay. This NF1 new mutation patient is a 31-year-old white male who exhibits no café-au-lait spots or axillary freckling, but has macrocephaly, Lisch nodules, and multiple cervical nerve root tumors (shown histologically to be neurofibromas) requiring surgical debulking. His parents have been examined and display no features of NF1 or NF2. With Eco RI (Fig. 5A), the patient has a normal Southern blot pattern, except that his 4.0-kb allele is fainter than expected and he also demonstrates an abnormal fragment of 4.5 kb, with approximately the same intensity as the 4.0-kb band. This 4.5-kb band is not present in the parents. Similarly, with three other enzymes, an abnormal fragment approximaely 0.5 kb larger than expected is seen. For Pst I, the involved 12-kb fragment is the same one that contains the t(17;22) breakpoint (*19*). These abnormal bands have not been seen in other NF1 or unaffected individuals so far examined. As a confirmation, the family members were sampled again, and the same results were obtained. The family was studied with three highly polymorphic VNTR probes (*26*), which showed no indication of incorrect paternity. Collectively, these data imply that this patient has a novel mutation that appears to be an insertion of approximately 0.5 kb close to or within an exon of *NF1LT*. This new mutation, along with the evidence showing that the t(17;22) breakpoint interrupts the gene, and the PCR data showing that *NF1LT* expression is absent in the t(1;17) hybrid DCR1, suggests that *NF1LT* is the *NF1* gene.

Complete sequencing of P5 and B3A revealed an overlap of 507 bp, with the combined sequence being 2012 bp (Fig. 6). A single open reading frame extending from the beginning of P5 across nearly the entire sequence shows that B3A is located at the 3′ end and that transcription occurs toward the telomere. At the 3′ end of B3A, a stop codon occurs 181 bp from the end. However, no polyadenylation signal or poly(A) tail is evident, which implies that part of the 3′ untranslated region is missing. Comparisons of this DNA and protein sequence with the entries in Genbank (*27*) or the NBRF and SWISS-PROT databases (*28*) did not show significant similarity with any known sequence. A hydropathy plot (*29*) of the amino acid sequence revealed a primarily hydrophilic polypeptide. Other analyses failed to reveal any other recognizable motifs except for two potential N-glycosylation sites and three possible nuclear localization signals (Fig. 6). Thus the current sequence holds few clues as to the function of this gene. However, as the Northern blots indicate, much more of this transcript remains to be cloned and sequenced. Perhaps then, structural domains or similarity with other genes will become evident.

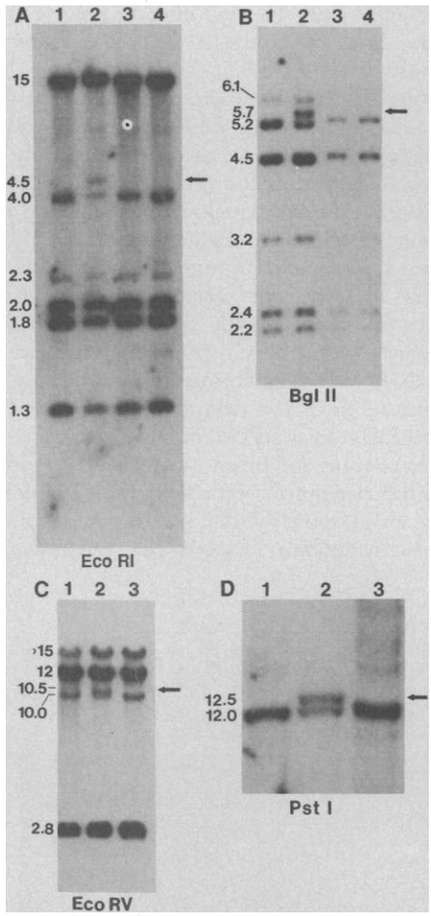

Fig. 5. Southern blots showing a 0.5-kb insertion in DNA of a new mutation NF1 patient. DNA samples from the patient and parents and a normal individual were digested with the indicated enzymes, transferred to Hybond, and hybridized with P5. In all panels, lane 1 contains DNA from the father; lane 2 is patient DNA; lane 3 is the mother's DNA; and lane 4 (when present) is DNA from an unaffected individual. Sizes are shown to the left of the bands, in kilobases, and the aberrant fragments are indicated by arrows on the right. In (D), where the gel has been run longer than usual to resolve large fragments, only the top bands are shown; P5 also hybridizes to numerous smaller Pst I fragments not involved in the rearrangement in this patient. No abnormal fragments were seen in this patient with Hind III, Xba I, Pvu II, Bcl I, Msp I, Nco I, and Nsi I.

Our conclusion that *NF1LT* is the gene involved in NF1 is based on several lines of evidence. First, this gene is disrupted by the t(17;22) breakpoint. As shown by RNA analysis of the DCR1 hybrid, the *NF1LT* gene is functionally disrupted by the t(1;17) NF1 translocation as well. More compelling is the identification of a 0.5-kb insertion in a new mutation NF1 patient. This insertion is located at least 10 kb away from the previously proposed candidate genes *EVI2* and *NF1-c2* (*14, 18*). Also in support of *NF1LT* as the *NF1* gene is its large transcript size, which is consistent with the high mutation rate [approximately 10^{-4} per allele per generation (*2*)] seen in this disease. This rate is similar to that for Duchenne muscular dystrophy (*6*), which has a transcript of similar size.

The 2.0-kb cloned portion of the *NF1LT* cDNA contains at least six exons (Fig. 7). The 5′ exon of the cloned transcript lies between the NF1 translocation breakpoints, within 12 kb of t(17;22). The size and number of genomic fragments detected by P5 and B3A on Southern blots indicates that the 2.0 kb of cloned cDNA span at least 33 kb of genomic DNA. Thus it is unlikely that the remaining 5′ *NF1LT* exons could lie entirely between the translocation breakpoints.

From the RNA-PCR expression analysis of hybrids (Fig. 4D), both NF1 translocations inactivate *NF1LT*, and the 0.5-kb insertion is likely to do so also. This is consistent with the hypothesis that *NF1LT* functions as a tumor suppressor gene rather than a dominant oncogene, and would place *NF1LT* in the same class as retinoblastoma (*RB*) (*7*), the putative Wilms tumor gene (*9*), and the *P53* gene (*30*). This inactivation mechanism is consistent with the recent identification of an NF1 patient carrying a constitutional deleted proximal 17q in which the deleted fragment forms a minichromosome (*25*). This minichromosome, which contains the *NF1* region, is lost in approximately 5 percent of somatic cells, leaving those cells hemizygous at the *NF1* locus. On the other hand, heterozygosity studies in tumors, which have provided significant evidence for the tumor suppressor mechanism in retinoblastoma (*7*) and Wilms tumor (*9*), have not supported this model in NF1 (*31*), potentially for three reasons: (i) The benign neurofibromas in NF1 are made up of a mixture of cell types and are not clonal (*32*), rendering analysis of loss of heterozygosity uninterpretable in these tumors; (ii) the malignant tumors often do show reduction to homozygosity for chromosome 17 (*31*), but this may reflect events at the *P53* locus (*30*) on 17p; and (iii) given the large size of *NF1LT*, it is possible that

the somatic event that disrupts the normal allele in the process of tumor development is commonly an independent mutation; this would not be reflected by a loss of heterozygosity for flanking markers.

The apparent ubiquitous expression of *NF1LT* mRNA, determined by sensitive RNA-PCR techniques, is in contrast to the predilection for neural crest–derived tissues to be predominantly involved in the disease. This same phenomenon has been observed for *RB* (*7*), which is expressed in a wide variety of tissues, though tumors are primarily restricted to retinal cells and bone. Presumably the site specificity of the disease phenotype reflects interactions between the *NF1LT* gene product and other cellular proteins, such that homozygous loss of *NF1LT* function is more likely to result in tumor development in neural crest tissues than in other environments. This widespread expression of *NF1LT* potentially explains the modest but significant increase in risk of other tumors, such as leukemia, rhabdomyosarcoma, and common solid tumors (*1, 33*), in NF1.

We have not yet localized the 5′ end of *NF1LT*, and it is conceivable that *NF1LT* could extend for hundreds of kilobases in the 5′ direction. Based on the observation that CpG islands often lie at the 5′ ends of genes (especially housekeeping genes) (*34*), one could look to the next 5′ CpG island as a potential site for the promoter of *NF1LT*. This island has been previously cloned in the Not I linking clone 17L1 (*35*) (Fig. 1). It lies approximately 150 kb centromeric to the t(1;17) breakpoint and shows strong conservation across species.

It is reasonable to conclude that *NF1LT* represents the *NF1* gene and that the phenotype of the disease results solely from altered expression of this gene product. However, it is still possible that the two previously identified genes *EVI2* and *NF1-c2*, or any other genes embedded in the region spanned by *NF1LT*, could play a role in NF1. There are very few previous examples of functional genes embedded within introns in higher eukaryotes. Genes transcribed from opposite strands of the same region in mammals (*36*) overlap at the 5′ or 3′ ends and occasionally share some exon sequence. These pairs of genes generally seem unrelated, although they may share regulatory mechanisms. A gene of unknown function has been described which is embedded in intron 22 of the human Factor VIII gene, on the antisense strand (*37*). Originally found by its association with a CpG island, this embedded gene has its entire 1.8-kb transcribed sequence contained in one exon, but its expression does not seem to be correlated with that of Factor VIII. The clearest previous examples of embedded genes occur in *Drosophila* (*38, 39*). These examples raise the possibility that antisense

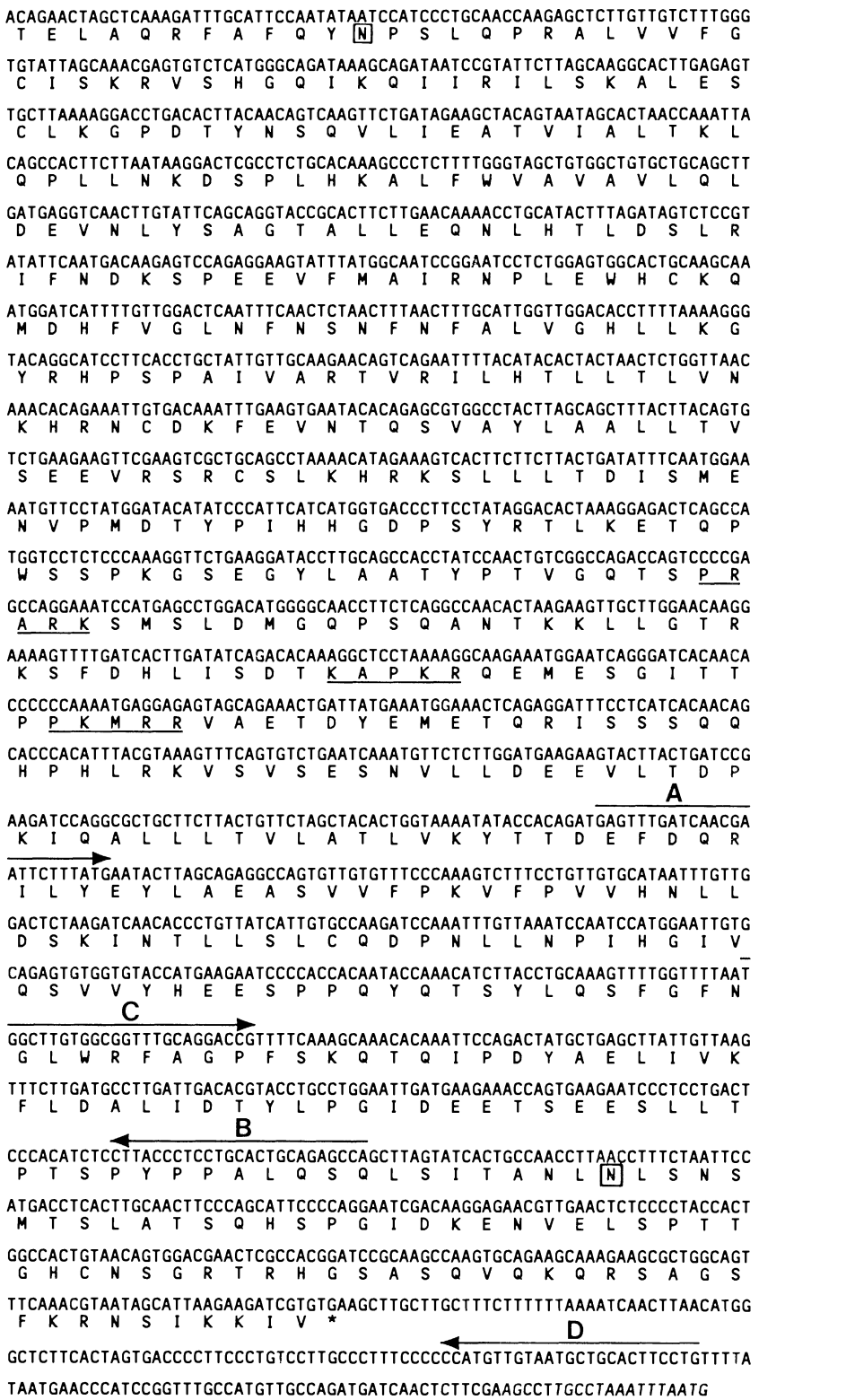

Fig. 6. Partial nucleotide sequence of *NF1LT* cDNA with deduced amino acid sequence. DNA sequence was obtained by double-stranded plasmid sequencing of P5 and B3A in Bluescript using the Sanger dideoxy method with ^{35}S as outlined in the Sequenase kit (U.S. Biochemicals, version 2.0). Sequencing was done with T3 and T7 primers as well as internal primers designed from sequence previously obtained. The two sets of primers (A, B and C, D) used in the PCR experiments (Fig. 4) are indicated. Potential N-glycosylation sites are boxed, and three possible nuclear localization signals (*46*) are underlined.

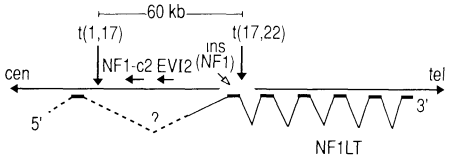

Fig. 7. Schematic exon map of *NF1LT*. The *NF1* region is represented, showing the locations of translocation breakpoints and genes *EVI2* and *NF1-c2*. Exons of *NF1LT* (not to scale) are represented by the black boxes below the solid line; the lines connecting these indicate splicing events. Since P5 hybridizes to seven genomic Eco RI fragments (Fig. 2) and contains a single Eco RI site, it includes at least six exons. The dashed lines and question mark indicate that the precise location of further 5′ exon sequences is unknown, but it is likely that at least some exons are centromeric to the t(1;17) breakpoint. The 0.5-kb insertion in the new mutation NF1 patient is located within the hatched region.

RNA may regulate the expression of complex loci, a phenomenon that is well documented in prokaryotes (*40*). In the case of *NF1LT*, one can envision steric hindrance in simultaneous transcription of *NF1LT* and the other two genes, or alternatively, double-stranded RNA could form by *EVI2* or *NF1-c2* RNA binding to the unprocessed *NF1LT* transcript, resulting in degradation of the duplex. Thus homozygous loss of *NF1LT* expression might lead to increased expression of *EVI2* and *NF1-c2*, and these gene products might ultimately be responsible for some of the NF1 phenotypic features. A simpler mechanism whereby *NF1LT* alterations alone lead to the NF1 phenotype is, however, probably more likely (*13, 17, 18*).

REFERENCES AND NOTES

1. T. Crump, *Adv Neurol* **29**, 259 (1981); V. M. Riccardi, *N Engl J Med* **305**, 1617 (1981); V. M. Riccardi and J. E. Eichner, *Neurofibromatosis Phenotype, Natural History, and Pathogenesis* (Johns Hopkins Univ. Press, Baltimore, MD, 1986), p. 305.
2. F. W. Crowe, W. T. Schull, J. F. Neel, *A Clinical, Pathological, and Genetic Study of Multiple Neurofibromatosis* (Thomas, Springfield, IL, 1956).
3. S. H. Orkin, *Cell* **47**, 845 (1986).
4. F. Ruddle, *Am J Hum Genet* **36**, 944 (1984).
5. B. Royer-Pokora *et al*., *Nature* **322**, 32 (1986)
6. A. P. Monaco *et al*., *ibid* **323**, 646 (1986); M. Koenig *et al*., *Cell* **50**, 509 (1987)
7. S. H. Friend *et al*., *Nature* **323**, 643 (1986); R. A. Weinberg, *Blood* **74**, 529 (1989).
8. E. R. Fearon *et al*., *Science* **247**, 49 (1990).
9. K. M. Call *et al*., *Cell* **60**, 509 (1990); M. Gessler *et al*., *Nature* **343**, 774 (1990).
10. J M Rommens *et al*., *Science* **245**, 1059 (1989); J. R. Riordan *et al*., *ibid*., p. 1066, B. S Kerem *et al*., *ibid*., p 1073.
11. D. Barker *et al*., *ibid* **236**, 1100 (1987); B. R. Seizinger *et al*., *Cell* **49**, 589 (1987); D E. Goldgar, P Green, D M. Parry, J J. Mulvihill, *Am J Hum Genet* **44**, 6 (1989)
12. M A. Schmidt, V V. Michels, W. Dewald, *Am J Med Genet* **28**, 771 (1987); D. H Ledbetter, D. C. Rich, P O'Connell, M Leppert, J. C. Carey, *ibid* **44**, 20 (1989), A. G. Menon *et al*., *Genomics* **5**, 245 (1989), F. S Collins, P O'Connell, B. A J Ponder, B R Seizinger, *Trends Genet* **5**, 217 (1989).
13. J. W Fountain *et al*., *Science* **244**, 1085 (1989); P. O'Connell *et al*., *ibid*., p 1087
14. P. O'Connell *et al*., *Genomics*, in press
15. D T Burke, G F Carle, M V. Olson, *Science* **236**, 806 (1987); E D Green and M. V Olson, *Proc Natl Acad Sci U S A* **87**, 1213 (1990)
16. D Marchuk *et al*., in preparation.
17. A. M Buchberg *et al*., *Oncogene Res* **2**, 149 (1988)
18. R Cawthon *et al*., in preparation
19. M. R. Wallace *et al*., in preparation
20. M Scott and K. Fischbeck (University of Pennsylvania Medical Center) constructed this library from human cauda equina RNA. This library is partially oligo(dT)-primed and partially random-primed, with the inserts cloned into the Eco RI site of λZAP (Stratagene) Clones (7×10^5) were plated on XL1-Blue cells and screened by usual methods (*43*).
21. Field-inversion gel electrophoresis of YAC genomic DNA was performed in a 1 0 percent low-melt agarose gel under conditions that separated the 270-kb YAC from the yeast chromosomes (160 V, 65 hours, 4°C; forward ramp, 6 to 48 seconds; reverse ramp, 2 to 16 seconds). The YAC was cut out of the gel, equilibrated in digestion buffer, and digested with Hinc II. After reequilibration in TE, the agarose was diluted in three volumes of water and melted at 68°C Labeling of the YAC was done in the diluted low-melt agarose according to (*22*), except that 0.5 mCi was used in a final volume of 500 μl. After removing unincorporated counts using a spin column, probe was preannealed with human placental DNA at a final concentration 1 mg/ml in 0.1 M NaCl for 15 minutes at 65°C Phage lifts on nitrocellulose filters were prehybridized overnight in 6× SSC, 2× Denhardt's, 1 mM EDTA, 0.5% SDS Hybridization was in the same solution for 48 hours Filters were washed to a final stringency of 0.2× SSC, 0 1% SDS at 65°C (D A Marchuk *et al*., in preparation)
22. A. P. Feinberg and B Vogelstein, *Anal Biochem* **137**, 266 (1984)
23. D. T. Bonthron, A F Markham, D Ginsburg, S H. Orkin, *J Clin Invest* 76, 894 (1985)
24. R. A Gibbs, P N Nguyen, L J. McBride, S M Koepf, C. T Caskey, *Proc Natl Acad Sci U S A* **86**, 1919 (1989)
25. L. B. Andersen, N Tommerup, J Koch, *Cytogenet Cell Genet*., in press; L B Andersen *et al*., in preparation
26. The three VNTR probes were pYNZ22, PYNH24, and pEKMDA2 [Y. Nakamura *et al*., *Science* **235**, 1616 (1987)]
27. C. Burks *et al*., *Methods Enzymol* **183**, 3 (1990).
28. W C. Barker, D. G George, L. T. Hunt, *ibid*., p. 31, P Kahn and G. Cameron, *ibid*., p. 23
29. J. Kyte and R F Doolittle, *J Mol Biol* **157**, 105 (1982)
30. S. J. Baker *et al*., *Science* **244**, 217 (1989); J M. Nigro *et al*., *Nature* **342**, 705 (1989); C. A. Finlay, P. W. Hinds, A J Levine, *Cell* **30**, 1083 (1989).
31. B. R. Seizinger *et al*., *Proc Natl Acad Sci U.S.A*., in press; G R. Skuse, B. A Kosciolek, P. T. Rowley, *Genes Chromosomes Cancer* **1**, 36 (1989); T Glover *et al*., in preparation
32. P. J. Fialkow, R W. Sagebriel, S. M. Gartler, D. L. Rimoin, *N. Engl J Med* **284**, 298 (1971).
33. J. L. Bader and R. W Miller, *J. Pediatr* **92**, 925 (1978); E. A. McKeen, J. Bodurtha, A. T. Meadows, E. C Douglass, J. J. Mulvihill, *ibid* **93**, 992 (1978); S A. Sørensen, J. J. Mulvihill, A. Nielsen, *N Engl J Med* **314**, 1010 (1986).
34. A P. Bird, *Nature* **321**, 209 (1986); S. Lindsay and A P Bird, *ibid*. **327**, 336 (1987).
35. M R. Wallace, J. W. Fountain, A. M. Brereton, F. S. Collins, *Nucleic Acids Res* **17**, 1665 (1989).
36. J P. Adelman, C. T. Bond, J. Douglass, E. Herbert, *Science* **235**, 1514 (1987); M. A. Lazar, R. A. Hodin, D. S. Darling, W. W Chin, *Mol Cell Biol* **9**, 1128 (1989), N. Miyajima *et al*., *Cell* **57**, 31 (1989).
37. B. Levinson, S. Kenwrick, D. Lakich, G. Hammonds, Jr., J. Gitschier, *Genomics* 7, 1 (1990).
38. S. Henikoff, M A. Keene, K. Fechtel, J. W. Fristrom, *Cell* **44**, 33 (1986).
39. C.-n. Chen, T Malone, S. K. Beckendorf, R. L. Davis, *Nature* **329**, 721 (1987).
40. M Inouye, *Gene* 72, 25 (1988); R. W. Simons, *ibid*., p. 35.
41. M L. Drumm *et al*., *Genomics* **2**, 346 (1988).
42. P. van Tuinen, D. C. Rich, K. M. Summers, D. H. Ledbetter, *ibid* **1**, 374 (1987).
43. J Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).
44. M. S. Roth, J H Antin, E. L Bingham, D. Ginsburg, *Blood* **74**, 882 (1989).
45. D Ginsburg *et al*., *Proc Natl Acad Sci U S A*. **86**, 3723 (1989)
46. C V. Dang and W M. F. Lee, *J Biol. Chem* **264**, 18019 (1989)

47. M. R. Wallace, D A. Marchuk, and L. B. Andersen contributed equally to this work. We thank the families with NF1 who participated in this study; the National Neurofibromatosis Foundation for their encouragement and a Young Investigator Award (M R W); M. Scott and K. Fischbeck for the human peripheral nerve cDNA library; D. Ginsburg for the B lymphoblast cDNA library; M. Powell and J Trent for melanoma RNA blots; D. Ledbetter for the NF1 translocation hybrids; T. Glover for the neurofibrosarcoma cell line; D. Freije for suggesting the idea for YAC screening of cDNA libraries; D Schlessinger and M. Olson for facilitating the screening of the Washington University YAC library to derive A113D7; M. Gellard of the University of Michigan Cancer Center Tissue Procurement Facility for crucial human tissue samples; S. Hayes, K Murphy, L Biesecker, M. Bruce, H. B. Clarke, S Norton, and D Osborn for technical assistance, R White and members of his research group for sharing information and probes; D. Ginsburg, T. Glover, and R. Boyce for helpful discussions; and B. Sandri for assistance in preparing the manuscript. Supported by National Institutes of Health grant NS23410 (F.S C.); the University of Copenhagen, Denmark (L.B.A.); and the Howard Hughes Medical Institute (F S.C. and D.A.M.)

12 June 1990; accepted 27 June 1990