GENOMICS 11, 931–940 (1991)

cDNA Cloning of the Type 1 Neurofibromatosis Gene: Complete Sequence of the *NF1* Gene Product

DOUGLAS A. MARCHUK, ANN M. SAULINO, ROXANNE TAVAKKOL, MANJU SWAROOP, MARGARET R. WALLACE, LONE B. ANDERSEN, ANNA L. MITCHELL, DAVID H. GUTMANN, MARK BOGUSKI,* AND FRANCIS S. COLLINS

*Departments of Human Genetics and Internal Medicine and the Howard Hughes Medical Institute, The University of Michigan Medical Center, Ann Arbor, Michigan 48109; and *National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894*

Received June 26, 1991; revised July 26, 1991

Von Recklinghausen neurofibromatosis, or type 1 neurofibromatosis (NF1), is a common autosomal dominant disorder characterized by abnormalities in multiple tissues derived from the embryonic neural crest. Portions of the gene have been recently identified by positional cloning, and sequence analysis has shown homology to the GTPase activating protein (GAP) family. In this report we present the results of an extensive cDNA walk resulting in the cloning of the complete coding region of the *NF1* transcript. Analysis of the sequences reveals an open reading frame of 2818 amino acids, although alternatively spliced products may code for different protein isoforms. The gene extends for approximately 300 kb on chromosome 17, with its promoter in a CpG-rich island. © 1991 Academic Press, Inc.

INTRODUCTION

Von Recklinghausen neurofibromatosis (NF1) is one of the most common inherited disorders in humans, with an incidence of about 1 in 3000, and affecting all ethnic groups (Crowe *et al.*, 1956; Riccardi, 1981; Riccardi and Eichner, 1986). The spontaneous mutation rate is very high, with 30–50% of cases being new mutations. This leads to a calculated mutation rate of 1/10,000, which is about 100-fold higher than the usual mutation rate for a single locus.

The clinical features of the disorder can be quite variable, even among affected family members, indicating that factors other than the specific mutation inherited must play a role in the disease phenotype. Diagnostic criteria for NF1 have been established (Stumpf *et al.*, 1987) and with careful examination it is rare to identify an obligate carrier who does not meet these criteria (Riccardi and Lewis, 1988). Diagnostic features include café-au-lait spots, Lisch nodules, and neurofibromas in the majority of gene carriers. Some individuals may also have learning disorders, more deeply placed plexiform neurofibromas, seizures, or skeletal abnormalities. The risk of malignancy is also increased, especially for optic glioma, neurofibrosarcoma, and brain tumors.

Recently, a gene was identified by positional cloning showing mutations in individuals affected with NF1 (Cawthon *et al.*, 1990a; Viskochil *et al.*, 1990; Wallace *et al.*, 1990). Further cloning and partial sequence analysis demonstrated that the gene product contains a domain showing approximately 30% similarity to the catalytic domains of yeast IRA1 and IRA2 proteins and the mammalian GTPase activating protein (GAP) (Xu *et al.*, 1990a). GAP is a cytosolic protein that catalyzes the conversion of active GTP-bound *ras* p21 to the inactive GDP-bound form (Trahey and McCormick, 1987; Trahey *et al.*, 1988; Vogel *et al.*, 1988). The IRA genes of *Saccharomyces cerevisiae* encode negative regulators of the yeast RAS genes that are homologs of mammalian GAP (Tanaka *et al.*, 1989, 1990, 1991). It was subsequently shown that the GAP-related domain of the *NF1* gene product can also interact with human and yeast *ras* p21 to down-regulate its activity (Ballester *et al.*, 1990; Martin *et al.*, 1990; Xu *et al.*, 1990b).

Previous reports of cDNA cloning of *NF1* were based on incomplete fragments of the transcript, which is approximately 13 kb by Northern blotting (Wallace *et al.*, 1990). In this report we present the complete amino acid sequence of the *NF1* gene product as deduced from cDNA clones containing the entire coding region of the gene. The location and extent of the *NF1* gene on the genomic map of chromosome 17 is also determined.

MATERIALS AND METHODS

Isolation of NF1 cDNA Clones

Five different cDNA libraries were used in the cDNA walk. A human fetal brain cDNA library,

0888-7543/91 $3.00
Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.

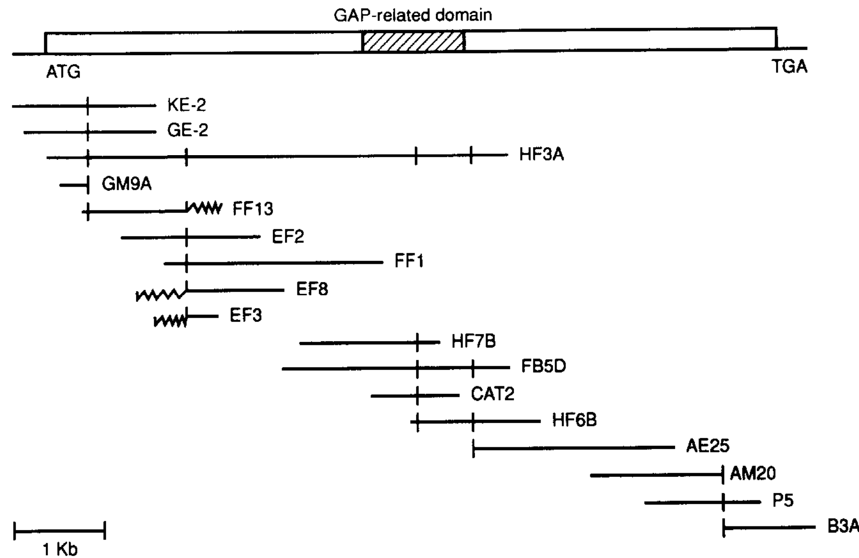

FIG. 1. A schematic diagram representing the cDNA walk in the *NF1* gene. The open reading frame is represented by the wide region bordered by the ATG and TGA codons, and the extent of the GAP-related domain used in complementation studies (Ref. (3)) is indicated. Clones are listed below the schematic of the transcript; straight lines represent authentic transcript and jagged lines represent cocloning events. *Eco*RI sites are represented by vertical lines. Clones B3A and P5 have previously been described (Ref. (62)). Clones AE25, KE-2, and GE-2 were isolated from an endothelial cell cDNA library (Ref. (21)). Clones HF6B, FB5D, HF7B, EF3, EF8, FF1, EF2, FF13, and HF3A were isolated from a fetal brain cDNA library (Stratagene, No. 936206). Clone CAT2 was isolated from the same library by PCR from total phage lysate from the library (Ref. (3)). Clone AM20 was isolated from a human brain (medulla) cDNA library (Clontech, No. HL1091a), and clone GM9A was isolated from a fetal muscle cDNA library (Ref. (30)).

oligo(dT) and random primed, was obtained from Stratagene, La Jolla, CA (No. 936206). Adult human brain (occipital pole and medulla) cDNA libraries, random and oligo(dT) primed, were obtained from Clontech, Palo Alto, CA (No's. HL1091a and HL1089). An oligo d(T) primed human fetal muscle library is described in Koenig *et al.* (1987), and a random primed endothelial cell library is described in Ginsburg *et al.* (1985). Typically, 500,000 plaques of each library were plated and screened (Benton and Davis, 1977) using an aqueous hybridization consisting of 6× SSC, 2× Denhardt's solution, 1 m*M* EDTA, and 0.5% SDS at 65°C. Washes were in 2×, 1×, and if needed, 0.2× SSC, 0.1% SDS at 65°C. Positive plaques were purified, and subcloned into Bluescript plasmid (Stratagene) or rescued as plasmid per λ ZAP instructions (Stratagene) in the case of the fetal brain library.

Sequence Analysis of Clones

Double-stranded sequencing of plasmid clones was performed using Sequenase Version 2.0 (U.S. Biochemicals, Cleveland, OH) per instructions. Sequence compilation was aided by the IBI/Pustell sequence analysis program package (International Biotechnologies, Inc., New Haven, CT). Analysis of the amino acid sequence was performed with the University of Wisconsin Genetics Computer Group protein analysis package (Devereux *et al.*, 1984).

Primer Extension

Total RNA was isolated from fresh human brain (frontal lobe) and melanoma cell line SK-MEL-23 (Carey *et al.*, 1976) as described in Sambrook *et al.* (1989). Polyadenylated RNA was isolated from melanoma total RNA using the FastTrack mRNA isolation kit (Invitrogen Corp., San Diego, CA). For primer extension, an oligomer (5′ AGAGGCAAGGAGAGGGTCTGTG) was synthesized, kinased with ^{32}P, and extended off of brain (total) or melanoma (poly(A)$^+$) RNA (Boorstein and Craig, 1989). Products were analyzed on a 6% denaturing polyacrylamide gel.

RESULTS

Figure 1 shows a cDNA walk extending from the 3′ end of the *NF1* gene. The initial clones P5 and B3A were isolated as previously described (Wallace *et al.*, 1990). Since the *NF1* transcript has been shown to be ubiquitously expressed (Buchberg *et al.*, 1990; Wallace *et al.*, 1990), cDNA walking proceeded in multiple cDNA libraries in order to maximize chances of finding positives. Library sources included fetal muscle,

```
   1 CCCAGCCTCCTTGCCAACGCCCCCTTTCCCTCTCCCCCTCCCGCTCGGCGCTGACCCCCCATCCCCACCCCCGTGGGAACACTGGGAGCCTGCACTCCACAGACCCTCTCCTTGCCTCTT

 121 CCCTCACCTCAGCCTCCGCTCCCCGCCCTCTTCCCGGCCCAGGGCGCCGGCCCACCCTTCCCTCCGCCGCCCCCCGGCCGCGGGGAGGACATGGCCGCGCACAGGCCGGTGGAATGGGTC
   1                                                                                             M  A  A  H  R  P  V  E  W  V

 241 CAGGCCGTGGTCAGCCGCTTCGACGAGCAGCTTCCAATAAAAACAGGACAGCAGAACACACATACCAAAGTCAGTACTGAGCACAACAAGGAATGTCTAATCAATATTTCCAAATACAAG
  11  Q  A  V  V  S  R  F  D  E  Q  L  P  I  K  T  G  Q  Q  N  T  H  T  K  V  S  T  E  H  N  K  E  C  L  I  N  I  S  K  Y  K

 361 TTTTCTTTGGTTATAAGCGGCCTCACTACTATTTTAAAGAATGTTAACAATATGAGAATATTTGGAGAAGCTGCTGAAAAAAATTTATATCTCTCTCAGTTGATTATATTGGATACACTG
  51  F  S  L  V  I  S  G  L  T  T  I  L  K  N  V  N  N  M  R  I  F  G  E  A  A  E  K  N  L  Y  L  S  Q  L  I  I  L  D  T  L

 481 GAAAAATGTCTTGCTGGGCAACCAAAGGACACAATGAGATTAGATGAAACGATGCTGGTCAAACAGTTGCTGCCAGAAATCTGCCATTTTCTTCACACCTGTCGTGAAGGAAACCAGCAT
  91  E  K  C  L  A  G  Q  P  K  D  T  M  R  L  D  E  T  M  L  V  K  Q  L  L  P  E  I  C  H  F  L  H  T  C  R  E  G  N  Q  H

 601 GCAGCTGAACTTCGGAATTCTGCCTCTGGGGTTTTATTTTCTCTCAGCTGCAACAACTTCAATGCAGTCTTTAGTCGCATTTCTACCAGGTTACAGGAATTAACTGTTTGTTCAGAAGAC
 131  A  A  E  L  R  N  S  A  S  G  V  L  F  S  L  S  C  N  N  F  N  A  V  F  S  R  I  S  T  R  L  Q  E  L  T  V  C  S  E  D

 721 AATGTTGATGTTCATGATATAGAATTGTTACAGTATATCAATGTGGATTGTGCAAAATTAAAACGACTCCTGAAGGAAACAGCATTTAAATTTAAAGCCCTAAAGAAGGTTGCGCAGTTA
 171  N  V  D  V  H  D  I  E  L  L  Q  Y  I  N  V  D  C  A  K  L  K  R  L  L  K  E  T  A  F  K  F  K  A  L  K  K  V  A  Q  L

 841 GCAGTTATAAATAGCCTGGAAAAGGCATTTTGGAACTGGGTAGAAAATTATCCAGATGAATTTACAAAACTGTACCAGATCCCACAGACTGATATGGCTGAATGTGCAGAAAAGCTATTT
 211  A  V  I  N  S  L  E  K  A  F  W  N  W  V  E  N  Y  P  D  E  F  T  K  L  Y  Q  I  P  Q  T  D  M  A  E  C  A  E  K  L  F

 961 GACTTGGTGGATGGTTTTGCTGAAAGCACCAAACGTAAAGCAGCAGTTTGGCCACTACAAATCATTCTCCTTATCTTGTGTCCAGAAATAATCCAGGATATATCCAAAGACGTGGTTGAT
 251  D  L  V  D  G  F  A  E  S  T  K  R  K  A  A  V  W  P  L  Q  I  I  L  L  I  L  C  P  E  I  I  Q  D  I  S  K  D  V  V  D

1081 GAAAACAACATGAATAAGAAGTTATTTCTGGACAGTCTACGAAAAGCTCTTGCTGGCCATGGAGGAAGTAGGCAGCTGACAGAAAGTGCTGCAATTGCCTGTGTCAAACTGTGTAAAGCA
 291  E  N  N  M  N  K  K  L  F  L  D  S  L  R  K  A  L  A  G  H  G  G  S  R  Q  L  T  E  S  A  A  I  A  C  V  K  L  C  K  A

1201 AGTACTTACATC  1212
 331  S  T  Y  I
```

FIG. 2. cDNA sequence of the 5′ portion of the *NF1* transcript. The sequence listed has not been previously published and ends where the previously published sequence begins (Ref. (65)). The sequence was compiled from clones KE-2, GE-2, GM9A, EF2, FF13, and HF3A. Both strands were sequenced at least once to complete the sequence. The nucleotide and deduced amino acid sequence are numbered along the left column. The start codon is underlined, and the upstream in-frame stop codon is boxed. The position of the oligonucleotide used for primer extension (Fig. 3) is shown by an arrow. The position of the first intron is indicated by a triangle.

fetal brain, adult brain (occipital pole and medulla), and endothelial cells. Walks proceeded sequentially by isolation of positive phage clones using the most 5′ cDNA insert. The positive clones were characterized by restriction mapping using *Eco*RI and Southern blot analysis using previously isolated inserts. The phage clones were subcloned into Bluescript (Stratagene) and the ends were sequenced to anchor the position of the clones to the transcript map. The cycle was repeated for each walk. Underrepresented regions in any given library were overcome by crossing into another library. The entire transcript as represented in the clones was sequenced multiple times and both strands were sequenced at least once for all previously unpublished sequence.

As the cDNA walk neared completion, a very GC-rich region of the transcript that contained an abnormally high concentration of the dinucleotide CpG, as well as rare cutting restriction endonuclease sites *Eag*I, *Nar*I, and *Sac*II, was encountered at the 5′ end. These sites had been previously placed on the pulsed-field map of this region using the linking clone 17L1 (Fountain *et al.,* 1989a). This clone was isolated from a *Not*I linking library constructed from DNA from flow-sorted chromosome 17 (Wallace *et al.,* 1989), and contains the sequences flanking both sides of a genomic *Not*I site. This probe was originally used to detect a translocation breakpoint within the *NF1* gene, which narrowed the search for the gene to a region of only about 300 kb (Fountain *et al.,* 1989b). Southern blots using the CpG-rich cDNAs as probes against 17L1 demonstrated that the most 5′ sequences obtained were indeed located in the centromeric half of this clone (17L1B), approximately 300 kb from the 3′ stop codon (data not shown).

The most 5′ cDNA clone, KE-2, isolated from the endothelial cell cDNA library, contained an in-frame stop codon (Fig. 2). Downstream from this stop codon, the first ATG fits the rules for a proper translational start (Koząk, 1986). Overlapping sequences have been found in cDNA clones from three different tissues (fetal muscle, fetal brain, and endothelial cells). We propose that this ATG codon represents the authentic start codon, giving the protein a total of 2818 amino acids with a predicted molecular weight of 327 kDa.

To determine whether a substantial portion of the 5′ end of this transcript remained uncloned, a primer

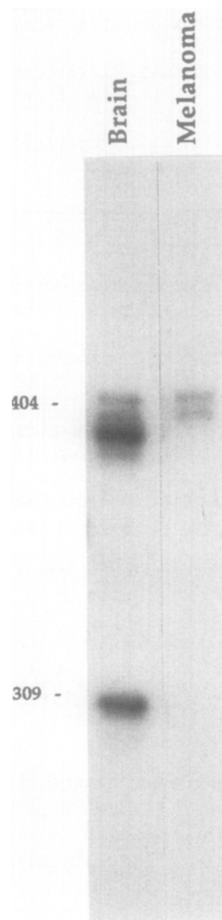

FIG. 3. Primer extension off of human brain (frontal lobe) total RNA and melanoma cell line SK-MEL-23 poly(A)$^+$ RNA. The primer was chosen upstream from the proposed start codon at a position shown in Fig. 2. A group of four major products are seen in the brain RNA ranging in size from approximately 380 to 410 bases. A second site is seen at about 300 base from the primer. The melanoma cell line shows only the 410 and 400 base products, even after longer exposure.

extension was performed using human brain (frontal lobe) total RNA and melanoma cell line SK-MEL-23 poly(A)$^+$ RNA. A reverse transcription primer was chosen that was 5′ of the proposed start codon. Figure 3 shows the result of this analysis. A series of four bands ranging in size from 380 to 410 bp is seen in the primer extension from brain RNA. A second prominent band of 300 bp is also seen. Primer extension from the melanoma RNA shows the top two bands at 400 and 410 bp, but does not show the lower band at 300 bp. We have cloned and sequenced 119 bp from the 5′ position of this primer, indicating that at most only 291 bp remains uncloned. Within the cloned sequence lies the start codon and upstream in-frame stop codon. These results indicate that the entire coding region of the transcript has been cloned and sequenced.

Sequencing of the proximal half of the *Not*I linking clone 17L1 (17L1B) demonstrated that the 5′ cDNA sequences from nucleotide 1 to 270 exist in this region of the genome as a single continuous exon. We conclude from this that exon 1 of this transcript contains a majority of sequence that is 5′ untranslated. If another exon is spliced in upstream from this, it would consist entirely of noncoding sequence. We also conclude that the transcriptional start site and the promoter region probably exist in the half linking clone 17L1B.

In the course of screening for cDNA clones that extended beyond our most 5′ clone, two alternative sequences were discovered. Both begin at position 270, the position of the first splice junction, and are represented by single cDNA clones from a fetal brain cDNA library. Further analysis of these clones has revealed that neither represent authentic NF1 mRNA. One sequence appeared to be derived from an unspliced message, as it contains a perfect splice-acceptor consensus sequence, a pyrimidine stretch of 20 bases, and a lariat formation consensus sequence (Sharp, 1987). This has been confirmed by designing primers that extend across the splice junction and by showing that these primers will amplify the expected fragment using genomic DNA as a template (data not shown). A second unusual clone diverges at the exact same position, yet has a different sequence. This clone has been shown to be the result of a rare trans-splicing event since the new sequence does not map to chromosome 17. We therefore conclude that there is no evidence for alternative splicing at the extreme 5′ end of the gene.

We have been unable to characterize the 3′ end of the *NF1* transcript, as a poly(A) tail has not been found in any cDNA clone. Our previous sequence analysis (Wallace *et al.*, 1990) has shown the proper position of the stop codon. Downstream from this the sequence is very A rich, with some regions that are capable of priming with oligo(dT) during construction of the cDNA libraries. The *NF1* transcript has been estimated to be 13 kb by its migration on a Northern blot (Wallace *et al.*, 1990). To date we have cloned and sequenced 9 kb of this message. The primer extension results (Fig. 3) indicate that the majority of uncloned sequence from this transcript must arise from a long (approximately 4 kb) 3′ untranslated region. Alternatively, our estimates of transcript size may be incorrect, as size estimates in this range of Northern blotting are difficult.

Two alternatively processed forms of this primary transcript have been discovered. A 54-bp insertion coding for an additional 18 amino acids near the 3′ end of the transcript has been described previously (Cawthon *et al.*, 1990). A 63-bp insertion coding for an additional 21 amino acids within one of the most con-

```
   1 MAAHRPVEWV QAVVSRFDEQ LPIKTGQQNT HTKVSTEHNK ECLINISKYK FSLVISGLTT ILKNVNNMRI FGEAAEKNLY LSQLIILDTL EKCLAGQPKD
 101 TMRLDETMLV KQLLPEICHF LHTCREGNQH AAELRNSASG VLFSLSCNNF NAVFSRISTR LQELTVCSED NVDVHDIELL QYINVDCAKL KRLLKETAFK
 201 FKALKKVAQL AVINSLEKAF WNWVENYPDE FTKLYQIPQT DMAECAEKLF DLVDGFAEST KRKAAVWPLQ IILLILCPEI IQDISKDVVD ENNMNKKLFL
 301 DSLRKALAGH GGSRQLTESA AIACVKLCKA STYINWEDNS VIFLLVQSMV VDLKNLLFNP SKPFSRGSQP ADVDLMIDCL VSCFRISPHN NQHFKICLAQ
 401 NSPSTFHYVL VNSLHRIITN SALDWWPKID AVYCHSVELR NMFGETLHKA VQGCGAHPAI RMAPSLTFKE KVTSLKFKEK PTDLETRSYK YLLLSMVKLI
 501 HADPKLLLCN PRKQGPETQG STAELITGLV QLVPQSHMPE IAQEAMEALL VLHQLDSIDL WNPDAPVETF WEISSQMLFY ICKKLTSHQM LSSTEILKWL
 601 REILICRNKF LLKNKQADRS SCHFLLFYGV GCDIPSSGNT SQMSMDHEEL LRTPGASLRK GKGNSSMDSA AGCSGTPPIC RQAQTKLEVA LYMFLWNPDT
 701 EAVLVAMSCF RHLCEEADIR CGVDEVSVHN LLPNYNTFME FASVSNMMST GRAALQKRVM ALLRRIEHPT AGNTEAWEDT HAKWEQATKL ILNYPKAKME
 801 DGQAAESLHK TIVKRRMSHV SGGGSIDLSD TDSLQEWINM TGFLCALGGV CLQQRSNSGL ATYSPPMGPV SERKGSMISV MSSEGNADTP VSKFMDRLLS
 901 LMVCNHEKVG LQIRTNVKDL VGLELSPALY PMLFNKLKNT ISKFFDSQGQ VLLTDTNTQF VEQTIAIMKN LLDNHTEGSS EHLGQASIET MMLNLVRYVR
1001 VLGNMVHAIQ IKTKLCQLVE VMMARRDDLS FCQEMKFRNK MVEYLTDWVM GTSNQAADDD VKCLTRDLDQ ASMEAVVSLL AGLPLQPEEG DGVELMEAKS
1101 QLFLKYFTLF MNLLNDCSEV EDESAQTGGR KRGMSRRLAS LRHCTVLAMS NLLNANVDSG LMHSIGLGYH KDLQTRATFM EVLTKILQQG TEFDTLAETV
1201 LADRFERLVE LVTMMGDQGE LPIAMALANV VPCSQWDELA RVLVTLFDSR HLLYQLLWNM FSKEVELADS MQTLFRGNSL ASKIMTFCFK VYGATYLQKL
1301 LDPLLRIVIT SSDWQHVSFE VDPTRLEPSE SLEENQRNLL QMTEKFFHAI ISSSSEFPPQ LRSVCHCLYQ VVSQRFPQNS IGAVGSAMFL RFINPAIVSP
1401 YEAGILDKKP PPRIERGLKL MSKILQSIAN HVLFTKEEHM RPFNDFVKSN FDAARRFFLD IASDCPTSDA VNHSLSFISD GNVLALHRLL WNNQEKIGQY
1501 LSSNRDHKAV GRRPFDKMAT LLAYLGPPEH KPVADTHWSS LNLTSSKFEE FMTRHQVHEK EEFKALKTLS IFYQAGTSKA GNPIFYYVAR RFKTGQINGD
1601 LLIYHVLLTL KPYYAKPYEI VVDLTHTGPS NRFKTDFLSK WFVVFPGFAY DNVSAVYIYN CNSWVREYTK YHERLLTGLK GSKRLVFIDC PGKLAEHIEH
1701 EQQKLPAATL ALEEDLKVFH NALKLAHKDT KVSIKVGSTA VQVTSAERTK VLGQSVFLND IYYASEIEEI CLVDENQFTL TIANQGTPLT FMHQECEAIV
1801 QSIIHIRTRW ELSQPDSIPQ HTKIRPKDVP GTLLNIALLN LGSSDPSLRS AAYNLLCALT CTFNLKIEGQ LLETSGLCIP ANNTLFIVSI SKTLAANEPH
1901 LTLEFLEECI SGFSKSSIEL KHLCLEYMTP WLSNLVRFCK HNDDAKRQRV TAILDKLITM TINEKQMYPS IQAKIWGSLG QITDLLDVVL DSFIKTSATG
2001 GLGSIKAEVM ADTAVALASG NVKLVSSKVI GRMCKIIDKT CLSPTPTLEQ HLMWDDIAIL ARYMLMLSFN NSLDVAAHLP YLFHVVTFLV ATGPLSLRAS
2101 THGLVINIIH SLCTCSQLHF SEETKQVLRL SLTEFSLPKF YLLFGISKVK SAAVIAFRSS YRDRSFSPGS YERETFALTS LETVTEALLE IMEACMRDIP
2201 TCKWLDQWTE LAQRFAFQYN PSLQPRALVV FGCISKRVSH GQIKQIIRIL SKALESCLKG PDTYNSQVLI EATVIALTKL QPLLNKDSPL HKALFWVAVA
2301 VLQLDEVNLY SAGTALLEQN LHTLDSLRIF NDKSPEEVFM AIRNPLEWHC KQMDHFVGLN FNSNFNFALV GHLLKGYRHP SPAIVARTVR ILHTLLTLVN
2401 KHRNCDKFEV NTQSVAYLAA LLTVSEEVRS RCSLKHRKSL LLTDISMENV PMDTYPIHHG DPSYRTLKET QPWSSPKGSE GYLAATYPTV GQTSPRARKS
2501 MSLDMGQPSQ ANTKKLLGTR KSFDHLISDT KAPKRQEMES GITTPPKMRR VAETDYEMET QRISSSQQHP HLRKVSVSES NVLLDEEVLT DPKIQALLLT
2601 VLATLVKYTT DEFDQRILYE YLAEASVVFP KVFPVVHNLL DSKINTLLSL CQDPNLLNPI HGIVQSVVYH EESPPQYQTS YLQSFGFNGL WRFAGPFSKQ
2701 TQIPDYAELI VKFLDALIDT YLPGIDEETS EESLLTPTSP YPPALQSQLS ITANLNLSNS MTSLATSQHS PGIDKENVEL SPTTGHCNSG RTRHGSASQV
2801 QKQRSAGSFK RNSIKKIV
```

FIG. 4. The complete amino acid sequence of the *NF1* gene product. The sequence was deduced from the open reading frame of sequenced clones from a cDNA walk. Boxed regions correspond to the most statistically significant regions of similarity among the GAP family of proteins with the invariant residues marked with stars (Ref. (63)), some of which appear to be functionally significant in a yeast complementation assay (D. Gutmann, unpublished observations). Residues underlined with a single line are potential cAMP-dependent protein kinase recognition sites (Ref. (22)). Residues that are double underlined represent a potential tyrosine phosphorylation recognition sequence (Ref. (15)). The position of a 21-amino-acid insertion (ATCHSLLNKATVKEKKENKKS) representing an alternatively spliced product is shown with a dark triangle (Andersen *et al.*, manuscript in preparation.) The position of an 18-amino-acid insertion (ASLPCSN-SAVFMQLFPHQ) representing an alternatively spliced product is shown by a open triangle (Ref. (64)). We found three regions of our nucleotide sequence at variance with previously published sequence (Ref. (64)), two resulting in changes in the amino acid sequence. Residue number 496 in our clones shows an ATG methionine codon rather than an ATA isoleucine codon. Another sequence variation at residue 1183 shows an CTG leucine codon rather than the previously published CTC. Our clones also lacked an extra CAT histidine codon after residue number 1555. The latter two changes noted agree with those of Martin *et al.* (32), from their sequence of a PCR clone of the GAP-related domain region.

served regions of the GAP-related domain has been discovered (Fig. 4; Andersen *et al.*, manuscript in preparation). The significance of this insertion to the *ras* GTPase activation will be discussed elsewhere.

Figure 4 shows the complete amino acid sequence of the primary *NF1* transcript. Boxed areas indicate the three blocks of homology most conserved between the GAP family of proteins (Wang *et al.*, 1991). The positions of the alternatively spliced exons and their sequence are shown. There are no SH2 or SH3 domains (*src* homology domains), which are present in GAP (Koch *et al.*, 1991). The protein shows no apparent membrane-spanning region and is predicted to be cytosolic by discriminant analysis (Klein *et al.*, 1985). A potential leucine zipper is present beginning at amino acid residue 1834, but this region is not predicted to be in an α-helical conformation due to the presence of a proline in the middle of the repeat. Six potential cAMP-dependent protein kinase phosphorylation sites and a single potential tyrosine phosphorylation

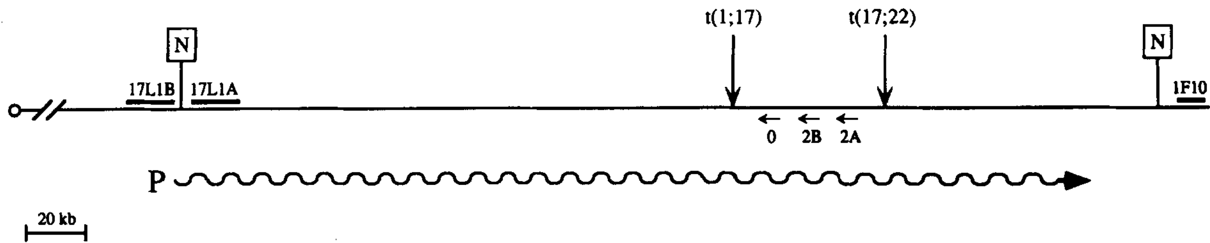

FIG. 5. Extent of the *NF1* transcript on the genomic map of chromosome 17. The data were generated by mapping cDNAs against the pulsed-field restriction map of the region (Ref. (19)). *Not*I sites representing the positions of undermethylated CpG islands in genomic DNA are shown with a boxed N. The 5′ end of the transcript begins in the centromeric half of the *Not*I linking clone 17L1B (Refs. (19,20,61)). It extends through the position of a t(1;17) NF1 translocation breakpoint (Refs. (20,37)) and beyond the position of a t(17;22) NF1 translocation breakpoint (Ref. (37)). The transcript ends before the second CpG island. The gene therefore extends a maximum of 300 kb. Three other transcripts are embedded within a single intron and are transcribed off the complementary strand (D. Marchuk, unpublished observations; Refs. (64,65)). The three genes are EVI2A (Refs. (38,13)), EVI2B (Ref. (14)), and OMgp (Refs. (34,59)).

site are present. The sequence shows no significant homology to the recently described *Bcr*-related GAP family, which includes *n*-chimaerin, and GAP^{rho} (Diekmann *et al.*, 1991) and possibly the p85 of bovine brain phosphatidylinositol 3-kinase (Otsu *et al.*, 1991). There is also no homology to the recently described *rap*1GAP (Rubinfeld *et al.*, 1991).

The size of the *NF1* gene has been determined by mapping cDNA clones back to the pulsed-field restriction map of the region (Fountain *et al.*, 1989a). The 5′ end is just beyond the *Not*I site within the linking clone 17L1B, the first intron beginning 81 bp centromeric to the *Not*I site. The most 3′ clone we have isolated does not extend beyond the *Not*I site defining the next CpG island, and defines a maximum gene size of approximately 300 kb (Fig. 5). This assumes that the remainder of the 3′ untranslated region yet uncloned exists in a single exon. We have not characterized all of the intron–exon borders of the gene, but would estimate by the number of bands on a genomic Southern blot that it contains in excess of 30 exons. The three previously described embedded genes (EVI2A, EVI2B, and OMgp) are transcribed from the opposite strand and are contained within a single intron (D. Marchuk, unpublished observations; Xu *et al.*, 1990a).

DISCUSSION

Initial partial sequences of the 3′ end of the *NF1* gene (Cawthon *et al.*, 1990a; Wallace *et al.*, 1990) revealed very little in the way of sequence homologies that could provide a clue to the function of the gene product. Further cloning and sequence analysis

TABLE 1

Results of Database Searching with Individual NF1GRP Domains

Query sequence	Related sequence	Karlin score	*p*-value	No. of segments matched
N-terminal domain	Yeast IRA2	66	0.0011	10
	Yeast IRA1	72	0.12	1
GAP-related domain	Bovine GAP	81	2.9×10^{-14}	3
	Yeast IRA2	120	6.5×10^{-13}	3
	Yeast IRA1	112	8.2×10^{-09}	3
C-terminal domain	Yeast IRA2	92	7.1×10^{-11}	2
	Yeast IRA1	80	8.5×10^{-07}	2

Note. The BLASTP program (Ref. (2)) was used to search a composite database consisting of all nonidentical protein sequences from the following databases: NBRF/PIR (release 28.0), SWISS-PROT (release 18.0) GenPept (release 64.3), GenPept (daily update, 6/20/91), and NCBI's GenInfo Backbone (prerelease version 6/20/91). At the time of submission this composite database contained 14,884,150 residues in 55,571 sequences. The BLASTP program was used with default parameters except that the PAM250 matrix was used for scoring. Entries in the table represent a summary of the significant results. The Karlin scores and *p*-values were automatically computed according to Karlin and Altschul (26). The Karlin score given represents the highest-scoring segment and the *p*-value is that of the most significant segment (when multiple, nonoverlapping matching segments were identified). Thus, the *p*-values indicate how surprising it is to find such similarities in a sequence collection of this size. Significant matches between NF1GRP domains and the other proteins generally correlated with the homology blocks previously defined by multiple alignment analyses (Refs. (3, 63)). Retrieval codes for the sequences cited in this table refer to the NBRF/PIR database and are as follows: A30135 for IRA1, A35656 for IRA2, and S01966 for bovine GAP. The N-terminal, GAP-related, and C-terminal domains of NF1GRP correspond to residues 1–1132, 1133–1537, and 1538–2818, respectively, of the protein.

(Buchberg *et al.*, 1990; Xu *et al.*, 1990a) revealed homology to the mammalian GAP protein and yeast IRA1 and IRA2 gene products, which modulate the activity of the p21 *ras* protein in their respective hosts by accelerating the rate at which *ras* hydrolyzes GTP to become inactive *ras*-GDP. This provided the first glimpse of the function of *NF1;* like GAP, it may be an upstream regulatory protein for *ras* (or a *ras*-related protein) with its normal function being to down-regulate one or more members of the *ras* family involved in mitogenic signal transduction. This model received further support when it was shown that the proposed GAP-related domain of *NF1* (*NF1*-GRD) could complement loss of IRA function in yeast, and that it could stimulate *ras*-GTPase activity *in vivo* and *in vitro* (Ballester *et al.*, 1990; Martin *et al.*, 1990; Xu *et al.*, 1990b). We propose that the *NF1* gene product be named the *NF1*-GAP-related protein (NF1GRP; Gutmann *et al.*, 1991).

An alternate model of *ras*-NF1GRP interaction postulates that NF1GRP may instead (or in addition) be a downstream effector for *ras*. A downstream model has also been proposed for the related GAP. Mutations in the putative *ras* effector domain inactivate the transforming ability of *ras* and block GTPase activation by GAP, yet retain guanine nucleotide binding capacity (Adari *et al.*, 1988; Cales *et al.*, 1988). *NF1*-GRD interactions with effector and oncogenic mutants of *ras* have shown similar results, suggesting that NF1GRP may also interact with *ras* p21 through its effector domain and be a target of activated *ras* (Martin *et al.*, 1990; Xu *et al.*, 1990b; Bollag and McCormick, 1991). It should be cautioned, however, that none of these studies were done using full-length NF1GRP.

Nonetheless, the downstream effector model is attractive for its ability to account for the role of *ras* in certain cells of neuroectodermal origin. In the rat pheochromocytoma cell line PC12, activated *ras* induces differentiation and blocks proliferation (Bar-Sagi and Feramisco, 1985; Noda *et al.*, 1985). Inhibition of *ras* in these cells blocks neural differentiation normally induced by nerve growth factor (Hagag *et al.*, 1986; Szeberenyi *et al.*, 1990). Activated *ras* has also been shown to induce cell cycle arrest when introduced into rat Schwann cells (Ridley *et al.*, 1988). This is significant because Schwann cells may be the original cells that recruit other cell types in the formation of neurofibromas and neurofibrosarcomas (Ratner *et al.*, 1990; Sheela *et al.*, 1990). Therefore, if NF1GRP is the target (effector) of *ras*, then loss of NF1GRP function could lead to a block of normal differentiation resulting in unrestrained proliferation of Schwann cells.

Either of these two models, with NF1GRP as the upstream negative regulator or downstream effector, are consistent with *NF1* being a tumor suppressor gene, where the phenotype results from the loss of both alleles of the gene (Knudson, 1985). Previous studies of NF1 tumors did not show a consistent loss of heterozygosity for *NF1* in 17q11.2 (Skuse *et al.*, 1989; Menon *et al.*, 1990; Glover *et al.*, 1991), but these interpretations were made difficult by frequent losses on 17p, apparently reflecting the major role that loss of the p53 gene plays in tumor progression in this disorder (Nigro *et al.*, 1989; Menon *et al.*, 1990). More recent analyses have indicated that loss of heterozygosity involving only 17q can be demonstrated at least for some tumors (Skuse, 1990; E. Legius and T. Glover, personal communication; B. Ponder, personal communication).

The large *NF1* transcript and 300-kb gene size represent a large target for mutations. Assuming that the NF1 phenotype results from loss of function mutations, causative mutations may be dispersed throughout the coding and regulatory region. In the patients surveyed thus far, this seems to be the case (Cawthon *et al.*, 1990a; Wallace *et al.*, 1990; M. Wallace *et al.*, unpublished observations). The size of the gene alone, however, cannot fully account for the high mutation rate. At best, the possible target size is only a factor of 10 larger than other genes, whereas the mutation rate is about 100-fold higher than the usual rate for a single locus. There is presently not enough mutation data to determine whether there is a mutational hot spot within this region.

The complete sequence of the *NF1* gene product has provided few additional clues to its function. The lack of SH2 and SH3 domains are in contrast to GAP. Homologous to noncatalytic regions of the oncogene *src*, these domains are thought to direct interactions with phosphotyrosine proteins involved in signal transduction (Koch *et al.*, 1991). Their absence in NF1GRP implies that NF1GRP and GAP are not interchangeable in the cell, and that NF1GRP is probably not directly modulated through tyrosine phosphorylation by activated growth factor receptors. The potential sites for tyrosine and serine/threonine phosphorylation (Fig. 4) may mean that an intermediate between the activated receptor and NF1GRP may modulate its activity, since there is evidence that NF1GRP is phosphorylated on serine and threonine residues (J. Downward, personal communication). A potential candidate for this intermediate could be one of the members of the ERK family, which are activated by tyrosine phosphorylation by nerve growth factor and are themselves serine/threonine protein kinases (Boulton *et al.*, 1991). Certainly, the large size of the product in relation to the small portion conferring GAP activity indicates that other domains may be involved in modulating the *ras*-GTPase activity of this protein, or carrying out entirely different func-

tions. The sequence homology with yeast IRA1 and IRA2 (Table 1) extends beyond the GAP catalytic domain toward both termini, with strong homology in the carboxy terminus. The homology, both globally and within each domain, is most significant with the IRA2 protein. There may be therefore more extensive functional homology with these members of the GAP family than with GAP itself. It may be useful to think of the NF1GRP as consisting of three domains; an amino-terminus of unknown function, a GAP-related middle domain, and a carboxy-terminus related to the IRA2 and IRA1 gene products. Unfortunately, the functions of these domains of the IRA gene products are not known even in this simple eukaryote. For NF1GRP, interaction with other factors in the pathway leading to differentiation, such as the low-affinity nerve growth factor receptor and the *trk* oncogene product, or factors that mediate between these and NF1GRP, may be localized to these yet undefined regions. Interactions at these other domains may also ultimately provide a clue as to the reasons why a mutation in a ubiquitously expressed gene reveals its character predominantly in cells derived from the embryonic neural crest.

It is presently unclear how the alternatively spliced transcripts and the corresponding protein isoforms play their unique role in NF1GRP function. We have not yet conducted an exhaustive search of the tissues where each of these transcripts can be found. We know that at least in some tissues, however, their expression is not mutually exclusive. These alternative forms may play a role in the diverse clinical manifestations of the disorder. Germline mutations in some of the alternatively spliced exons may give rise to some of the more unusual NF1 phenotypes.

ACKNOWLEDGMENTS

We thank D. Siemieniak for help with the GCG package of protein sequence analysis tools. An endothelial cell cDNA library was the gift of D. Ginsburg and a fetal muscle cDNA library was the gift of F. Boyce and L. Kunkel. We thank N. Dracopoli and J. Fountain for the melanoma strain SK-MEL-23. We thank J. Koh for help with primer extensions and B. Sandri for assistance with the figures. This work was supported by the Howard Hughes Medical Institute, a Young Investigator Award from the National Neurofibromatosis Foundation (M.R.W.), and the National Institutes of Health Grants NS23410 (F.S.C.) and HG00018 (D.A.M.).

REFERENCES

1. Adari, H., Lowy, D. R., Willumsen, B. M., Der, C. J., and McCormick, F. (1988). Guanosine triphosphate activating protein (GAP) interacts with the p21 *ras* effector binding domain. *Science* **240:** 518–521.
2. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.
3. Ballester, R., Marchuk, D., Boguski, M., Saulino, A., Letcher, R., Wigler, M., and Collins, F. (1990). The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. *Cell* **63:** 851–859.
4. Bar-Sagi, D., and Feramisco, J. (1985). Microinjection of the *ras* oncogene protein into PC12 cells induces morphological differentiation. *Cell* **42:** 841–848.
5. Benton, W. D., and Davis. R. W. (1977). Screening lambdagt recombinant clones by hybridization to single plaques in situ. *Science* **196:** 180–182.
6. Bollag, G., and McCormick, F. (1991). Differential regulation of rasGAP and neurofibromatosis gene product activities. *Nature* **351:** 576–579.
7. Boorstein, W. R., and Craig, E. A. (1989). Primer extension analysis of RNA. In "Methods in Enzymology" (J. E. Dahlberg and J. N. Abelson, Eds.), Vol. 180, pp. 347–369, Academic Press, San Diego.
8. Boulton, T. G., Nye, S. H., Robbins, D. J., Ip, N. Y., Radziejewska, E., Morgenbesser, S. D., DePinho, R. A., Panayotatos, N., Cobb, M. H., and Yancopoulos, G. D. (1991) ERKs: A family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. *Cell* **65:** 663–675.
9. Buchberg, A., Cleveland, L., Jenkins, N., and Copeland, N. (1990). Sequence homology shared by neurofibromatosis type-1 gene and IRA-1 and IRA-2 negative regulators of the RAS cyclic AMP pathway. *Nature* **347:** 291–294.
10. Cales, C. Hancock, J., Marshall, C., and Hall, A. (1988). The cytoplasmic protein GAP is implicated as the target for regulation by the *ras* gene product. *Nature* **332:** 548–551.
11. Carey, T. E., Takahashi, T., Resnick, L. A., Oettgen, H. F., and Old, L. J. (1976). Cell surface antigens of human malignant melanoma: mixed hemadsorption assays for humoral immunity to cultured autologous melanoma cells. *Proc. Natl. Acad. Sci. USA* **73:** 3278–3282.
12. Cawthon, R., Weiss, R., Xu, G., Viskochil, D., Culver, M., Stephens, J., Robertson, M., Dunn, D., Gesteland, R., O'Connell, P., and White, R. (1990a). A major segment of the neurofibromatosis type 1 gene; cDNA sequence, genomic structure, and point mutations. *Cell* **62:** 193–201.
13. Cawthon, R., O'Connell, P., Buchberg, A. M., Viskochil, D., Weiss, R., Culver, M., Stevens, J., Jenkins, N., Copeland, N., and White, R. (1990b). Identification and characterization of transcripts from the neurofibromatosis 1 region: The sequence and genomic structure of EV12 and mapping of other transcripts. *Genomics* **7:** 555–565.
14. Cawthon, R. M., Andersen, L. B., Buchberg, A. M., Xu, G., O'Connell, P., Viskochil, D., Weiss, R. B., Wallace, M. R., Marchuk, D. A., Culver, M., Stevens, J., Jenkins, N. A., Copeland, N. G., Collins, F. S., and White, R. (1991). cDNA sequence and genomic structure of EV12B, a gene lying within an intron of the neurofibromatosis type 1 gene. *Genomics* **9:** 446–460.
15. Cooper, J. A., Esch, F. S., Taylor, S. S., and Hunter, T. (1984). Phosphorylation sites in enolase and lactate dehydrogenase utilized by tyrosine protein kinases *in vivo and in vitro. J. Biol. Chem.* **259:** 7835–7841.
16. Crowe, F., Schull, W., and Neel, J. (1956). "A Clinical, Pathological, and Genetic Study of Multiple Neurofibromatosis," Thomas, Springfield, IL.
17. Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12:** 387–395.
18. Diekmann, D., Brill, S., Garrett, M. D., Totty, N.,

HSUAN, J., MONFRIES, C., HALL, C., LIM, L., AND HALL, A. (1991). *Bcr* encodes a GTPase-activating protein for p21 *ras*. *Nature* **351:** 400–402.

19. FOUNTAIN, J. W., WALLACE, M. R., BRERETON, A. M., O'CONNELL, P., WHITE, R. L., RICH, D., LEDBETTER, D. H., LEACH, R. J., FOURNIER, R. E. K., MENON, A. G., GUSELLA, J. F., BARKER, D., STEPHENS, K., AND COLLINS, F. S. (1989a). Physical mapping of the von Recklinghausen neurofibromatosis region on chromosome 17. *Am. J. Hum. Genet.* **44:** 58–67.

20. FOUNTAIN, J. W., WALLACE, M. R., BRUCE, M. A., SEIZINGER, B. R., MENON, A. G., GUSELLA, J. F., MICHELS, V. V., SCHMIDT, M. A., DEWALD, G. W., AND COLLINS, F. S. (1989b). Physical mapping of a translocation breakpoint in neurofibromatosis. *Science* **244:** 1085–1087.

21. GINSBURG, D., HANDIN, R. I., BONTHRON, D. T., DONLON, T. A., BRUNS, G. A. P., LATT, S. A., AND ORKIN, S. H. (1985). Human von Willebrand factro (VWF): Isolation of cDNA clones and chromosomal localization. *Science* **228:** 1401–1406.

22. GLASS, D. B., EL-MAGHRABI, M. R., AND PILKIS, S. J. (1986). Synthetic peptides corresponding to the site phosphorylated in 6-phosphofructo-2-kinase/fructose-2, 6-bisphosphatase as substrates of cyclic nucleotide-dependent protein kinases. *J. Biol. Chem.* **261:** 2987–2993.

23. GLOVER, T. W., STEIN, C. K., LEGIUS, E., ANDERSEN, L. B., BRERETON, A., AND JOHNSON, S. (1991). Molecular and cytogenetic analysis of tumors in von Recklinghausen neurofibromatosis. *Genes Chrom. Cancer* **3:** 62–70.

24. GUTMANN, D. H., WOOD, D. L., AND COLLINS, F. S. (1991). Identification of the neurofibromatosis type 1 gene product. *Proc. Natl. Acad. Sci. USA*, in press.

25. HAGAG, N., HALEGOUA, S., AND VIOLA, M. (1986). Inhibition of growth factor induced differentiation of PC12 cells by microinjection of antibody to *ras* p21. *Nature* **319:** 680–682.

26. KARLIN, S., AND ALTSCHUL, S. F. (1990). Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. USA* **87:** 2264–2268.

27. KLEIN, P., KANEHISA, M., AND DELISI, C. (1985). The detection and classification of membrane-spanning proteins. *Biochim. Biophys. Acta* **815:** 468–476.

28. KNUDSON, A. G. (1985). Hereditary cancer and antioncogenes. *Cancer Res.* **45:** 1437–1443.

29. KOCH, C. A., ANDERSON, D., MORAN, M. F., ELLIS, C., AND PAWSON, T. (1991). SH2 and SH3 domains: Elements that control interactions of cytoplasmic signaling proteins. *Science* **252:** 668–674.

30. KOENIG, M., HOFFMAN, E. P., BERTELSON, C. J., MONACO, A. P., FEENER, C., AND KUNKEL, L. M. (1987). Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50:** 509–517.

31. KOZAK, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44:** 284–292.

32. MARTIN, G. A., VISKOCHIL, D., BOLLAG, G., MCCABE, P. C., CROSIER, W. J., HAUBRUCK, H., CONROY, L., CLARK, R., O'CONNELL, P., CAWTHON, R. M., INNIS, M. A., AND MCCORMICK, F. (1990). The GAP-related domain of the neurofibromatosis type 1 gene product interacts with *ras* p21. *Cell* **63:** 843–849.

33. MENON, A. G., ANDERSON, K. M., RICCARDI, V. M., CHUNG, R. Y., WHALEY, J. M., YANDELL, D. W., FARMER, G. E., FRIEMAN, R. N., LEE, J. K., LI, F. P., BARKER, D. F., LEDBETTER, D. H., KLELCER, A., MARTUZA, R. L., GUSELLA, J. R., AND SEIZINGER, B. R. (1990). Chromosome 17p deletions and p53 gene mutations associated with the formation of malignant neurofibromas in von Recklinghausen neurofibromatosis. *Proc. Natl. Acad. Sci. USA* **87:** 5435–5439.

34. MIKOL, D. D., GULCHER, J. R., AND STEFANSSON, K. (1990). The oligodendrocyte-myelin glycoprotein belongs to a distinct family of proteins and contains the HNK-1 carbohydrate. *J. Cell Biol.* **110:** 471–479.

35. NIGRO, J. M., BAKER, S. J., PREISINGER, A. C., JESSUP, J. M., HOSTETTER, R., CLEARY, K., BIBNER, S. H., DAVIDSON, N., BAYLIN, S., DEVILEE, P., GLOVER, T., COLLINS, F. S., WESTON, A., MOLDALI, R., HARRIS, C. C., AND VOGELSTEIN, B. (1989) Mutations in the p53 gene occur in diverse human tumor types. *Nature* **342:** 705–708.

36. NODA, M., KO, M., OGURA, A., LIU, D., AMANO, T., TANSKO, T., AND IKAWA, Y. (1985). Sarcoma viruses carrying *ras* oncogenes induce differentiation-associated properties in a neuronal cell line. *Nature* **318:** 73–75.

37. O'CONNELL, P., LEACH, R., CAWTHON, R., CULVER, M., STEVENS, J., VISKOCHIL, D., FOURNIER, R. E. K., RICH, D., LEDBETTER, D., AND WHITE, R. (1989). Two von Recklinghausen neurofibromatosis translocations map within a 600-kb segment of 17q11.2. *Science* **244:** 1087–1088.

38. O'CONNELL, P., VISKOCHIL, D., BUCHBERG, A. M., FOUNTAIN, J., CAWTHON, R. M., CULVER, M., STEVENS, J., RICH, D. C., LEDBETTER, D. H., WALLACE, M., CAREY, J. C., JENKINS, N. A., COPELAND, N. G., COLLINS, F. S., AND WHITE R. (1990). The human homolog of murine Evi-2 lies between two von Recklinghausen neurofibromatosis translocations. *Genomics* **7:** 547–554.

39. OTSU, M., HILES, I., GOUT, I., FRY, M. J., RUIZ-LARREA, F., PANAYOTOU, G., THOMPSON, A., DHAND, R., HSUAN, J., TOTTY, N., SMITH, A. D., MORGAN, S. J., COURTNEIDGE, S. A., PARKER, P. J., AND WATERFIELD, M. D. (1991). Characterization of two 85 kd proteins that associate with receptor tyrosine kinases, middle-T/pp60 c-*src* complexes, and P13 kinase. *Cell* **65:** 91–104.

40. RATNER, N., LIEBERMAN, M. A., RICCARDI, V. M., AND HONG, D. (1990). Mitogen accumulation in von Recklinghausen neurofibromatosis. *Ann. Neurol.* **27:** 496–501.

41. RICCARDI, V. M. (1981). Von Recklinghausen neurofibromatosis. *N. Engl. J. Med.* **305:** 1617–1627.

42. RICCARDI, V. M., AND EICHNER, J. E. (1986). "Neurofibromatosis: Phenotype, Natural History, and Pathogenesis," Johns Hopkins Univ. Press, Baltimore, MD.

43. RICCARDI, V. M., AND LEWIS, R. A. (1988). Penetrance of von Recklinghausen neurofibromatosis: A distinction between predecessors and descendants. *Amer. J. Hum. Genet.* **42:** 284–289.

44. RIDLEY, A. J., PATERSON, H. F., NOBLE, M., AND LAND, H. (1988). *ras*-mediated cell cycle arrest is altered by nuclear oncogenes to induce Schwann cell transformation. *EMBO J.* **7:** 1635–1645.

45. RUBINFELD, B., MUNEMITSU, S., CLARK, R., CONROY, L., WATT, K., CROSIER, W. J., MCCORMICK, F., AND POLAKIS, P. (1991). Molecular cloning of a GTPase activating protein specific for the *Krev*-1 protein $p21^{rap1}$. *Cell* **65:** 1033–1042.

46. SAMBROOK, J., FRITSCH, E. F., AND MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

47. SHARP, P. A. (1987). Splicing of messenger RNA precursors. *Science* **235:** 766–771.

48. SHEELA, S., RICCARDI, V. M., AND RATNER, N. (1990). Angiogenic and invasive properties of neurofibroma Schwann cells. *J. Cell Biol.* **111:** 645–653.

49. SKUSE, G. R., KOSCIOLEK, B. A., AND ROWLEY, P. T. (1989). Molecular genetic analysis of tumors in von Recklinghausen neurofibromatosis: Loss of heterozygosity for chromosome 17. *Genes Chrom. Cancer* **1:** 36–41.

50. SKUSE, G. R. (1990). Type 1 neurofibromatosis gene: Correction. *Science* **250:** 1749.

51. STUMPF, D. A., ALKSNE, J. F., ANNEGERS, J. F., BROWN, S. S., CONNEALLY, P. M., HOUSMAN, D., LEPPERT, M., MILLER, J. P., MOSS, M. L., PILEGGI, A. J., RAPIN, I., STROHMAN, R. C., SWANSON, L. W., AND ZIMMERMAN, A. (1987). Neurofibromatosis. *In* "NIH Consensus Development Conference Statement," Vol. 6, No. 12.

52. SZEBERENYI, J., CAI, H., AND COOPER, G. (1990). Effect of a dominant inhibitory Ha-*ras* mutation on neuronal differentiation of PC12 cells. *Mol. Cell Biol.* **10:** 5324–5332.

53. TANAKA, K., MATSUMOTO, K., AND TOH-E, A. (1989). IRA1, an inhibitory regulator of the RAS/cyclic AMP pathway in *Saccharomyces cerevisiae. Mol. Cell Biol.* **9:** 757–768.

54. TANAKA, K., NAKAFUKU, M., TAMANOI, F., KAZIRO, Y., MATSUMOTO, K., AND TOH-E, A. (1990). *IRA2,* a second gene of *Saccharomyces cerevisiae* that encodes a protein with a domain homologous to mammalian ras GTPase activating protein. *Mol. Cell Biol.* **10:** 4303–4313.

55. TANAKA, K., LIN, B. K., WOOD, D. R., AND TAMANOI, F. (1991). IRA2, an upstream negative regulator of RAS in yeast, is a RAS GTPase-activating protein. *Proc. Natl. Acad. Sci. USA* **88:** 468–472.

56. TRAHEY, M., AND MCCORMICK, F. (1987). A cytoplasmic protein stimulates normal N-*ras* p21 GTPase, but does not affect oncogenic mutants. *Science* **238:** 542–545.

57. TRAHEY, M., WONG, G., HALENBECK, R., RUBINFELD, B., MARTIN, G., LADNER, M., LONG, C., CROSIER, W., WATT, K., KOTHS, K., AND MCCORMICK, F. (1988). Molecular cloning of two types of GAP from complementary DNA from human placenta. *Science* **242:** 1697–1700.

58. VISKOCHIL, D., BUCHBERG, A. M., XU, G., CAWTHON, R. M., STEVENS, J., WOLFF, R. K., CULVER, M., CAREY, J. C., COPELAND, N. G., JENKINS, N. A., WHITE, R., AND O'CONNELL, P. (1990). Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* **62:** 187–192.

59. VISKOCHIL, D., CAWTHON, R., O'CONNELL, P., XU, G., STEVENS, J., CULVER, M., CAREY, J., AND WHITE, R. (1991). The oligodendrocyte-myelin glycoprotein is embedded within the neurofibromatosis type 1 gene. *Mol. Cell Biol.* **11:** 906–912.

60. VOGEL, U., DIXON, R., SCHABER, M., DIEHL, R., MARSHALL, M., SCOLNICK, E., SIGAL, I., AND GIBBS, J. (1988). Cloning of bovine GAP and its interaction with oncogenic *ras* p21. *Nature* **335:** 90–93.

61. WALLACE, M. R., FOUNTAIN, J. W., BRERETON, A. M., AND COLLINS, F. S. (1989). Direct construction of a chromosome-specific *Not*I linking library from flow-sorted chromosomes. *Nucleic Acids. Res.* **17:** 1665–1677.

62. WALLACE, M. R., MARCHUK, D. A., ANDERSEN, L. B., LETCHER, R., ODEH, H. M., SAULINO, A. M., FOUNTAIN, J. W., BRERETON, A., NICHOLSON, J., MITCHELL, A. L., BROWNSTEIN, B., AND COLLINS, F. S. (1990). Type 1 neurofibromatosis gene: Identification of a large transcript disrupted in three NF1 patients. *Science* **249:** 181–186.

63. WANG, Y., BOGUSKI, M., RIGGS, M., RODGERS, L., AND WIGLER, M. (1991). Sar1, a gene from Schizosaccharomyces pombe encoding a GAP-like protein that regulates ras1. *Cell Reg.* **2:** 453–465.

64. XU, G., O'CONNELL, P., VISKOCHIL, D., CAWTHON, R., ROBERTSON, M., CULVER, M., DUNN, D., STEVENS, J., GESTELAND, R., WHITE, R., AND WEISS, R. (1990a). The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* **62:** 599–608.

65. XU, G., LIN, B., TANAKA, K., DUNN, D., WOOD, D., GESTELAND, R., WHITE. R., WEISS, R., AND TAMANOI, F. (1990b) The catalytic domain of the neurofibromatosis type 1 gene product stimulates *ras* GTPase and complements *ira* mutants of *S. cerevisiae. Cell* **63:** 835–841.