GENOMICS 13, 672–680 (1992)

A Yeast Artificial Chromosome Contig Encompassing the Type 1 Neurofibromatosis Gene

DOUGLAS A. MARCHUK,*,†,[1] ROXANNE TAVAKKOL,*,† MARGARET R. WALLACE,*,†,[2] BERNARD H. BROWNSTEIN,‡ PATRICIA TAILLON-MILLER,‡ CHIN-TO FONG,§ ERIC LEGIUS,* LONE B. ANDERSEN,*,† THOMAS W. GLOVER,* AND FRANCIS S. COLLINS*,†

*Department of Human Genetics and the †Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan 48109; ‡Center for Genetics in Medicine, Washington University School of Medicine, St. Louis, Missouri 63130; and the §Department of Pediatric Genetics, University of Rochester Medical Center, Rochester, New York 14642

Received December 26, 1991; revised March 31, 1992

The yeast artificial chromosome (YAC) system (Burke *et al.*, 1987, *Science* 236: 806–812) allows the direct cloning of large regions of the genome. A YAC contig map of approximately 700 kb encompassing the region surrounding the type 1 neurofibromatosis (*NF1*) locus on 17q11.2 has been constructed. A single YAC containing the entire *NF1* locus has been constructed by homologous recombination in yeast. In the process of contig construction a novel method of YAC end rescue has been developed by YAC circularization in yeast and plasmid rescue in bacteria. YACs containing homology to the *NF1* region but mapping to another chromosome have also been discovered. Sequences of portions of the homologous locus indicate that this other locus is a nonprocessed pseudogene.

INTRODUCTION

Von Recklinghausen neurofibromatosis (NF1) is one of the more common inherited disorders in humans, with an incidence of about 1 in 3000 (Crowe *et al.*, 1956; Riccardi, 1981; Riccardi and Eichner, 1986). The clinical features of the disorder involve various manifestations in tissues derived largely from the embryonic neural crest and include café-au-lait spots, Lisch nodules, and neurofibromas in the majority of affected individuals. Some individuals may also have learning disabilities, more deeply placed plexiform neurofibromas, seizures, or skeletal abnormalities. The risk of malignancy is also increased, especially for optic glioma, neurofibrosarcoma, and brain tumors.

The *NF1* gene has been localized to 17q11.2 by genetic linkage analysis (Barker *et al.*, 1987; Seizinger *et al.*, 1987; Goldgar *et al.*, 1989) and physical mapping studies using patients showing balanced chromosomal translocations involving this region (Fountain *et al.*, 1989b; O'Connell *et al.*, 1989). To further, localize and clone the gene yeast artificial chromosome (YAC) clones from the surrounding region were derived. In the course of this work, a gene showing mutations in individuals affected with *NF1* was identified (Cawthon *et al.*, 1990; Viskochil *et al.*, 1990; Wallace *et al.*, 1990b). One of these YACs was instrumental in the cloning and identification of the *NF1* gene (Wallace *et al.*, 1990b). We now present the complete cloning of the *NF1* region in YACs and the construction of a single YAC containing the entire *NF1* gene by homologous recombination.

MATERIALS AND METHODS

Yeast artificial chromosome clones. Clones of *Saccharomyces cerevisiae* containing YACs were isolated at the Center for Genetics in Medicine (St. Louis) using the PCR-based method of screening the library (Green and Olson, 1990a). Individual clones were grown on AHC medium lacking uracil and tryptophan (Brownstein *et al.*, 1989). Liquid DNA preps of total yeast DNA were prepared according to the methods of Sherman *et al.* (1986). DNA in agarose blocks for pulsed-field gel electrophoresis was prepared according to Brownstein *et al.* (1989).

Pulsed-field gel electrophoresis. Sizes of YAC inserts were determined by CHEF gel electrophoresis (Chu *et al.*, 1986) under the following conditions: 1% agarose gel, $\frac{1}{2}$X TBE, 175–200 V, for 24 h with a 25- to 50-s ramp, using the CHEF-DR II (Bio-Rad). Rare-cutter restriction mapping was done using field-inversion gel electrophoresis (Carle *et al.*, 1986) in a conventional gel box (IBI Model HRH) to enable the use of longer gels and thus expand the usable portion of the gel. Conditions were as follows: 1% agarose gel, $\frac{1}{2}$X TBE, 160 V using a forward ramp of 6–60 s and a reverse ramp of 2–20 s.

Restriction digestion. Agarose blocks (approximately 50 μl volume, or 5×10^8 cells) were washed extensively in TE (five times) and then equilibrated overnight in the appropriate digestion buffer. A total of 30–60 units of enzyme per block was used and digestion was allowed to go for 1 h for partial digest and 4–6 h for complete digests. Reactions

[1] Present address: Department of Radiation Oncology, Division of Cancer Biology, University of Michigan Medical School, Ann Arbor, MI 48109.

[2] Present address: Department of Pediatrics, Division of Genetics, Center for Mammalian Genetics, University of Florida Health Science Center, Gainesville, FL 32610-0296.

0888-7543/92 $5.00

TABLE 1

YAC Clones in the *NF1* Region

Probe	YAC clones	Size (kb)	Comments
17L1A	yA43A9	290	
	yB62G2	480, 170	Two YACs, 480 contains probe and is chimera
c1F10	yA113D7	270	
	yA140G3	210	
	yA140G10	210–140	Unstable (sib of A140G3?)
	yA220G3	140	
pEH1	yA113D11	270	Identical to A113D7
	yB164C5	130	
	yB227C3	270	Chimera
	yD8F4	550	Chimera
	yD28G8	250	Chimera

were stopped by immersing the block in 50 m*M* EDTA for 15 min and then immediately loaded on the gel.

Southern blots. Both conventional and pulsed-field gels were transferred to GeneScreen (Dupont) by alkaline (1.5 *M* NaCl and 0.5 *M* NaOH) transfer. Subsequent treatment of the blots and hybridization and wash conditions were as previously described (Drumm *et al.*, 1988). For partial restriction digests of the YAC clones, probes were the 1.0-kb *Bam*HI–*Ava*I fragment of pBR322 for mapping from the URA end of the YAC vector and the 1.1-kb *Bam*HI–*Pst*I fragment of pBR322 for mapping from the centromere end of the YAC vector. The pBlur8 probe insert (Deininger *et al.*, 1981) was used to probe the complete digest to confirm the number and size of the fragments.

YAC end rescue. A variety of techniques were employed to retrieve the ends of the YACs. Initially, a plasmid rescue approach was used to clone the centromeric end of each YAC. Genomic DNA from each YAC clone was digested with either *Nde*I or *Xho*I and recircularized (ligated) under dilute conditions (final concentration of DNA of 20 ng/μl). The DNA was precipitated and transformed into high-efficiency (greater than 10^8 cfu/μg DNA) competent *Escherichia coli* cells. Resulting colonies were found to contain inserts from the centromeric vector arm to the first *Nde*I or *Xho*I site in the YAC insert.

To clone both ends of the YAC using a similar plasmid rescue approach, a YAC circularization approach was applied. The YAC circularization vector pPM680 (Garza *et al.*, 1989) was cut with *Hin*dIII and *Sal*I and transformed into spheroplasts of the YAC-containing strains (Burgers and Percival, 1987). Transformants were selected for lysine prototrophy and subsequently screened for uracil auxotrophy. Agarose blocks were made from potential circular YACs and analyzed by FIGE. Total genomic DNA from the circular YAC clones was digested with *Sac*I and then recircularized (ligated) under dilute conditions. The DNA was precipitated and transformed into high-efficiency competent *E. coli* cells. Transformants have both ends of the original YAC insert ligated at the first *Sac*I site from each end. Alternatively, either end of the insert could be rescued alone by digestion with *Hin*dIII or *Cla*I for the URA end or *Nde*I, *Bam*HI, or *Xho*I for the CEN end of the YAC. A third method used for recovering ends of YACs was an adaptation of inverse PCR (Silverman *et al.*, 1989, 1991).

RESULTS

Initial Screening of the YAC Clones

The first screening of the CGM YAC library involved two probes derived from 17L1A and c1F10, which flanked the *NF1* region as defined by two balanced chromosome translocations involving 17q11.2 (Fountain *et al.*, 1989b; O'Connell *et al.*, 1989). Two YACs were recovered from the 17L1A screen and five from the c1F10 screen (Table 1). One of the YAC 17L1A clones, yB62G2, was shown to contain two YACs, and the 480-kb YAC that contained sequences from the probe 17L1A was subsequently shown to be chimeric. Both ends of yA43A9 were rescued and shown to lie on chromosome 17.

Because the YACs from the c1F10 screening all seemed to be contained within yA113D7 by fingerprint analysis and sizing on CHEF gels, this YAC was chosen for further study. Plasmid rescue of both ends indicated that this YAC was entirely contained on chromosome 17.

The distal end of yA43A9 (AN1) did not map to yA113D7 nor did the proximal end of yA113D7 map to yA43A9, indicating that a gap existed in the *NF1* region YAC contig. Comparison of the sizes of the YACs with the physical map of the region (Fountain *et al.*, 1989a) indicated that the gap was approximately 60 kb in size. Two new probes were used to rescreen the YAC libraries. One was derived from clone pAN1, the distal end of yA43A9. The other was derived from a jump clone, pEH1, within the region between the two translocation breakpoints (Wallace *et al.*, 1990a).

Eight clones were obtained from the screen using AN1 as a probe. *Alu* fingerprint analysis indicated that all of these were related, yet none of them had a pattern that indicated that they overlapped with either yA43A9 or yA113D7 (data not shown). Subsequent analysis indicates that all of these YACs derive from a region on either chromosome 14 or 22 (see below). Five clones were derived from the screening using pEH1 sequences. One (yD8F4) also contained the sequences of AN1 and thus closed the gap. Both ends of the YAC were rescued and it was found that this YAC was chimeric distal to the probe c1F10.

A total of 11 YACs that contained portions of the *NF1* region were obtained. Their sizes were determined by CHEF gel analysis. Table 1 shows the pertinent information on the YACs. It should be emphasized that of the YACs from the 1F10 screening, only yA113D7 has been extensively analyzed, as fingerprinting analysis indicated that the other YACs appeared to be contained within this larger YAC. However, it does not follow that all of these are nonchimeric, as has been shown for other YACs that show similar *Alu* fingerprints (Gaensler *et al.*, 1991).

YAC End Cloning by Plasmid Rescue Schemes

Initial characterization of the YACs involved fingerprinting to determine the overall relatedness of the YACs and end rescue to determine whether the YACs were chimeric or not. Initially all YACs had the CEN end cloned by a plasmid end rescue scheme using either *Nde*I or *Xho*I as the enzyme for the junction point. This method was almost always successful using *Nde*I. This finding suggested that a plasmid end rescue scheme for both ends would be successful if a circular YAC was the starting point. The YAC circularization vector of Garza *et al.* (1989) was used to create circular YACs by homolo-

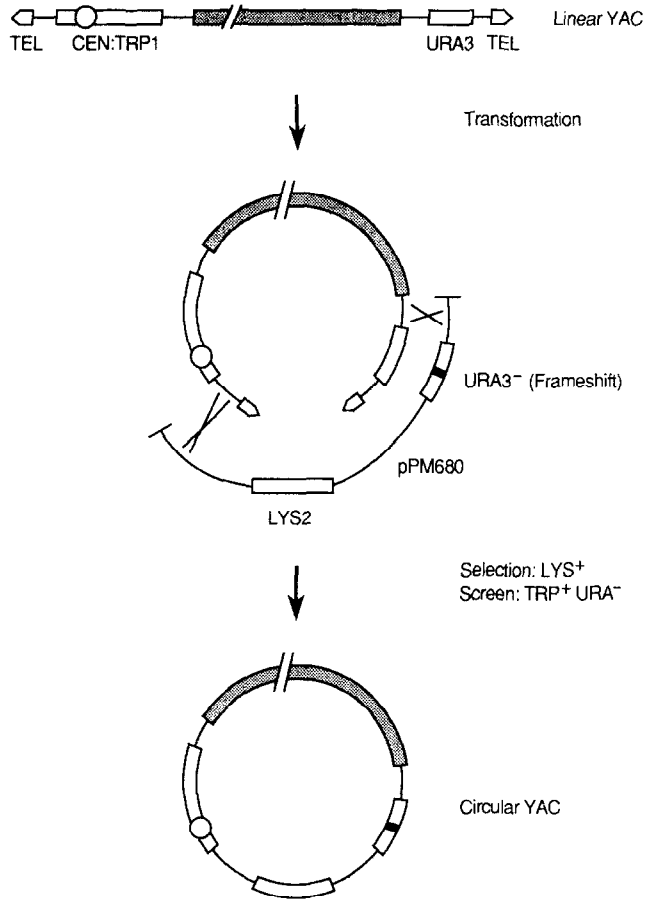

FIG. 1. YAC circularization by homologous recombination *in vivo*. Plasmid pPM680 (Garza *et al.*, 1990) is linearized by digestion with *Sal*I and *Hin*dIII and transformed into spheroplasts of a YAC-containing yeast clone. Transformants are selected for lysine prototrophy and subsequently screened for tryptophan prototrophy and uracil auxotrophy. The resulting clone contains a YAC that has lost its telomeres but has been circularized.

gous recombination in yeast (Fig. 1). After selecting for lysine prototrophy and screening for uracil auxotrophy, high-molecular-weight DNA was prepared from colonies to prove that they contained circular YACs. Field-inversion gel electrophoresis conditions were chosen so that circular YACs did not enter the gel. The blot was probed with radiolabeled pBR322. Figure 2 shows that for the circular YAC, the hybridizing band stays in the well. As a further proof, digestion of the circular YAC with a restriction enzyme that cuts only once in the insert produces a hybridizing band identical in size to that of the original linear YAC.

To ensure that gross rearrangements did not occur during the circularization process, comparisons were made of fingerprint patterns derived from linear and circular forms of the same YAC. Figure 3 shows identical Southern blot patterns using an *Alu* repeat probe for both forms of the YAC. This indicates that no major deletions or internal rearrangements can be detected upon circularization.

For double-end rescue, liquid DNA from these circular YACs was prepared and digested with *Sac*I, religated under dilute conditions that favor recircularization, and transformed into high-efficiency competent *E. coli* (Fig. 4). The resulting plasmids were shown to contain sequences at the vector–insert junction by comparison of sequences obtained by other methods. Rescue of each end individually was also accomplished by digestion with the enzymes listed under Materials and Methods.

Restriction Analysis of the YACs

The three YACs that were the minimal set necessary to define the contig were subjected to restriction mapping using a battery of rare-cutting restriction endonucleases. The results are shown in Fig. 5. The CpG islands are readily visible as regions of about 5 kb that contain many of the enzyme sites tested for. Comparison of the published genomic map (Fountain *et al.*, 1989a) indicates that many of the sites present in the YACs between the CpG islands must be methylated in na-

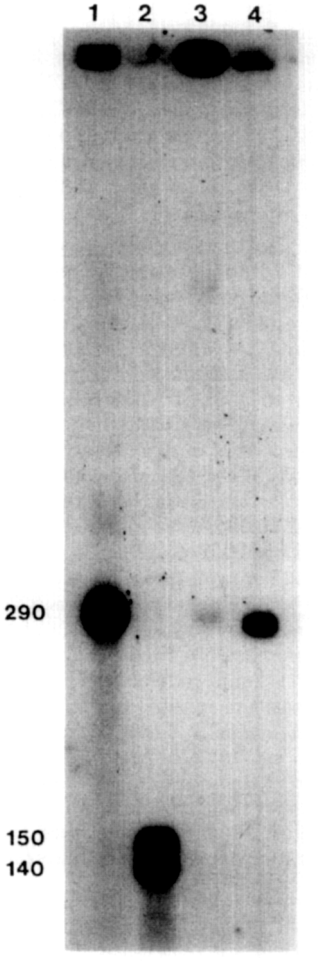

FIG. 2. Proof of YAC circularization. Clone yA43A9 was circularized and the linear and circular forms were analyzed by field-inversion gel electrophoresis. Lane **1** shows linear yA43A9 probed with pBR322. It migrates at 290 kb. Lane **2** shows the same linear YAC digested with *Mlu*I, which cuts at a single unique site in the insert. Lane **3** shows uncut circular yA43A9. Lane 4 shows *Mlu*I cut circular yA43A9. The faint hybridization at 290 kb in lane **3** can be accounted for by YAC linearized by shearing during agarose block formation.

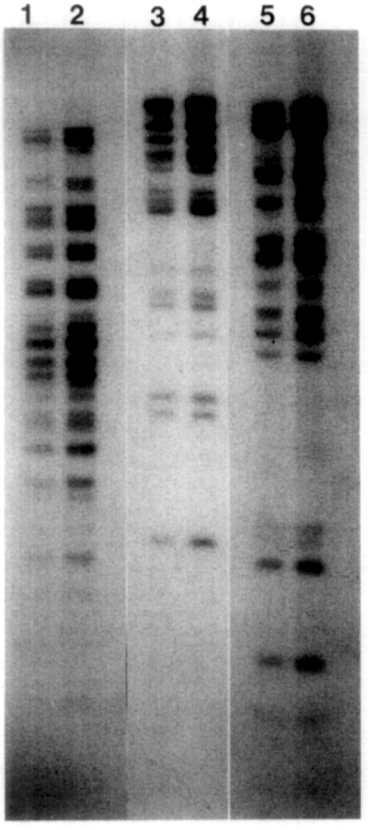

FIG. 3. Lack of gross rearrangement of insert DNA during YAC circularization. Conventional Southern blot *Alu* repeat "fingerprints" of total yeast DNA cut with *Eco*RI lanes **1** and **2**; *Eco*RV, lanes **3** and **4**; or *Pst*I, lanes **5** and **6**. All three enzymes have identical sites within the vector sequence in both linear and circular YAC constructs. The first lane in each pair consists of linear yA43A9 DNA and the second of circular yA43A9.

tive genomic DNA and therefore are not cleavable. Otherwise the native genomic and YAC maps look very similar.

Construction of an NF1 YAC

Since none of the 11 YACs isolated in this region contained the entire *NF1* gene, an *NF1* YAC was constructed by homologous recombination in yeast according to the method of Green and Olson (1990b). Clone yD8F4, in host strain AB1380 (MAT a), was crossed with the host strain AB1610 (MAT α) and sporulated, and tetrads were dissected. Haploid spores of the α mating type that also contained yD8F4 were selected. This new clone was named yD8F4α. Clone yD8F4α was crossed with yA43A9, and diploids were selected and sporulated. Haploids were picked by random spore analysis (Treco, 1989). A PCR assay to find a spore containing sequences across the *NF1* gene is shown in Fig. 6. One spore in 33 initially tested met the criterion of being a recombinant between yA43A9 and yD8F4α. CHEF gel analysis of the resulting clone indicated that it contained a new YAC of approximately 780 kb (Fig. 7). It must be noted that this YAC, like the parental yD8F4, is chimeric. However, characterization of the recombinant YAC using *NF1* cDNAs as hybridization probes indicated that the entire coding region exists in an unrearranged form at this level of detection.

NF1 Pseudogene

The eight YAC clones from the pAN1 screening that did not map to chromosome 17 were subjected to further analysis. The trivial explanation was that the end clone pAN1 was not from chromosome 17, but this was unlikely because the probe used hybridizes to a band on a chromosome 17-containing hybrid, and the PCR primers derived from the probe amplify DNA from a chromosome 17-specific hybrid, but not from the rodent background (data not shown). However, using a somatic

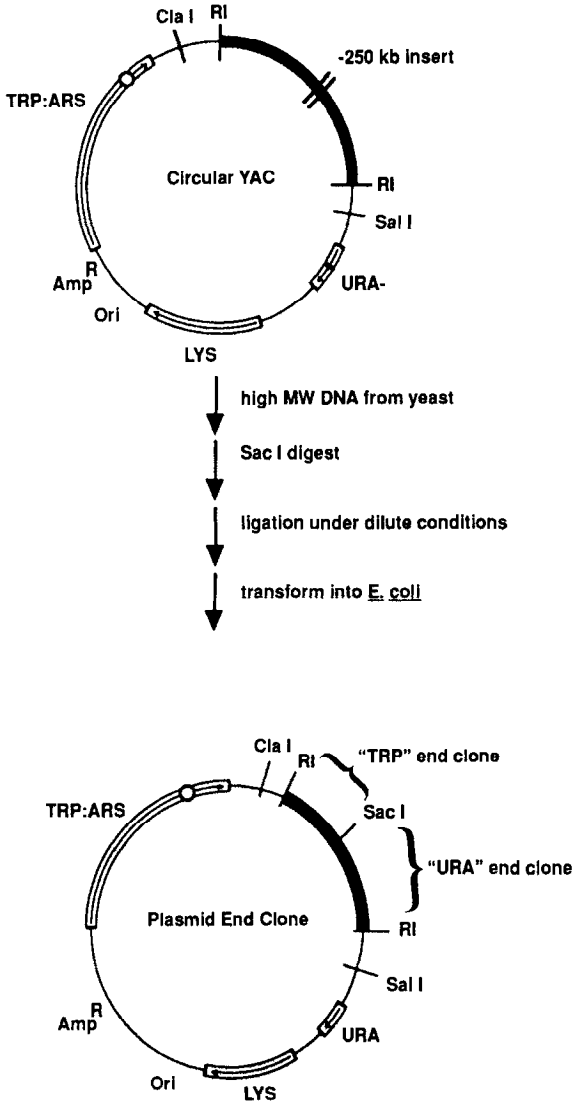

FIG. 4. Plasmid rescue of YAC ends from circular YAC constructs. To rescue both ends of the YAC, high-molecular-weight DNA is prepared from the circular YAC clone. This is digested with *Sac*I, which does not cut within the circular vector sequences. After ligation under dilute conditions, the DNA is transformed into *E. coli*. The resulting plasmid contains both ends of the YAC ligated at a unique *Sac*I site.

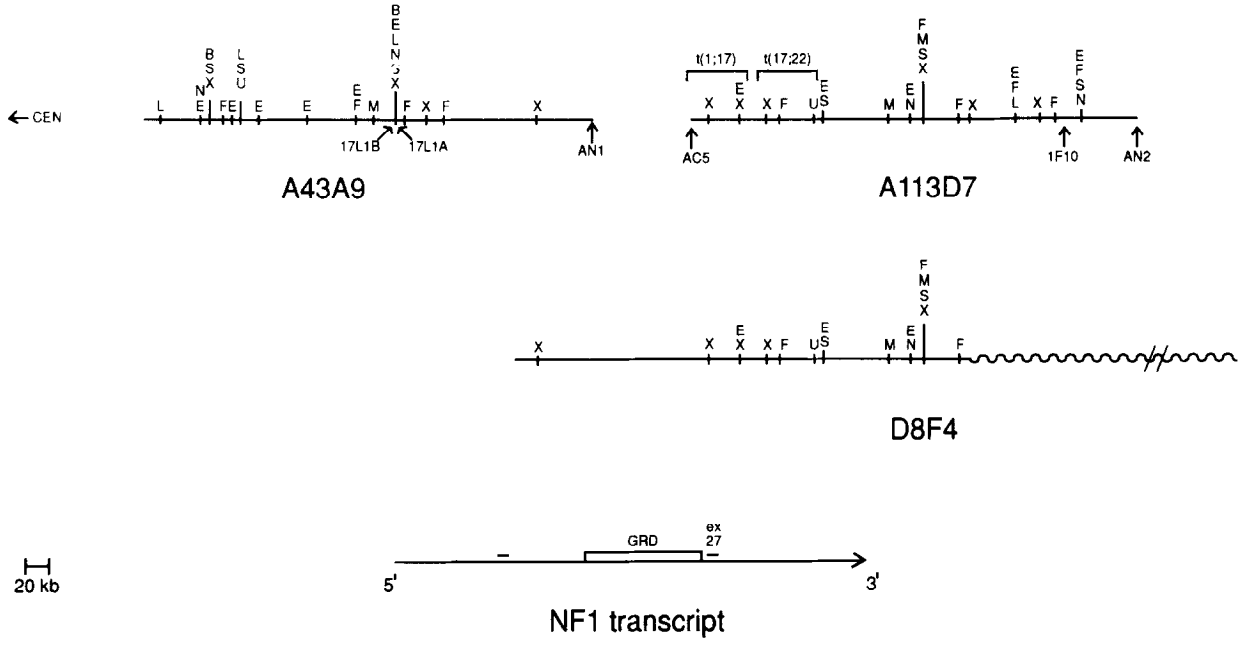

FIG. 5. An *NF1* YAC contig. The three clones necessary to define the contig are shown. The YACs were restriction mapped by field-inversion gel electrophoresis and indirect end labeling using YAC vector arm probes. Clone yD8F4 is chimeric in the regions shown by the wavy line. The key for restriction sites is B, *Bss*HII; E, *Eag*I; F, *Sfi*I; L, *Sal*I; M, *Mlu*I; N, *Not*I; S, *Sac*II; U, *Nru*I; and X, *Xho*I. The positions of probes 17L1A and B (Fountain *et al.*, 1989a), c1F10 (O'Connell *et al.*, 1989), and YAC ends AN1, AN2, and AC5 are shown. The extent and direction of the *NF1* transcript are shown (Marchuk *et al.*, 1991). The GAP-related domain (Xu *et al.*, 1990) is shown with a box (GRD) and the approximate locations of the NF1 translocation breakpoints are shown. As an aid to the discussion of the *NF1* pseudogene, the bars indicate the positions of the two *NF1* regions shown in Fig. 10 with homology to the pseudogene. One corresponds to exon 27 and the other to a region upstream from the GAP-related domain.

cell hybrid mapping panel, this probe showed additional homology on chromosomes other than 17. This probe was known to consist only of intron sequence. A search for homologous loci was subsequently attempted using a genomic clone containing a nearby *NF1* exon. Using the rescued YAC end probe pAC5, which contains an exon (corresponding to exon 27; R. White, personal communication) just downstream of the GAP-related domain of *NF1* (Xu *et al.*, 1990), four other loci that contained regions of homology were discovered (Fig. 8). Using a somatic cell hybrid mapping panel, these were mapped to chromosomes 14 (two bands), 15, and 22. The presence of a similar-size cross-hybridizing band of 4.3 kb located on chromosomes 14 and 22 (Fig. 8) was confirmed using monochromosome-specific cell hybrids (data not shown).

The YACs were then probed with an *NF1* cDNA clone containing exon 27 (Fig. 9). Comparison of the Southern blot patterns using a number of probes against the somatic cell hybrids with that of the YACs indicated that this other locus resided on either chromosome 14 or 22. In every case tested, identical fragments were seen in hybrids containing chromosomes 14 and 22 and the YACs. Comparison of sequenced regions from somatic cell hybrids containing either chromosome 14 or 22 indicated that the sequence of the YACs more closely resembled that of chromosome 14, although some differences were seen between the sequence derived from the YAC and that derived from the chromosome. Since these differences may be due to polymorphism, it is uncertain whether the eight YAC clones derive from chromosome 14 or 22, or some from each.

To determine whether the locus on the YACs contains a functional gene, further analysis of the cross-hybridizing regions was undertaken. Two regions with homology to *NF1* exons were subcloned from one of the YACs and sequenced. One region corresponds to exon 27 of *NF1*, and the other is from coding sequence 5′ of the GAP-related domain in a region where we do not have complete genomic sequence information (see Fig. 5). In both regions, the YAC sequence was highly homologous to the *NF1* exon but contained base pair changes (Fig. 10). More significantly, the sequenced regions contained either a single base pair deletion (Fig. 10A) or an insertion (Fig. 10B). For the region shown in Fig. 10B, three stop codons were encountered, two created downstream from the single base pair insertion. The homology with the *NF1* gene extends into the introns on either side for the region shown in Fig. 10A (we do not know the intron sequence for the region of the *NF1* gene shown in 10B), and the homology is as strong in the limited amount of intron sequence obtained as in the exon (data not shown). We interpret this to mean that this locus represents a nonprocessed pseudogene.

The entire complement of *NF1* cDNA sequence (Marchuk *et al.*, 1991) was then used as a probe to survey the pseudogene present on the YACs for other regions of homology to *NF1* coding sequences. No significant

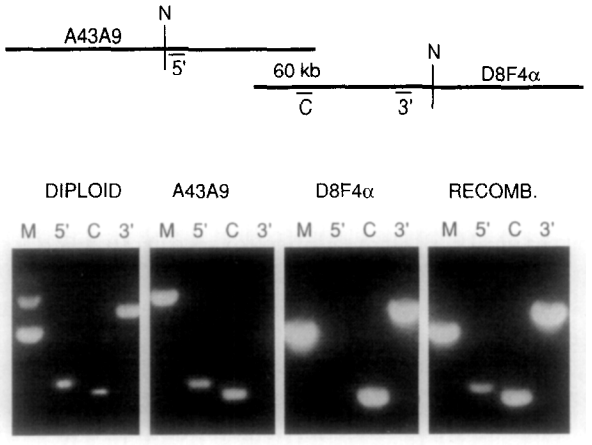

FIG. 6. Construction of an *NF1* YAC by homologous recombination *in vivo*. Parental clones yA43A9 and an α mating type derivative of yD8F4 were mated and a diploid was selected for. The diploid was sporulated and colonies from individual spores were subjected to the PCR analysis shown. Four primer pairs were used, one (set of three) to determine mating type (Huxley *et al.*, 1990), designated M, and probes spanning the 5′, central (C), and 3′ portions of the *NF1* gene. A haploid recombinant was found that contains all three regions of the *NF1* gene by PCR analysis. The *NF1* PCR primer pairs used are as follows: 5′ primer pair: CGGATCCTCCCCCGCGGCTGCCTCAGGCTCTG and CGGATCCCAGGTCACTCATCCCCATTTCCAAG; the central primer pair: GGCATGAAAGTCTGAAGTCTAATCTC and GAGTTTTATCTGGTACTAGAAGCATAGCTG; the 3′ primer pair: CATCGGATCCATATCTGTTTTATCATCAGGAGG and CATCGGATCCAAGTAAAATGGAGAAAGGAACTGG. The 5′ and 3′ primer pairs contain *Bam*HI cloning sites at the 5′ end of the primers.

cross-hybridizing bands were visible, other than those previously seen, using this broad survey approach with large cDNA inserts. To more carefully probe the pseudogene for homologous regions that might have been missed, the *NF1* exons with known intron borders (Cawthon *et al.*, 1990; D. Marchuk and A. Martin-Gallardo, unpublished) were individually amplified by PCR and used as probes against the YACs (corresponding to exons 26–39; R. White, personal communication). Of these, only exon 27 (corresponding to the region shown in Fig. 10A) showed any cross-hybridization. We have also been unable to find any homology to the GAP-related domain (Xu *et al.*, 1990) within the pseudogene present on the YACs, despite the fact that homologous sequences are present for *NF1* sequence on either side of this domain. It is unclear therefore whether these regions of the pseudogene have been deleted, have been rearranged, or have diverged beyond the point of cross-hybridization.

DISCUSSION

The 11 *NF1* YACs from 17q11.2 form a contig of almost 700 kb. No attempt was made to extend the contig further in either direction since in the course of this work the *NF1* gene was identified (Cawthon *et al.*, 1990; Viskochil *et al.*, 1990; Wallace *et al.*, 1990b) and found to be within the boundaries of the contig (Marchuk *et al.*, 1991). A total of 11 YAC clones were isolated from the library by screening with three different probes. Two of these 11 (yA113D7 and yA113D11) are definitely sib clones from near-adjacent positions in the gridded library since they show identical sizes, fingerprints, and rescued ends. Four of the 10 independent YACs have been shown to be chimeric, which is within the range that others have found with this YAC library (Gaensler *et al.*, 1991; Green and Olson, 1990b; Silverman *et al.*, 1989, 1991). However, this is a minimum estimate, since four of the YACs isolated from the c1F10 screen were not extensively characterized for chimerism because by *Alu* fingerprint analysis they appeared to be completely contained within another YAC.

Comparisons with the published restriction map (Fountain *et al.*, 1989a) and the restriction map of the YACs show substantial similarity. The CpG islands are clearly visible on both maps as regions where multiple methylation-sensitive restriction endonucleases cut. The YAC clones show some sites that are not present in the genomic maps. These are most likely due to sites that are authentic but methylated in native genomic

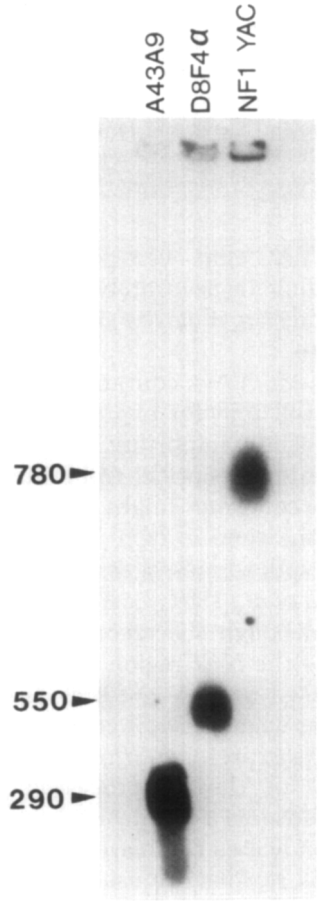

FIG. 7. CHEF gel analysis of an *NF1* YAC recombinant. Agarose blocks from the parental strains yA43A9 and yD8F4, along with the recombinant YAC shown in Fig. 6, were subjected to CHEF gel analysis and probed with pBR322. The extent of the overlap between yA43A9 and yD8F4 is approximately 60 kb, resulting in a 780-kb recombinant.

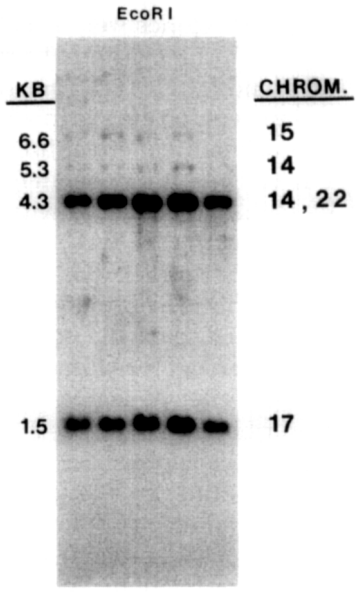

FIG. 8. *Chromosomal location of the NF1-related sequences.* Probe pAC5, a 1.5-kb *Eco*RI fragment rescued from one of the ends of yA113D7, contains most of the exon 27 sequence (base pairs 4895–4983 of GenBank M82814) plus approximately 1.4 kb of 3′ flanking intron. This probe was used against a panel of 20 unrelated individuals (five shown here) to ensure that RFLPs were not being scored. Chromosomal locations for each cross-hybridizing band were determined by comparison with a Southern blot using a somatic cell hybrid mapping panel (data not shown). The loci corresponding to the to 5.3- and 6.6-kb bands are less intense due to lower homology to the intron sequence of the probe. The YACs shown in Fig. 9 show the 4.3-kb band, identical to the closely related loci on chromosomes 14 and 22.

DNA. Since the YAC maps were generated from defined end points and could be accurately assigned without the need for double digests, it can be argued that these maps are more accurate.

The recombinant YAC containing the entire *NF1* locus is a potential tool for expression studies of this gene. This is especially important if alternative splicing generates different polypeptides from this locus, because any given cDNA construct might not reflect the entire contribution of the gene *in vivo*. There have been two such alternative spliced messages reported thus far (Xu *et al.*, 1990; Nishi *et al.*, 1991; L. Andersen, unpublished observations). Although the recombinant YAC is chimeric, it contains the *NF1* region intact and should express all the known mRNA isoforms arising from this locus. It is unclear what effect, if any, the chimeric DNA might have on the expression of the gene.

The discovery of the nonprocessed *NF1* pseudogene(s) on chromosome 14 and/or 22 contained within the AN1 YACs indicates that care must be taken when these homologous regions are used as probes to screen for authentic *NF1* mutations in patients. This is complicated by the fact that the locus at chromosome 15 has also been determined to be a pseudogene (E. Legius, personal communication). In addition, there may be other loci with homology to *NF1* that have yet to be discovered, since we have made no attempt to screen for all related sequences in the genome, other than to use exon 27 of the *NF1* gene as a probe.

It is possible that these and any other loci that exist are all pseudogenes and that they represent a reservoir of mutations that can be crossed into the *NF1* locus by interchromosomal gene conversion. This might partially explain the abnormally high mutation rate at this locus (Crowe *et al.*, 1956). Precedence for mutation arising as a result of gene conversion between a gene and a pseudogene exists with the steroid 21-hydroxylase gene and congenital adrenal hyperplasia (Harada *et al.*, 1987; Higashi *et al.*, 1988). However the 21-OHase gene and pseudogene are paired in a tandem repeat with the C4 complement gene and pseudogene. If gene conversion events with unlinked *NF1* pseudogenes contribute to mutation, a more complex explanation than simple mismatch pairing would have to be invoked to account for gene conver-

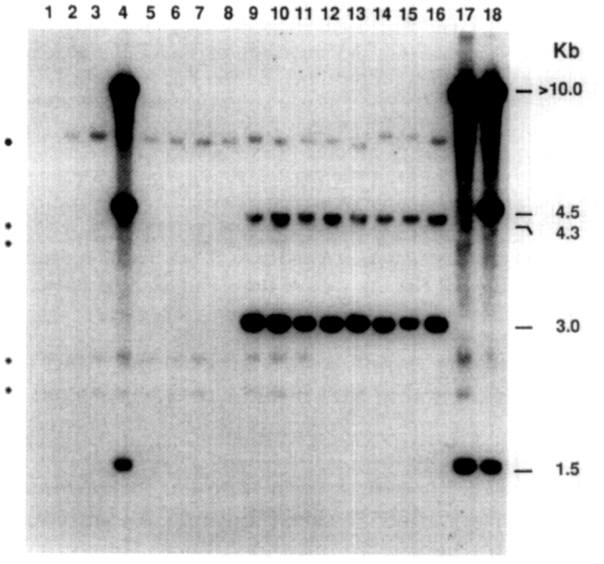

FIG. 9. A second set of YACs containing a locus with homology to the *NF1* gene. DNA from all the YACs derived from screenings of the Washington University YAC libraries were digested with *Eco*RI and probed with cDNA clone pFB5D (Marchuk *et al.*, 1991), which corresponds to nucleotides 3309–5488 of the *NF1* cDNA sequence of GenBank M82814. This clone contains the entire exon 27 of the *NF1* gene. The samples are as follows in numerical order; AB1380, yA43A9, yB62G2, yA113D7, yA140G3, yA220G3, yA140G10, y1F10Y1, yAN1, yA10C9, yB258G2, yA280G10, yA144H3, yA276B3, yB89C2, yA272H3, yB227C3, and yB164C5. Clones y1F10 and yAN1 were isolated from a neuroblastoma YAC library (Schneider *et al.*, 1991). Lane **1** contains yeast host strain AB1380. Lanes **2–8** and **17–18** contain YAC clones that map to chromosome 17. Strongly hybridizing bands of approximately 10, 4.5, and 1.5 kb are seen with this probe and are identical to those seen with a hybrid cell line containing only human chromosome 17. Clones in lanes **2, 3,** and **5–8** do not overlap with the probe and thus do not show these bands. Lanes **9–16** show YAC clones derived from a screen using clone pAN1, rescued from the end of yA43A9. All the AN1-derived YACs show hybridizing bands at 4.3 and 3.0 kb. These bands also appear in somatic cell hybrid containing only chromosome 14 or 22. Faint hybridization seen in lanes **2–18** *(marked with a dot) is due to the presence of plasmid sequence in the* probe that cross-hybridizes to YAC vector arm sequence and migrates on the telomere-containing fragment of variable size within individual clones. Faint hybridization seen in all lanes (marked with asterisks) is due to homology to yeast sequences.

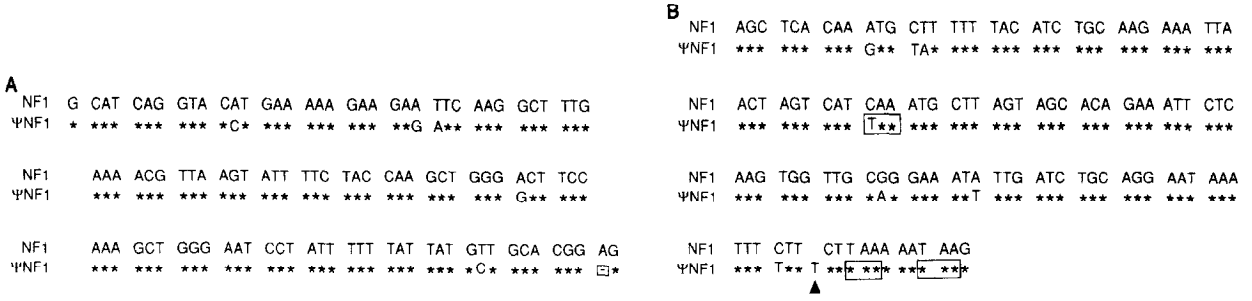

FIG. 10. Comparison of *NF1* and *NF1* pseudogene sequences in two homologous regions. The *NF1* sequence is shown on top with the pseudogene below. Identical bases are indicated by an asterisk. (**A**) This region corresponds to exon 27 and contains nucleotides 4873–4983 according to the *NF1* cDNA sequence, GenBank Accession No. M82814 (Marchuk *et al.*, 1991). The splice junctions were previously determined from the yA113D7 YAC end clone pAC5 (D. Marchuk and A. Martin-Gallardo, unpublished). The pseudogene contains a deletion of 1 base, shown with a box. (**B**) This region corresponds to nucleotides 1931–2056 of the *NF1* cDNA sequence as referenced above. A substitution (C to T) results in an in-frame stop codon, shown with a box. An insertion downstream of this region, shown with an arrowhead, results in a frameshift and the creation of two more stop codons, shown in boxes.

sion events. However, meiotic gene conversion between unlinked genes does occur in yeast (Jinks-Robertson and Petes, 1985).

The presence of unlinked *NF1*-related loci in the genome suggests the possibility that some of these may represent functional genes. If so, they may code for yet undiscovered tumor-suppressor genes related to *NF1*. In this regard it is interesting that the type 2 neurofibromatosis gene (NF2) maps to chromosome 22. These loci are currently being investigated for the presence of transcribed sequences.

ACKNOWLEDGMENTS

We thank D. Gutmann, S. Wilson-Gunn, and other members of the NF group in Michigan for helpful comments during the course of this work. A. Saulino, S. Hayes, and L. Gross provided technical assistance. M. Greenberg and E. Green provided expert advise on all aspects of yeast genetics. D. Hartl and D. Garza provided pPM680. R. White and P. O'Connell provided clone c1F10. A. Martin-Gallardo and C. Ventor provided sequence of clones pEH1 and pAC5. We thank S. Schneider and G. Brodeur for screening their neuroblastoma YAC library. This work was supported by NIH Grant NS23410 (F.S.C.) and NIH Fellowship HG00018 (D.A.M.).

REFERENCES

Barker, D., Wright, E., Nguyen, K., Cannon, L., Fain, P., Goldgar, D., Bishop, D. T., Carey, J., Baty, B., Kivlin, J., Willard, H., Waye, J. S., Greig, G., Leinwand, L., Nakamura, Y., O'Connell, P., Leppert, M., Lalouel, J.-M., White, R., and Skolnick, M. (1987). Gene for von Recklinghausen neurofibromatosis is in the pericentromeric region of chromosome 17. *Science* **236:** 1100–1102.

Brownstein, B. H., Silverman, G. A., Little, R. D., Burke, D. T., Korsmeyer, S. J., Schlessinger, D., and Olson, M. V. (1989). Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* **244:** 1348–1351.

Burgers, P. M., and Percival K. J. (1987). Transformation of yeast spheroplasts without cell fusion. *Anal. Biochem.* **163:** 391–397.

Burke, D. T., Carle, G. G., and Olson, M. V. (1987). Cloning of large segments of exogeneous DNA into yeast by means of artificial chromosome vectors. *Science* **236:** 806–812.

Carle, G. F., Frank, M., and Olson, M. V. (1986). Electrophonic separations of large DNA molecules by periodic inversion of the electric field. *Science* **232:** 65–68.

Cawthon, R., Weiss, R., Xu, G., Viskochil, D., Culver, M., Stephens, J., Robertson, M., Dunn, D., Gesteland, R., O'Connell, P., and White, R. (1990). A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomeic structure, and point mutations. *Cell* **62:** 193–201.

Chu, G. D., Vollrath, D., and Davis, R. W. (1986). Separation of large DNA molecules by contour clamped homogeneous electric fields. *Science* **234:** 1582–1585.

Crowe, F., Schull, W., and Neel, J. (1956). "A Clinical, Pathological, and Genetic Study of Multiple Neurofibromatosis," Thomas, Springfield, IL.

Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T., and Schmit, C. W. (1981). Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol.* **151:** 17–33.

Drumm, M. L., Smith, C. L., Dean, M., Cole, J. L., Iannuzzi, M. C., and Collins, F. S. (1988). Physical mapping of the cystic fibrosis region by pulsed-field gel electrophoresis. *Genomics* **2:** 346–354.

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., Stephines, K., and Collins, F. S. (1989a). Physical mapping of the van Recklinghausen neurofibromatosis region on chromosome 17. *Am. J. Hum. Genet.* **44:** 58–67.

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 break point in neurofibromatosis. *Science* **244:** 1085–1087.

Gaensler, K. M., Burmeister, M., Brownstein, B. H., Taillon-Miller, P., and Myers, R. M. (1991). Physical mapping of yeast artificial chromosomes containing sequences from the human β-globin gene region. *Genomics* **10:** 976–984.

Garza, D., Ajioka, J. W., Burke, D. T., and Hartl, D. L. (1989). Mapping the Drosophila genome with yeast artificial chromosomes. *Science* **246:** 641–646.

Goldgar, D. E., Green, P., Parry, D. M., and Mulvihill, J. J. (1989) Multipoint linkage analysis in neurofibromatosis type 1: An international collaboration. *Am. J. Hum. Genet.* **44:** 6–12.

Green, E. D., and Olson, M. V. (1990a). Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* **87:** 1213–1217.

Green, E. D., and Olson, M. V. (1990b). Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: A model for human genome mapping. *Science* **250:** 94–98.

Harada, F., Kimura, A., Iwanaga, T., Shimozawa, K., Yata, J., and Sasazuki, T. (1987). Gene conversion-like events cause steroid

21-hydroxylase deficiency in congenital adrenal hyperplasia. *Proc. Natl. Acad. Sci. USA* **84:** 8091–8094.

Higashi, Y., Tanae, A., Inoue, H., Hiromasa, T., and Fujii-Kurijama, Y. (1988). Aberrant splicing and missense mutation causing steroid 21-hydroxylase (P-450c21) deficiency in humans: Possible gene conversion products. *Proc. Natl. Acad. Sci. USA* **85:** 7486–7490.

Huxley, C., Green, E. D., and Dunham, I. (1990). Rapid assessment of *S. cerevisiae* mating type by PCR. *Trends Genet.* **6:** 236.

Jinks-Robertson, S., and Petes, T. D. (1985). High-frequency meiotic gene conversion between repeated genes on nonhomologous chromosomes in yeast. *Proc. Natl. Acad. Sci. USA* **82:** 3350–3354.

Marchuk, D. A., Saulino, A., Tavakkol, R., Swaroop, M., Wallace, M., Andersen, L., Mitchell, A., Gutmann, D., Boguski, M., and Collins, F. (1991). cDNA cloning of the type 1 neurofibromatosis gene: Complete sequence of the NF1 gene product. *Genomics* **11:** 931–940.

Nishi, T., Lee, P. S., Oka, K., Levin, V. A., Tanase, S., Morino, Y., and Saya, H., (1991). Differential expression of two types of the neurofibromatosis type 1 (*NF1*) gene transcripts related to neuronal differentiation. *Oncogene* **6:** 1555–1559.

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.

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

Riccardi, V. M., and Eichner, J. E. (1986). "Neurofibromatosis: Phenotype, Natural History, and Pathogenesis," Johns Hopkins Univ. Press, Baltimore, MD.

Schneider, S. S., Zehnbauer, B. A., Vogelstein, B., and Brodeur, G. M. (1991). Yeast artificial chromosome (YAC) vector cloning of the MYCN amplified domain in neuroblastomas. *Prog. Clin. Biol. Res.* **366:** 71–76.

Seizinger, B. R., Rouleau, G. A., Ozelius, L. J., Lane, A. H., Faryniarz, A. G., Chao, M. V., Huson, S., Korf, B. R., Parry, D. M., Pericak-Vance, M. A., Collins, F. S., Hobbs, W. J., Falcone, B. G., Iannazzi, J. A., Roy, J. C., St. George-Hyslop, P. H., Tanzi, R. E., Bothwell, M. A., Upadhyaya, M., Harper, P., Goldstein, A. E., Hoover, D. L., Bader, J. L., Breakefield, X. O., and Gusella, J. (1987). Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor gene. *Cell* **49:** 589–594.

Sherman, F., Fink, G. R., and Jicks, J. B. (1986). "Laboratory Course Manual for Methods in Yeast Genetics," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Silverman, G. A., Ye, R. D., Pollock, K. M., Sadler, J. E., and Korsmeyer, S. J. (1989). Use of yeast artificial chromosome clones for mapping and walking within human chromosome segment 18q21.3. *Proc. Natl. Acad. Sci. USA* **86:** 7485–7489.

Silverman, G. A., Jockel, J. I., Domer, P. H., Mohr, R. M., Taillon-Miller, P., and Korsmeyer, S. J. (1991). Yeast artificial chromosome cloning of a two-megabase-size contig within chromosomal band 18q21 establishes physical linkage between BCL2 and plasminogen activator inhibitor type-2. *Genomics* **9:** 219–228.

Treco, D. A. (1989). Growth and manipulation of yeast. *In* "Current Protocols in Molecular Biology" (F. A. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, Eds.), pp. 13.2.1–13.2.10, Greene and Wiley-Interscience, New York.

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.

Wallace, M. R., Andersen, L. B., Fountain, J. W., Odeh, H. M., Viskochil, D., Marchuk, D. A., O'Connell, P., White, R., and Collins, F. S. (1990a). A chromosome jump crosses a translocation breakpoint in the von Recklinghausen neurofibromatosis region. *Genes Chromosomes Cancer* **2:** 271–277.

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. (1990b). Type 1 neurofibromatosis gene: Identification of a large transcript disrupted in three NF1 patients. *Science* **249:** 181–186.

Xu, G., O'Connell, P., Viskochil, D., Cawthon, R., Robertson, M., Culver, M., Dunn, D., Stevens, J., Gesteland, R., White, R., and Weiss, R. (1990). The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* **62:** 599–608.