ARTICLE

Mechanism of spreading of the highly related neurofibromatosis type 1 (*NF1*) pseudogenes on chromosomes 2, 14 and 22

Mirjam Luijten[1], YingPing Wang[2], Blaine T Smith[2], Andries Westerveld[1], Luc J Smink[3], Ian Dunham[3], Bruce A Roe[2] and Theo JM Hulsebos[1]

[1]*Department of Human Genetics, Academic Medical Center, University of Amsterdam, The Netherlands;*
[2]*Department of Chemistry and Biochemistry, University of Oklahoma, Norman, OK, USA;* [3]*Sanger Centre, Wellcome Trust Genome Campus, Hinxton Hall, Cambridge, UK*

Neurofibromatosis type 1 (NF1) is a frequent hereditary disorder that involves tissues derived from the embryonic neural crest. Besides the functional gene on chromosome arm 17q, *NF1*-related sequences (pseudogenes) are present on a number of chromosomes including 2, 12, 14, 15, 18, 21, and 22. We elucidated the complete nucleotide sequence of the *NF1* pseudogene on chromosome 22. Only the middle part of the functional gene but not exons 21–27a, encoding the functionally important GAP-related domain of the NF1 protein, is presented in this pseudogene. In addition to the two known *NF1* pseudogenes on chromosome 14 we identified two novel variants. A phylogenetic tree was constructed, from which we concluded that the *NF1* pseudogenes on chromosomes 2, 14, and 22 are closely related to each other. Clones containing one of these pseudogenes cross-hybridised with the other pseudogenes in this subset, but did not reveal any *in situ* hybridisation with the functional *NF1* gene or with *NF1* pseudogenes on other chromosomes. This suggests that their hybridisation specificity is mainly determined by homologous sequences flanking the pseudogenes. Strong support for this concept was obtained by sequence analysis of the flanking regions, which revealed more than 95% homology. We hypothesise that during evolution this subset of *NF1* pseudogenes initially arose by duplication and transposition of the middle part of the functional *NF1* gene to chromosome 2. Subsequently, a much larger fragment, including flanking sequences, was duplicated and gave rise to the current *NF1* pseudogene copies on chromosomes 14 and 22. *European Journal of Human Genetics* (2000) **8**, 209–214.

Keywords: neurofibromatosis type 1; pseudogene; sequence analysis; evolution

Introduction

Neurofibromatosis type 1 (NF1) is a frequent autosomal disorder with an incidence of approximately 1 in 3000 births. The disease involves tissues derived from the embryonic neural crest and the most important clinical features of the disorder are multiple neurofibromas, café-au-lait spots and Lisch nodules of the iris. The *NF1* gene, located on chromosome 17q11.2, consists of 59 exons distributed over 350 kb of genomic DNA and encodes a protein of 2818 amino acids, called neurofibromin.[1] Neurofibromin contains a central domain showing homology to the catalytic domain of mammalian GTPase-activating proteins (GAPs).[2] The GAP-related domain (GRD), encoded by exons 21–27a, stimulates the intrinsic GTPase-activity of the proto-oncogene *ras* p21. By converting *ras* p21 from the GTP- to the GDP-bound form, the activity of *ras* p21 is down-regulated.

In recent years, numerous *NF1*-related sequences (pseudogenes) have been identified on various chromosomes, including 2, 12, 14, 15, 18, 21, and 22.[3–11] Most of the *NF1* pseudogenes have been mapped to pericentromeric regions.

The chromosome 2 *NF1* pseudogene has been localised in region 2q21, which is known to contain the remnant of an ancestral centromere.[12] *NF1*-related sequences are classified as non-processed pseudogenes, as partial sequence analysis revealed a conserved exon–intron organisation. On chromosomes 14 and 15, respectively two and three different *NF1* pseudogene copies have been found.[3–5,7,10,11]

In this study, we investigated the mechanism of spreading of *NF1* pseudogenes, in particular those on chromosomes 2, 14, and 22. In a previous report,[9] we showed that cosmids containing the chromosome 22 *NF1* pseudogene only cross-hybridise with the *NF1* pseudogene regions on chromosomes 2 and 14. No hybridisation was detected with *NF1*-related loci on other chromosomes or with the functional *NF1* gene. To explain this apparent lack of cross-hybridisation, we suggested that the *NF1*-related loci on chromosomes 2, 14, and 22 form a subset of *NF1* pseudogenes that are highly related to each other. To demonstrate this, we determined exactly which exons are represented in the *NF1* pseudogenes on chromosomes 2, 14, and 22 and we established their evolutionary relationship. In addition, we sequenced regions flanking the *NF1* pseudogenes on these chromosomes and show that these are highly homologous to each other.

Materials and methods
Southern blot analysis
Exons of the functional *NF1* gene were amplified by PCR using as template DNA extracted from a human × mouse hybrid cell line with chromosome 17 as the sole human contribution (GM10498B, NIGMS Human Genetic Mutant Cell Repository, Camden, NJ, USA). PCR primers and conditions were according to Li *et al.*[1] PCR products were recovered from low melting temperature agarose, labelled with α-[³²P]dCTP and hybridised to Southern blots containing digests with *Hin*dIII of total human DNA, and of DNA extracted from the chromosome 17-only hybrid cell line, from a chromosome 22-only human × mouse hybrid cell line (PgMe25NU), from a chromosome 14-only human × mouse cell line (Wegroth B3), from a chromosome 2-only human × hamster cell line (GM10826B, NIGMS Human Genetic Mutant Cell Repository, Camden, NJ, USA), from a mouse-only cell line (Wehi-TG) and from a hamster-only cell line (A3). Signal densities were quantified by laser scanning densitometry (Phosphorimager, Molecular Dynamics, Sunnyvale, CA, USA) using Imagequant Software v.30. Cell lines PgMe25NU, Wegroth B3 and Wehi-TG originated from and were kindly provided by A Geurts van Kessel, University of Nijmegen, Nijmegen, The Netherlands.

Screening of libraries
A chromosome 2 cosmid library, LL02NC02 (Lawrence Livermore National Laboratory, Human Genome Center, Livermore, CA, USA; UK HGMP Resource Centre, Hinxton, Cambridge, UK), was screened for *NF1*-related sequences using PCR products of exons 15, 17, and 19a of the functional *NF1* gene as probes as described previously.[13] A Human Female PAC Library, RPCI6 no. 709 (Pieter de Jong, Roswell
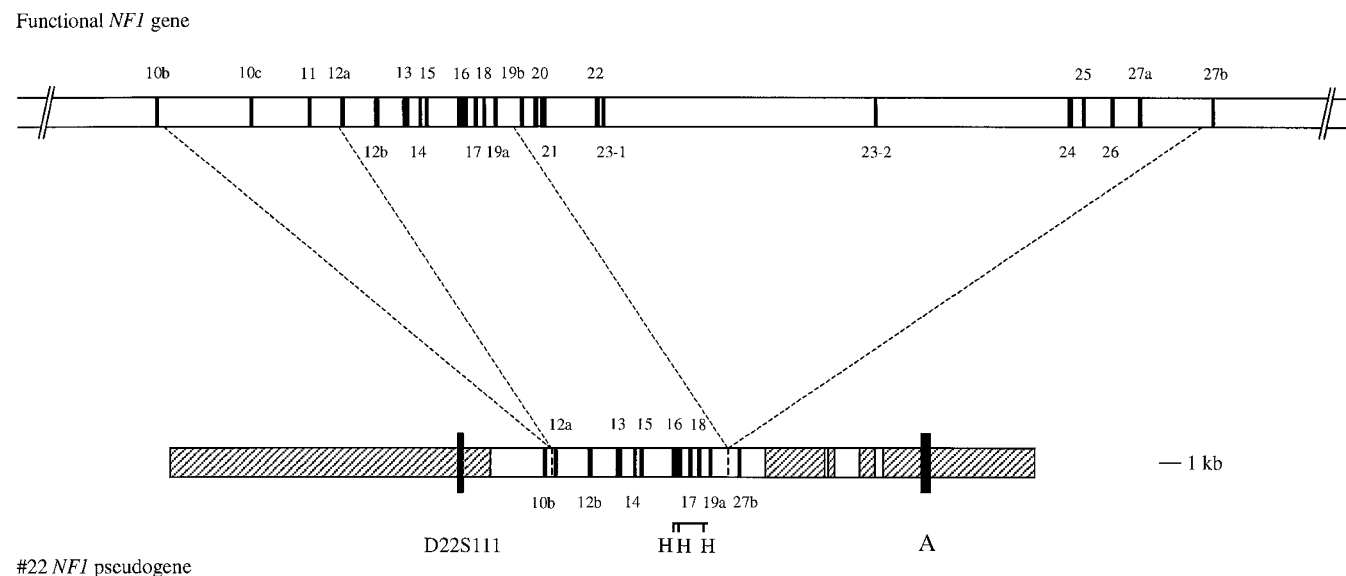
Figure 1 Genomic organization of the *NF1* pseudogene on chromosome 22 in cosmid 29A9 (GenBank Accession No. AC005374) and the corresponding part of the functional *NF1* gene. Exons and introns are denoted by filled and open boxes, respectively. Hatched boxes indicate sequences in cosmid 29A9 that are not related to those in the functional *NF1* gene. Cosmid 29A9 has an insert of 41958 bp. The *NF1*-homologous region starts at position 15686 and ends, except for a few additional regions of homology to intron 27b, at position 28934. The positions of region D22S111 (from 14122 to 14381) and of region A (from 36421 to 36848) are indicated. H: *Hin*dIII restriction sites in exon 16 and intron 18 (see text for details).

Park Cancer Institute; Resource Center of the German Human Genome Project at the Max-Planck-Institut for Molecular Genetics), was screened in an analogous way.

PCR amplification and DNA sequencing

Primer sequences for amplification of region A (see Figure 1) were as follows: forward, 5'-CTGTTCTTCAGTTAGAGCTTAT-TACTC-3' and reverse, 5'-CTCACATACATTATCTCATTTCA-TACT-3'. Both this region and marker D22S111 were amplified using standard PCR conditions.

Sequencing of cosmids 29A9, 29D1 and 127F6[9] was performed according to procedures described previously.[14] The derived sequences were compared to the sequence of the *NF1* locus[15] and the cDNA sequence according to the NNFF International NF1 Genetics Analysis Consortium (http://www.nf.org/nf1gene/nf1gene.cDNAtext.html). Selected regions in and flanking the *NF1* pseudogenes on chromosomes 2 and 14 were sequenced on an Applied Biosystems Model 377 automated sequencer using the Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer, Foster City, CA, USA).

Phylogenetic analysis

Phylogenetic analyses were performed on the exon-related sequences of the *NF1* pseudogenes on chromosomes 2, 14 and 22 and the corresponding sequences of the functional *NF1* gene using the PHYLIP package (http://evolution.genetics.washington.edu/phylip.html). The maximum parsimony tree was obtained by branch-and-bound algorithm and contiguous gaps were recoded as a single character prior to analysis. For the neighbour-joining method, the statistical support for each branch was determined in terms of branch length and bootstrap analysis, using 1000 iterations.

Results

In a previous study, we isolated cosmids that contain *NF1*-related sequences on chromosome 22.[9] FISH analysis with these cosmids using high stringency conditions not only showed a hybridisation signal in the pericentromeric region of chromosome arm 22q, but also in the pericentromeric region of chromosome arm 14q and in region 2q21, the latter known to contain the remnant of an ancestral centromere. As these cosmids did not hybridise to other *NF1*-related loci or to the functional gene on chromosome arm 17q, we suggested that the *NF1*-related loci on chromosomes 22, 14, and 2 are related to each other and form a subset of *NF1* pseudogenes.

To determine which *NF1*-related exons are present in the pseudogene on chromosome 22, we hybridised PCR products of a large number of exons of the functional *NF1* gene to Southern blots with digests of DNA extracted from a chromosome 22-only human × mouse cell line. Human-specific hybridising fragments were only detected for exons 10b, 12a–19a, and 27b (data not shown). The same

results were obtained for cosmids 29A9, 29D1, and 127F6, suggesting that the entire chromosome 22 *NF1* pseudogene is contained within these cosmids.

Next we determined the complete nucleotide sequence of these cosmids (GenBank Accession Nos AC002471, AC003064, AC005374). Only the middle part of the functional *NF1* gene is represented in this pseudogene (Figure 1). Exons 10c–11 and 19b–27a and their intervening sequences are absent. As sequences of introns 19a and 27a of the functional *NF1* gene are known,[1,15] the transition between their homologous sequences in the pseudogene could be determined exactly. At the 5' end of the chromosome 22 *NF1* pseudogene, intron 10a is represented by its last 2594 bp. At the 3' end, the first 2934 bp of intron 27b are present. However, the representation of intron 27b is disrupted three times by sequences that are not related to those in the functional *NF1* gene, consisting of 2875 bp, 299 bp and 694 bp, respectively (Figure 1).
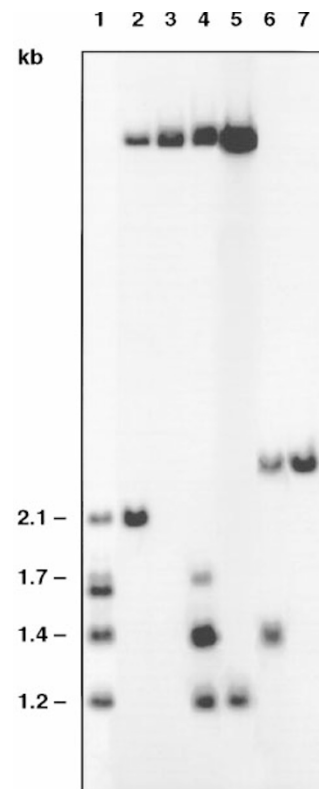
Figure 2 Southern blot of *Hin*dIII-digested total human DNA and DNAs extracted from human × mouse and human × hamster cell lines, hybridised with exons 16, 17 and 18 of the functional *NFI* gene. **Lane 1**: human lymphocytes; **Lane 2** chromosome 17-only human × mouse cell line GM10498B; **Lane 3** mouse-only cell line Wehi-TG; **Lane 4** chromosome 14-only human × mouse cell line Wegroth B3; **Lane 5** chromosome 22-only human × mouse cell line PgMe25Nu; **Lane 6** chromosome 2-only human × hamster cell line GM10826B; **Lane 7** hamster-only cell line A3.

For the *NF1* pseudogenes on chromosomes 2 and 14, Southern blot analysis was performed using chromosome 2-only and chromosome 14-only cell lines. Like for chromosome 22, human-specific hybridising fragments were only detected for exons 10b, 12a–19a, and 27b (data not shown). Except for exons 16, 17 and 18, all hybridising *Hin*dIII fragments observed with chromosomes 2 and 14 were of the same length as those with chromosome 22. Exons 16, 17, and 18 hybridised on chromosome 22 with a fragment of 1.2 kb, on chromosome 2 with a fragment of 1.4 kb and on chromosome 14 with fragments of 1.2, 1.4, and 1.7 kb (Figure 2). The three signals in the chromosome 14 lane were quantified. Assuming that the *Hin*dIII fragment of 1.7 kb represents one copy, we calculated that the fragments of 1.2 and 1.4 kb represent two and five copies, respectively. This suggests that eight copies of an *NF1* pseudogene, which may differ from each other, are present on this chromosome 14.

To obtain clones with sequences from the *NF1*-related locus on chromosome 2, we screened a chromosome 2 cosmid library (LL02NC02) with PCR products of exon 15, 17, and 19a. Only one positive cosmid was found, clone 180H4. All exons represented in the *NF1* pseudogene on chromosome 2 as well as the fusion of intron 19a to intron 27a were present in this cosmid and their sequences were determined (GenBank Accession Nos AF232292–AF232302).

Since a chromosome 14-specific library was not available, 15 clones containing *NF1* homologous loci were obtained by screening a human female PAC library (RPCI6). Upon Southern blot analysis with exons 16, 17, and 18, seven of these clones showed *Hin*dIII fragments with a length of 1.2, 1.4, or 1.7 kb. FISH analysis revealed that these clones all hybridise to the pericentromeric regions of chromosome arms 22q and 14q and to 2q21. In addition, the transition of intron 19a to intron 27a could be PCR-amplified from these

seven PACs. Sequence analysis showed that three clones were derived from chromosome 2. Two other clones carried sequences identical to those presented by Régnier *et al*[11] as φ*NF1*-14A (GenBank Accession Nos AF232248–AF232258) and φ*NF1*-14B (GenBank Accession Nos AF232259–AF232269). A third clone, which we named φ*NF1*-14C, showed a *Hin*dIII fragment for exons 16, 17, and 18 of the same length (1.4 kb) as the chromosome 2 *NF1* pseudogene (Figure 2), but sequence analysis revealed that this clone originated from chromosome 14 rather than from chromosome 2 (GenBank Accession Nos AF232270–AF232280). The fourth clone, φ*NF1*-14D (GenBank Accession Nos AF232281–AF232291), yielded the 1.7 kb *Hin*dIII fragment for exons 16, 17 and 18, which is present on chromosome 14, but absent on chromosomes 2 and 22 (Figure 2).

Compared with the exons in the functional *NF1* gene, several base substitutions, insertions and deletions occur in the exons represented in the pseudogenes on chromosomes 2, 14, and 22. Nevertheless, they share a high degree of identity, as shown in Figure 3. Phylogenetic analysis was performed by maximum parsimony using exonic sequences of the functional *NF1* gene and the *NF1* pseudogenes, with the functional gene as outgroup. Analysis by the neighbour-joining method resulted in the same evolutionary tree as was found by maximum parsimony and the nodes were supported under the bootstrap test. Although bootstrap analysis of the most parsimonious tree did not yield significant values, the overall results suggest that the *NF1* pseudogenes can be divided into two groups; one that is formed by the pseudogene on chromosome 2 and another by the pseudogenes on chromosomes 14 and 22.

To investigate whether the homologies between the various *NF1* pseudogene regions extend beyond the *NF1* gene duplicated segments, we determined the nucleotide sequence of flanking regions. For this purpose, primer pairs were
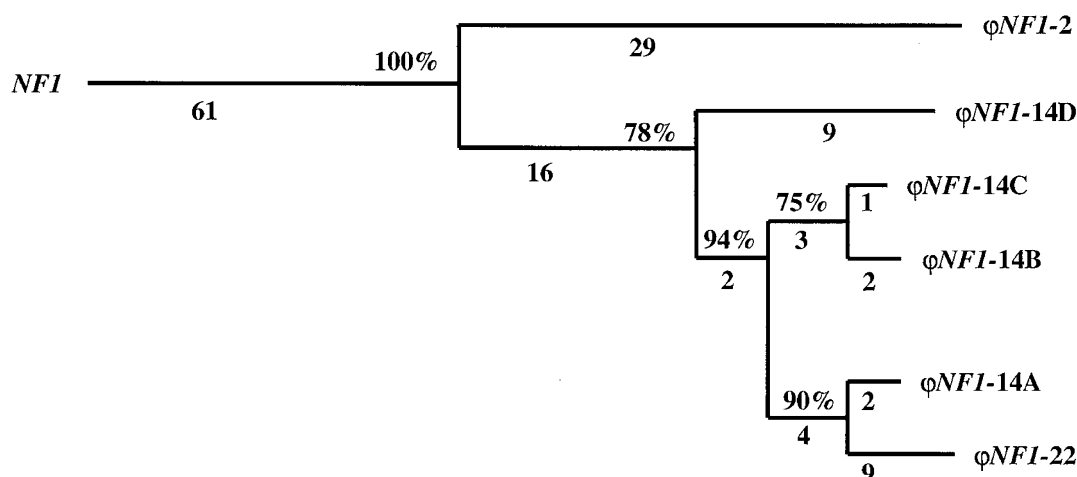

Figure 3 Phylogenetic tree of the functional *NF1* gene and its pseudogenes on chromosomes 2, 14, and 22. Branch lengths are indicated by numbers of nucleotide changes and are given below each branch. Bootstrap values are indicated above each branch as percentages of 1000 iterations.

developed on the basis of the complete sequence of the *NF1* pseudogene region on chromosome 22 and used to PCR-amplify the corresponding segments in the clones containing the chromosome 2- and chromosome 14-*NF1* pseudogenes. A BLAST search revealed that marker D22S111 (in 22q11.2) is located at about 1300 bp 5' of the *NF1*-homologous segment

on chromosome 22 (Figure 1). The primers for D22S111 and those for region A, at about 1800 bp 3' of the last segment with homology to the functional *NF1* gene (Figure 1), yielded PCR products in all clones. Each PCR product was sequenced and alignments of these sequences show at least 95% homology (Figure 4).
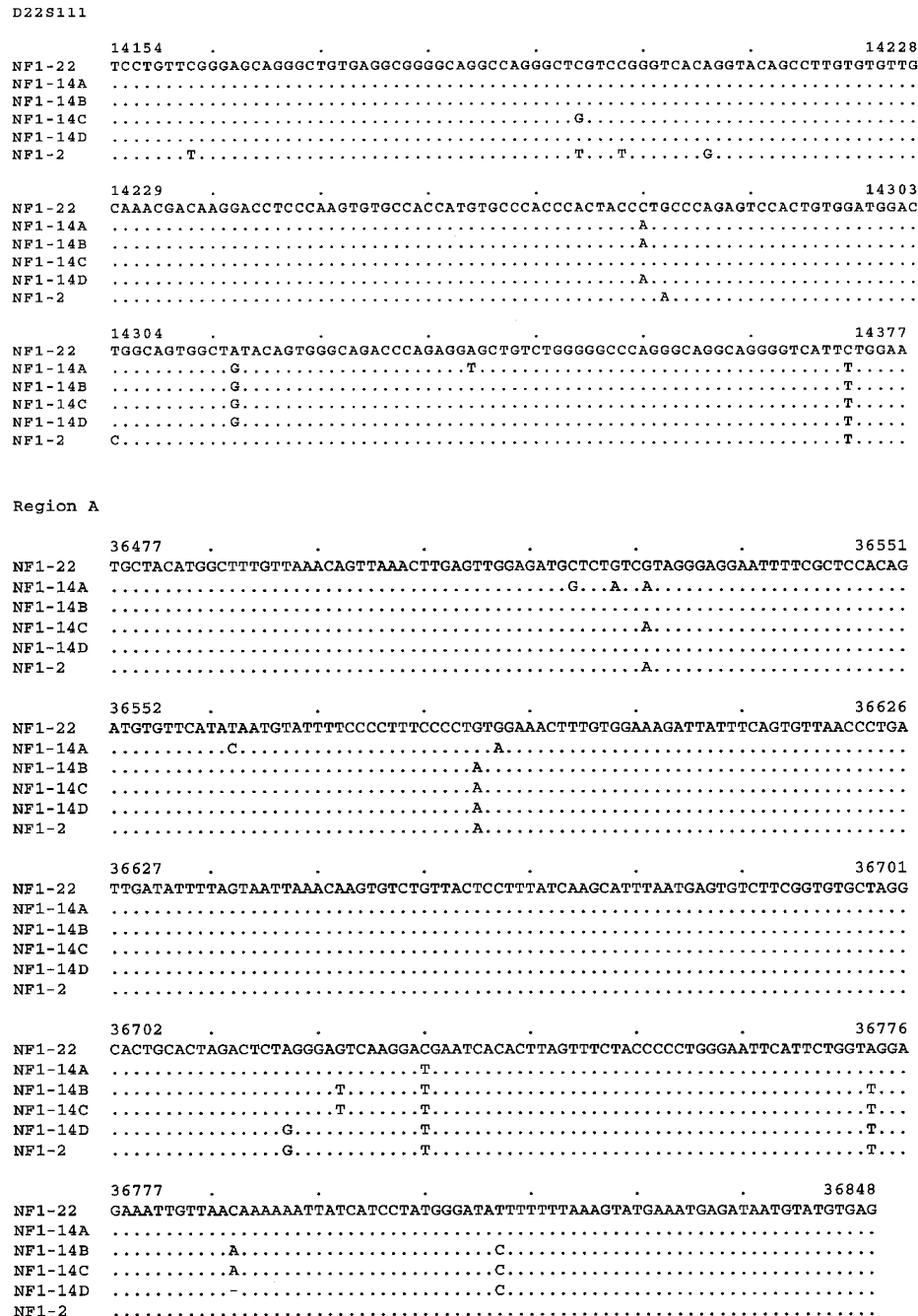
```
D22S111

         14154       .          .          .          .          .          .             14228
NF1-22   TCCTGTTCGGGAGCAGGGCTGTGAGGCGGGGCAGGCCAGGGCTCGTCCGGGTCACAGGTACAGCCTTGTGTGTTG
NF1-14A  ..........................................................................
NF1-14B  ..........................................................................
NF1-14C  .......................................................G..................
NF1-14D  ..........................................................................
NF1-2    .......T...............................................T...T.......G.......

         14229       .          .          .          .          .          .             14303
NF1-22   CAAACGACAAGGACCTCCCAAGTGTGCCACCATGTGCCCACCCACTACCCTGCCCAGAGTCCACTGTGGATGGAC
NF1-14A  ...........................................................A..............
NF1-14B  ...........................................................A..............
NF1-14C  ........................................................................
NF1-14D  ...........................................................A..............
NF1-2    .........................................................A................

         14304       .          .          .          .          .          .             14377
NF1-22   TGGCAGTGGCTATACAGTGGGCAGACCCAGAGGAGCTGTCTGGGGGCCCAGGGCAGGCAGGGGTCATTCTGGAA
NF1-14A  ..........G....................T.............................T.....
NF1-14B  ..........G..................................................T.....
NF1-14C  ..........G..................................................T.....
NF1-14D  ..........G..................................................T.....
NF1-2    C............................................................T.....


Region A

         36477       .          .          .          .          .          .             36551
NF1-22   TGCTACATGGCTTTGTTAAACAGTTAAACTTGAGTTGGAGATGCTCTGTCGTAGGGAGGAATTTTCGCTCCACAG
NF1-14A  .................................................G...A..A..................
NF1-14B  ..........................................................................
NF1-14C  ....................................................A.....................
NF1-14D  ..........................................................................
NF1-2    ....................................................A.....................

         36552       .          .          .          .          .          .             36626
NF1-22   ATGTGTTCATATAATGTATTTTCCCCTTTCCCCTGTGGAAACTTTGTGGAAAGATTATTTCAGTGTTAACCCTGA
NF1-14A  ...........C....................A.........................................
NF1-14B  ..............................A...........................................
NF1-14C  ..............................A...........................................
NF1-14D  ..............................A...........................................
NF1-2    ..............................A...........................................

         36627       .          .          .          .          .          .             36701
NF1-22   TTGATATTTTAGTAATTAAACAAGTGTCTGTTACTCCTTTATCAAGCATTTAATGAGTGTCTTCGGTGTGCTAGG
NF1-14A  ..........................................................................
NF1-14B  ..........................................................................
NF1-14C  ..........................................................................
NF1-14D  ..........................................................................
NF1-2    ..........................................................................

         36702       .          .          .          .          .          .             36776
NF1-22   CACTGCACTAGACTCTAGGGAGTCAAGGACGAATCACACTTAGTTTCTACCCCCTGGGAATTCATTCTGGTAGGA
NF1-14A  .........................T................................................T...
NF1-14B  ...................T.......T..............................................T...
NF1-14C  ...................T.......T..............................................T...
NF1-14D  ...............G...........T..............................................T...
NF1-2    ...............G...........T..............................................T...

         36777       .          .          .          .          .          .             36848
NF1-22   GAAATTGTTAACAAAAAATTATCATCCTATGGGATATTTTTTTAAAGTATGAAATGAGATAATGTATGTGAG
NF1-14A  .......................................................................
NF1-14B  ..........A....................C.......................................
NF1-14C  ..........A....................C.......................................
NF1-14D  ..........-....................C.......................................
NF1-2    .......................................................................
```

Figure 4 Alignments of sequences flanking the *NF1* pseudogenes on chromosomes 2, 14, and 22. Primers for marker D22S111 and for region A in, respectively, the 5' and 3' flanking regions of the *NF1* pseudogene on chromosome 22 (see Figure 1) were used to amplify by PCR the corresponding segments in the *NF1* pseudogene regions on clones derived from chromosomes 2 and 14. The resulting sequences were compared to the sequence of marker D22S111 and of region A on chromosome 22.

Discussion

We have determined the complete sequence of the *NF1* pseudogene on chromosome 22. As is shown in Figure 1, only the middle part of the functional *NF1* gene is represented and complex rearrangements accompanied its insertion. Exons 19b–27a were deleted, resulting in the fusion of intron 19a to intron 27a. This deletion includes exons 21–27a, which code for the functionally important GAP-related domain of the NF1 protein. Compared with the functional *NF1* gene, the exon–intron structure of its representation in the chromosome 22 *NF1* pseudogene is well conserved. The overall similarity between the sequenced regions of the functional gene and the pseudogene is 95%, with no difference in mutation rate between exons and introns.

We have identified two new *NF1* pseudogenes on chromosome 14 (φ*NF1*-14C and φ*NF1*-14D) in addition to the two already known copies on that chromosome (φ*NF1*-14A and φ*NF1*-14B).[4,7,11] The phylogenetic tree in Figure 3 demonstrates that the new pseudogenes are highly related to the other pseudogenes on chromosomes 2, 14, and 22. Moreover, they all contain the same exons and the transition of intron 19a to intron 27a, the latter marking the deletion of the GAP-related domain encoding exons. As such, they differ from other *NF1* pseudogenes, especially those on chromosome 15, in which representation of GAP-related domain encoding exons has been demonstrated.[3,5,10,11]

Based on their phylogenetic tree, Régnier *et al*[11] suggested a common ancestral sequence for the *NF1* pseudogenes on chromosomes 2, 14, 15, and 22. Duplication and transposition of this ancestral sequence would have resulted in the *NF1* pseudogenes that are now present on chromosomes 2 and 15. The results presented here indicate that this duplication process gave rise to *NF1* pseudogene copies that differed already in the presence (chromosome 2) or absence (chromosome 15) of the deletion that includes the GAP-related domain encoding exons. Duplication and transposition of the *NF1* pseudogene copy on chromosome 2 then may have resulted in the pseudogenes on chromosomes 14 and 22. Further duplication and transposition might then most likely have generated the multiple copies of *NF1* pseudogenes currently present on chromosome 14.

We demonstrated more than 95% sequence conservation in regions flanking the 5'- and 3'- ends of the *NF1* pseudogenes on chromosomes 2, 14, and 22 (Figure 4). This high homology strongly suggests that during evolution of the ancestral *NF1* pseudogene on chromosome 2, the sequences flanking this locus were also included in the subsequent duplication and transposition processes. The sequenced segments display no homology at all with the corresponding regions in the functional *NF1* gene on chromosome 17. This absence of homology may explain why our clones, contain-ing the *NF1* pseudogenes on chromosomes 2, 14, and 22, did not reveal any *in situ* hybridisation with the functional *NF1* gene on chromosome arm 17q, despite having 95% sequence homology in the represented part, or with other *NF1*-related loci. FISH experiments with neighbouring clones are in progress to determine whether or not these cross-hybridise with the *NF1* pseudogenes on chromosomes 2 and 14. These studies will enable estimation of the length of the duplicated and transposed segment.

References
1 Li Y, O'Connell P, Huntsman Breidenbach H *et al*: Genomic organization of the neurofibromatosis 1 gene (*NF1*). *Genomics* 1995; **25**: 9–18.
2 Xu G, O'Connell P, Viskochil D *et al*: The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* 1990a; **62**: 599–608.
3 Legius E, Marchuk DA, Hall BK *et al*: *NF1*-related locus on chromosome 15. *Genomics* 1992; **13**: 1316–1318.
4 Marchuk DA, Tavakkol R, Wallace MR *et al*: A yeast artificial chromosome contig encompassing the type 1 neurofibromatosis gene. *Genomics* 1992; **13**: 672–680.
5 Gasparini P, Grifa A, Origone P, Coviello D, Antonacci R, Rocchi M: Detection of a neurofibromatosis type 1 (*NF1*) homologous sequence by PCR: implications for the diagnosis and screening of genetic diseases. *Mol Cell Probes* 1993; **7**: 415–418.
6 Suzuki H, Ozawa N, Taga C, Kano T, Hattori M, Sakaki Y: Genomic analysis of a *NF1*-related pseudogene on human chromosome 21. *Gene* 1994; **147**: 277–280.
7 Purandare SM, Huntsman Breidenbach H, Li Y *et al*: Identification of neurofibromatosis 1 (*NF1*) homologous loci by direct sequencing, fluorescence *in situ* hybridization, and PCR amplification of somatic cell hybrids. *Genomics* 1995; **30**: 476–485.
8 Cummings LM, Trent JM, Marchuk DA: Identification and mapping of type 1 neurofibromatosis (*NF1*) homologous loci. *Cytogenet Cell Genet* 1996; **73**: 334–340.
9 Hulsebos TJM, Bijleveld EH, Riegman PHJ, Smink LJ, Dunham I: Identification and characterization of *NF1*-related loci on human chromosomes 22, 14 and 2. *Hum Genet* 1996; **98**: 7–11.
10 Kehrer-Sawatzki H, Schwickhardt GT, Rocchi M, Krone W: A third neurofibromatosis type 1 (*NF1*) pseudogene at chromosome 15q11.2. *Hum Genet* 1997; **100**: 595–600.
11 Régnier V, Meddeb M, Lecointre G *et al*: Emergence and scattering of multiple neurofibromatosis (*NF1*)-related sequences during hominoid evolution suggest a process of pericentromeric interchromosomal transposition. *Hum Mol Genet* 1997; **6**: 9–16.
12 Ijdo JW, Baldini A, Ward DC, Reeders ST, Wells RA: Origin of human chromosome 2: an ancestral telomere–telomere fusion. *Proc Natl Acad Sci USA* 1991; **88**: 9051–9055.
13 Hulsebos TJM, Gilbert DJ, Delattre O *et al*: Assignment of the βB1 crystallin gene (CRYBB1) to human chromosome 22 and mouse chromosome 5. *Genomics* 1995; **29**: 712–718.
14 Roe BA, Crabtree JS, Khan AS: *DNA Isolation and Sequencing: Essential Techniques Series*. John Wiley: New York, 1996.
15 Weiss RB, Dunn DM, Aoyagi A *et al*: Genbank AC004526, 1998.