AD_____

Award Number: DAMD17-98-1-8294

TITLE: Isolation of Genes from Chromosome Region Ip31 Involved in the Development of Breast Cancer

PRINCIPAL INVESTIGATOR: John K. Cowell, Ph.D.

CONTRACTING ORGANIZATION: Cleveland Clinic Foundation
Cleveland, Ohio 44195

REPORT DATE: September 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1999	Annual (1 Sep 98 - 31 Aug 99)

4. TITLE AND SUBTITLE
Isolation of Genes from Chromosome Region Ip31 Involved in the Development of Breast Cancer

5. FUNDING NUMBERS
DAMD17-98-1-8294

6. AUTHOR(S)
John K. Cowell, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cleveland Clinic Foundation
Cleveland, Ohio 44195

E-MAIL:
cowellj@ccf.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

Loss of heterozygosity (LOH) involving the p31 region of chromosome 1 has been reported as one of the most frequent genetic changes in human breast cancer. LOH is generally considered a mechanism for exposing recessive mutations in genes critical for tumorigenesis. We have begun creating a contiguous array of bacterial artificial chromosome (BAC) clones spanning the minimum region in 1p31 defined by LOH studies. These clones will facilitate the search for genes from the region which can be tested for mutations in breast tumors. One such gene, TTC4, has been identified, cloned and characterized. This gene contains a tetratricopeptide repeat motif which has been implicated in protein-protein binding and one member of this extended family of genes has been implicated in liver tumorigenesis. By establishing the exon-intron structure of TTC4 we have been able to perform an exon-by-exon survey of the gene in breast tumors. No mutations were found in any of the tumors indicating that TTC4 is probably not the gene in 1p31 responsible for breast tumorigenesis. In the absence of a clear indication about the function of the critical breast cancer gene, excluding candidate genes by mutation analysis is important in the overall search for genes from an extended genomic region.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
15

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

9-30-99

PI - Signature Date

Table of Contents

DAMD17-98-1-8294

INTRODUCTION

One of the hallmarks of tumor suppressor genes is that both copies must be inactivated for the development of the malignant phenotype. A phenomenon associated with this inactivation is that a single recessive mutation is exposed following loss of a chromosomal region containing the normal copy of gene. This event is usually manifested in tumors as loss of heterozygosity (LOH) when compared to normal tissue from the same individual. Genome screening for LOH in breast cancer has identified chromosome region 1p31 as one commonly (50%) involved in LOH indicating the site of a tumor suppressor gene. By comparison of LOH in a large number of tumors the minimal region has been defined by two markers which are 500-800 Kb apart. An extended region of loss spans 100-1500 Kb. The overall aim of this project is to characterize the region of chromosome 1p31 which shows consistent loss of heterozygosity in breast tumors by constructing a BAC contig across the region, searching for genes in this contig and then performing a mutation analysis to determine which ones are involved in tumorigenesis.

RESEARCH SUMMARY

We had previously established a YAC contig across the region and used this collection of large genomic clones as a mapping tool for newly identified genes in the region. One of these genes contained the EST marker 16a09, and was shown to map to the critical region of LOH. One of the original aims was to characterize this gene further. This turned out to be a more complicated procedure than expected. The original cDNA contained a 2.7 Kb insert which we sequenced by primer walking. An open reading frame was identified within the sequence. However when this gene was used as a probe to Northern blots a 2.0 Kb transcript was identified. This discrepancy in size implied that the original clone was probably chimeric or contained intronic sequences derived from a partially processed cDNA in the original preparation of the RNA used to create the library. To determine which part of the cDNA was not from the gene defined by 16a09 we isolated 3 different Hind III fragments from the largest cDNA clone and probed Northern blots containing RNA from a variety of tissues. The two most 3' fragments gave a transcript size of 2.0 Kb indicating that they were probably from the same gene. The 5' fragment did not identify a transcript on the Northern blot indicating it was not from a coding region at all. We therefore needed to isolate new cDNAs in order to establish the full-length sequence of this gene. Screening two different cDNA libraries using the two fragments which recognized a transcript size of 2.0 Kb, resulted in two further cDNAs being isolated which were also sequenced. Comparison with the 16a09 sequence revealed that they too were chimeric but allowed us to determine exactly at which point the sequence diverged between clones. Although there was sequence homology between the several cDNA clones isolated it was only at the 5' end where the sequence diverged. This demonstrated that many of the cDNAs we were isolating were chimeric. Since the 5' end of this gene was missing we therefore used 5' RACE and generated an approximately 500 bp fragment which was sequenced. In combination with the known consensus sequence this allowed us to identify an initiation codon, an open reading frame and a 3' untranslated region. This gene is 2005 bp long that was identical to the transcript seen on Northern blots. This somewhat protracted cloning exercise proved to be very time

consuming and delayed us in some of our other tasks (see below). From the DNA sequence for the 16a09 gene conceptual translation identified a tetratricopeptide repeat motif at the C-terminal end and a coiled coil motif at the N-terminal end of the reading frame. It has been shown that genes with these motifs are involved in protein-protein binding. The 16a09 gene is therefore a novel member of this family and is the fourth member of the human gene family. In consultation with the Human Gene Nomenclature Committee this gene was named TTC4. The description of the cloning of this gene was published recently in Genomics (see attached).

This family of genes has been associated with regulation of the cell cycle in lower organisms and has been associated with the development of liver tumors in rats. This observation suggested that TTC4 was a reasonably good candidate for the gene in 1p31 involved in breast cancer. Expression analysis showed that TTC4 is expressed almost in all tissues including normal breast. This is not unexpected for tumor suppressor genes. We, therefore, performed a preliminary mutation analysis in 8 breast tumor cells lines by using RT-PCR followed by an SSCP analysis of the PCR products. We were unable to identify band shifts in any of these cell lines indicating that mutant transcripts were not present. It was possible, however, that mutant RNAs would be unstable and so would not be detected in these types of analyses. In order to make a more detailed mutation analysis, the exon/intron structure of the gene was established from the DNA sequence of a BAC containing the TTC4 gene. It was shown that TTC4 contains 10 exons which are spread over 25 Kb of genomic DNA. All of the exons conformed to the AG.GT rule. Knowing the exon/intron boundaries made it possible for us to construct PCR primer pairs which could then be used to amplify each of the individual exons from genomic DNA for SSCP analysis. Using this approach we were still unable to detect any mutations in the coding region for this gene, although polymorphisms outside the coding region were readily identified. Although it is still possible that modifications in the promotor of this gene is, in some way, connected to an altered expression pattern in breast cancers, this would not be consistent with the observation of LOH which is generally thought to result in the exposure of a mutation in a recessive oncogene. Thus we feel that we have proven that TTC4 is not the target in the LOH region in 1p31. Although disappointing, excluding genes from involvement in tumorigenesis is still very valuable in the overall effort.

After the considerable effort expended on the analysis, of what was at first sight an excellent candidate gene, we have returned to complete the second task - that of completing the BAC contig across the region. For this we have begun with the BAC which contains TTC4 (see figure 1). Individual overlapping BACs have been identified which span 406 Kb of the region. We are currently using a limited number of other markers from the region to identify more BACs in order to construct a skeleton contig. End sequencing of these BACs will then provide PCR markers with which to develop a contiguous contig across the region that will provide the substrate for further gene isolation strategies over the next 12 months.

KEY RESEARCH ACCOMPLISHMENTS

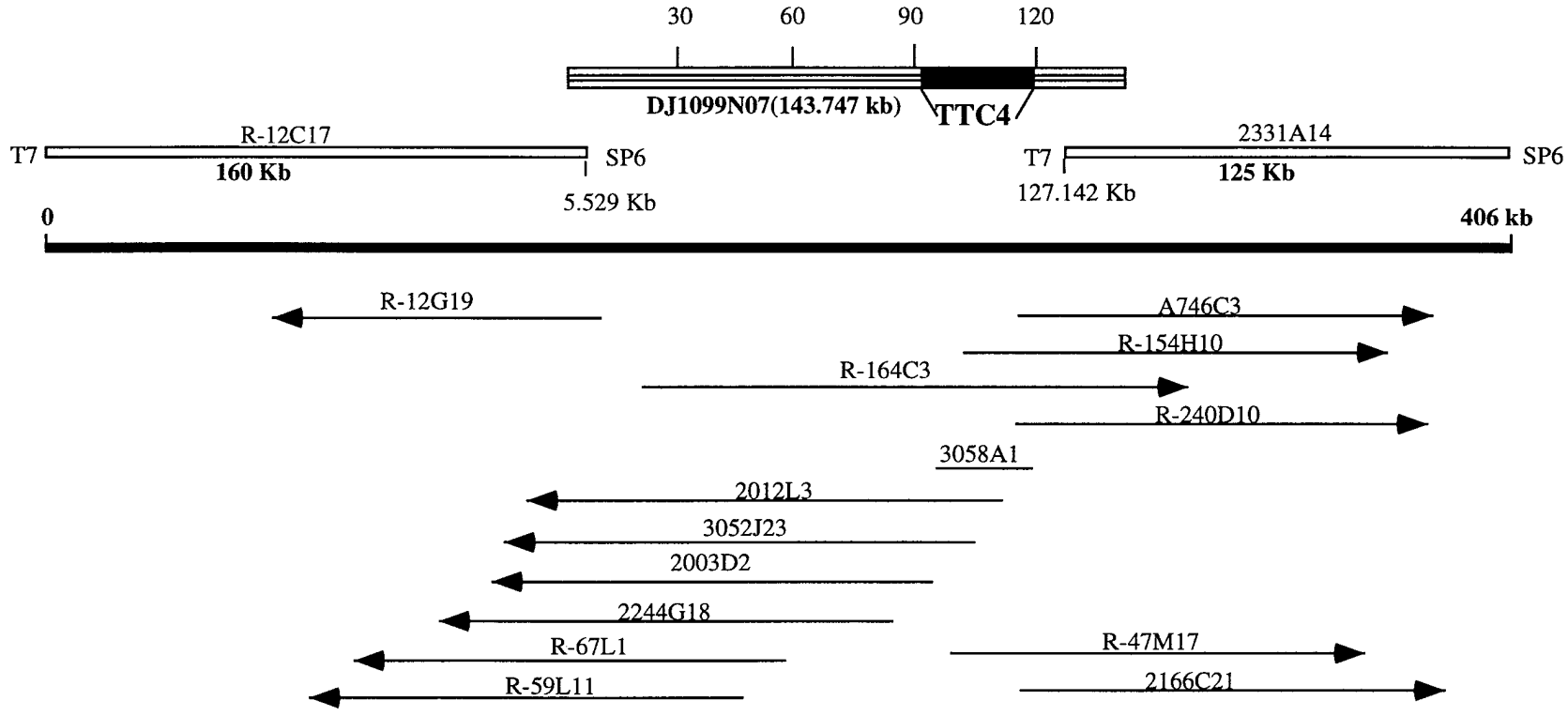
- Cloning and characterization of the TTC4 gene from 1p31.
- Exon by exon mutation analysis for the TTC4 gene in 28 breast tumors and cell lines.
- Construction of a 406 Kb BAC contig from the extended region of LOH in 1p31.

REPORTABLE OUTCOMES

Su G, Roberts T, Cowell JK. (1999) *TTC4*, a novel human gene containing the tetratricopeptide repeat and which maps to the region of chromosome 1p31 which is frequently deleted in sporadic breast cancer. Genomics 55; 157-163.

CONCLUSIONS

We have cloned, characterized, sequenced and performed mutation analysis in breast tumors for a novel gene, TTC4, which maps to the regions of chromosome 1p31 frequently deleted in breast cancers. Mutation analysis indicates that mutational inactivation of this gene is not involved in the development of these tumors. In order to identify more genes from the region we have established a 406 Kb contig of BACs which spans approximately 30-40% of the overall critical region.



BAC clones at region of chromosome 1p31

Figure 1.

Summary of the BAC contig in 1p31. DJ1099N07 contains the TTC4 gene and sequence data from this BAC allowed the identification of overlapping BACs R12C17 and 2331A14. This contig spans an approximate 400 Kb. Other BACs identified as a result of sequencing from single ends are shown below. The position of the arrowheads indicates BAC ends which have not yet been sequenced

Award Number: DAMD17-98-1-8294
John K. Cowell, Ph.D.

Genomics **55,** 157–163 (1999)
Article ID geno.1998.5633, available online at http://www.idealibrary.com on IDEAL®

TTC4, a Novel Human Gene Containing the Tetratricopeptide Repeat and Mapping to the Region of Chromosome 1p31 That Is Frequently Deleted in Sporadic Breast Cancer

Guanfang Su, Terry Roberts, and John K. Cowell[1]

Center for Molecular Genetics–NB20, Lerner Research Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, Ohio 44195

Received August 25, 1998; accepted October 19, 1998

The 1p31 region shows loss of heterozygosity in up to 50% of human breast cancers, indicating the presence of a tumor suppressor gene in this location. We have mapped six novel ESTs to a 15-Mb contig of yeast artificial chromosomes spanning the critical region of 1p31. One of these ESTs was localized within the contig to the region most commonly undergoing loss of heterozygosity in breast cancer. The corresponding gene sequence for this EST was established by cDNA cloning and RACE procedures. This gene is 2 kb long and contains a tetratricopeptide repeat motif and a coiled-coil domain. This family of genes has been implicated in a wide variety of functions, including tumorigenesis. This is the fourth member of the human gene family, and so we have named this gene *TTC4.* Northern blot analysis demonstrates a ubiquitous pattern of gene expression that includes breast tissue. A preliminary screen of human breast cancer cell lines shows that *TTC4* is expressed in all cases, but SSCP analysis of the coding region of this gene following RT-PCR failed to reveal any mutations. Clearly, because of its map location, a more extensive analysis is warranted to establish whether subtle mutations are present in breast cancers. © 1999 Academic Press

INTRODUCTION

The tetratricopeptide repeat (TPR) motif has been identified in a wide variety of genes from many different organisms and with very different functions (Murthy *et al.,* 1996). The TPR motif is a degenerate, 34-amino-acid repeat whose function is mostly unknown, although there are indications that, through formation of helical structures (Das *et al.,* 1998), the motif mediates protein–protein or protein–membrane interactions (Goebl and Yanagida, 1991; Sikorski *et al.,* 1991). Although the first TPR genes, which were identified in bacteria and lower eukaryotes, appeared to be involved in the regulation of cell division and RNA synthesis (Hirano *et al.,* 1990; Sikorski *et al.,* 1990), more recently it has been shown that members of this family are clearly related to physiological functions such as cellular stress responses (Nicolet and Craig, 1989; Honore *et al.,* 1992), the interferon response (Lee *et al.,* 1994), transport of proteins across both the mitochondrial (Steger *et al.,* 1990) and peroxisomal (Brocard *et al.,* 1994; Fransen *et al.,* 1995) membranes, transcription suppression (Tzamarias and Struhl, 1995; Schultz and Carlson, 1987), and *Drosophila* neurogenesis (Lamb *et al.,* 1995). Recently, three members of the human gene family have been isolated (Murthy *et al.,* 1996; Tsukahara *et al.,* 1996; Ohira *et al.,* 1996), although the exact function of these genes is not known.

The short arm of chromosome 1 has been shown to be involved in nonrandom genomic alterations in a variety of different human cancers (Schwab *et al.,* 1996). Observations of structural chromosome rearrangements and loss of heterozygosity (LOH) studies involving chromosome 1p have implied the presence of tumor suppressor genes along its length. In particular, LOH for the 1p31 region has been observed in up to 50% of human breast cancers (Hoggard *et al.,* 1995). Although the human EST sequencing project has generated large numbers of partial gene sequences, only a small proportion of these have so far been assigned to specific chromosomes, and of those that have a chromosomal assignment, many have not been localized to specific chromosome bands. We recently described the mapping of EST markers along the length of human chromosome 1 (Roberts *et al.,* 1996a) using a well-characterized panel of somatic cell hybrids (Roberts *et al.,* 1996b). As a result of this analysis we were able to assign groups of ESTs to five distinct regions of the short arm of chromosome 1. As part of our attempt to clone the translocation breakpoint in a patient with neuroblastoma (Mead and Cowell, 1995), we generated a 6-Mb contig of yeast artificial chromosomes (YACs) in the 1p22 region (Roberts *et al.,* 1998a, b). This contig was shown to link up with the 9-Mb contig described by

[1] To whom correspondence should be addressed. Telephone: 216 445 2688; fax: 216 444 7927; e-mail: cowellj@cesmtp.ccf.org.

0888-7543/99 $30.00

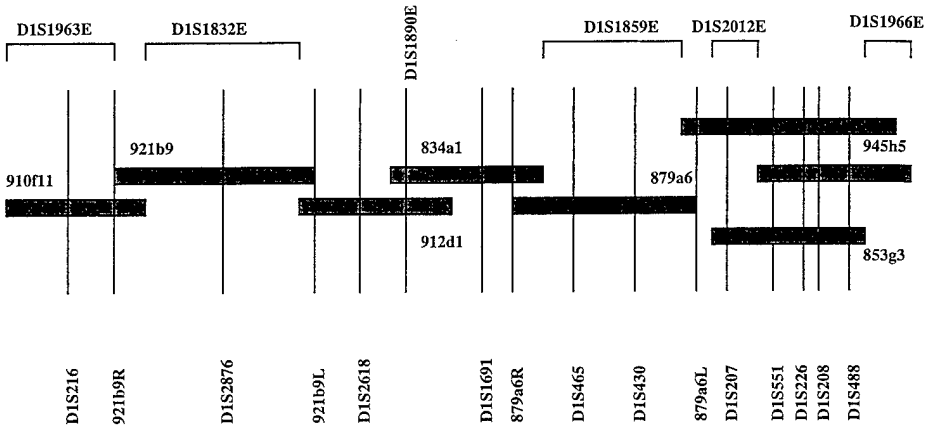


FIG. 1. Physical map of the 1p31 chromosome region frequently deleted in human breast cancer. The extent of the overlapping YACs spanning the minimally deleted region is shown by the horizontal shaded bars. Landmark STS markers from the region are shown below. The location of the six newly mapped ESTs is shown above.

Hoggard *et al.* (1995), which spanned the region of 1p31 frequently shown to undergo LOH in breast tumors. In our original mapping project we assigned 38 novel ESTs, derived from 30 different genes to the 1p21–p31 region. Six of these ESTs mapped to the 6-Mb contig described by Roberts *et al.* (1998a) in 1p22. We have now analyzed the YACs from the contig described by Hoggard *et al.* (1995) with the remaining ESTs from the 1p21–p31 region and sublocalized a further 6 ESTs within this YAC contig. One of these ESTs lay within the minimal region of LOH in breast cancer. Here we describe the isolation and characterization of this gene, *TTC4,* which is a novel member of the TPR family.

MATERIALS AND METHODS

Molecular analyses. cDNA libraries and Northern blots were screened using standard hybridization procedures with ^{32}P-labeled probes (Sambrook *et al.,* 1989). Northern blots were obtained from Clontech. Filter hybridization was carried out with the ExpressHyb procedure (Clontech).

5′ RACE. RACE was carried out with the GIBCO kit as described by the manufacturers using 2 μg of total RNA from either normal brain or liver (Clontech). Two gene-specific primers (GSP1 and GSP2) were designed. The GSP1, 5′-CTTGGCTTGTCTCCAGTCAA-3′, was 927–947 bp of cDNA 487410 from the 5′ end. The GSP2, 5′-GTGCAGATCCTGATTTGAATGC-3′, was homologous to bp 1077–1198 of cDNA 487410 from the 5′ end. Both of the GSPs were used in combination with the adapter primer (AP1) to amplify the 5′ end of the gene. The PCR reaction was performed in 1.5 mM magnesium chloride, 0.2 mM each of the four deoxynucleotide triphosphates, 0.3 mM of each primer, 5 μl of 10× PCR reaction buffer, and 1 U of *Taq* DNA polymerase (GIBCO BRL) in a volume of 50 μl. RACE products were cloned into pCR2.1 using the TA cloning procedure (Invitrogen) and transformed into JM109 competent cells (Promega). The sequencing analysis of the 5′-RACE PCR product was achieved by using M13 reverse and forward (−20) primers.

PCR and RT-PCR. RNA was isolated using Trizol reagent. Using Superscript II (GIBCO) and random hexamers (Promega), 1–2 μg of total RNA or 500 ng of mRNA was reverse transcribed. All PCRs were performed using *Taq* polymerase and the buffer supplied by GIBCO with 0.2 mM dNTPs and 0.2 μM primers. PCR products were analyzed on 3% agarose gels. SSCP was carried out as described by Hogg *et al.* (1992). After the digested RT-PCR products were denatured at 80°C for 5 min, the samples were loaded on the SSCP gel and separated by electrophoresis for 14 h at 6 W.

RESULTS

To map subregionally the 1p22–p32 ESTs identified by Roberts *et al.* (1996a), we selected a minimal set of 8 YACs that spanned 9 Mb in chromosome region 1p31 (Fig. 1). DNA from these YACs was used as the substrate for PCR with the ESTs from 1p22–p31 that did not map to the 1p22 contig described by Roberts *et al.* (1996a). Six of these ESTs were present in the YAC contig (Fig. 1). In the latest release of the human genome map, none of these ESTs has been previously assigned a map location. Of particular interest was the fact that one of these ESTs, D1S1859E, mapped to YAC 879a6, which is central to the overlapping region most commonly undergoing LOH in breast cancer (Hoggard *et al.,* 1995). Because this EST is a potential candidate for the breast cancer gene in 1p31, we characterized it more fully.

The D1S1859E marker was derived from cDNA clone 487410 (ATCC). This clone contains a 2.7-kb insert (data not shown). To facilitate sequencing, we digested this clone with *Hin*dIII/*Not*I, generating three fragments, 0.4, 0.9, and 1.4 kb long. These fragments were subcloned, and the full 2644-bp sequence was generated. A poly(A) tail was identified at the 3′ end, and a 933-bp open reading frame was found between nt 877 and 1809. When cDNA 487410 was used to probe a multiple tissue Northern blot, an approximately 2-kb transcript (Fig. 2) was observed in most tissues. Because this transcript was 600 bp smaller than the cDNA clone insert, the individual subclones were also hybridized to Northern blots. Only the 0.9- and 1.4-kb fragments generated hybridization signals, both of which were 2 kb long. This result indicated that the original 487410 clone was probably chimeric. The 0.9- and 1.4-kb fragments were then used to screen both fetal brain and adult liver cDNA libraries. A 1.7-kb clone (FC4) was isolated from the fetal brain library

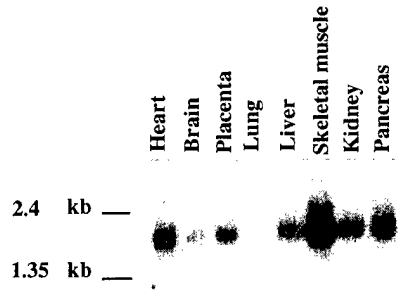


FIG. 2. Northern blot analysis of the distribution of tissue expression of the *TTC4* gene. The multiple tissue blot was probed with cDNA 487410 and detects a 2-kb transcript in all cases. Expression in lung was detectable at low levels.

using the 0.9-kb probe, and a 3.8-kb clone (LC4) was isolated from the fetal liver library using the 1.4-kb probe. These two new clones were sequenced, and, when they were compared with cDNA 487410, it was clear that both of these cDNAs were also chimeric; the 950 bp at the 3′ end of FC4 was identical to the 5′ end of 487410, whereas the 1784 bp at the 3′ end of LC4 was identical to the 3′ end of 487410 (Fig. 3). Database homology searches with the 5′ end of LC4 revealed homologies with the 3′ untranslated region of a gene on chromosome 22. Thus, as a result of this analysis, the 5′ end of the gene that maps to chromosome 1p31 was still not accounted for. To find the 5′ end we used the Marathon-ready brain cDNA RACE approach. Two gene-specific primers (GSP) were designed from the 5′ end of 487410 within the region showing homology with LC4 and FC4 (Fig. 3). Thus, GSP1 annealed to nt 928–947 and GSP2 annealed to nt 1077–1098. These GSPs were used in combination with the AP1 provided in the RACE procedure. The two resultant bands (Fig. 3), which were 232 and 382 bp long, were subcloned into the pCR 2.1 plasmid vector and then these clones were sequenced. The 232-bp product showed complete homology to the 382-bp fragment at the 5′ end (Fig. 3). The sequence at the 3′ end of each of these clones was homologous to 487410 (Fig. 3). When the 382-bp fragment was used to probe the same Northern blot used previously, a 2-kb transcript was identified (data not shown). When the 382-bp fragment was incorporated into the 3′ sequence of the 487410 cDNA a 2005-bp gene sequence was generated (Genbank Accession No. AF073887). Conceptual translation of this sequence identified a 1071-bp (356 amino acid) open reading frame spanning nt 34–1104 (Fig. 4). This gene contains a tetratricopeptide repeat motif at the 5′ end of the gene (Fig. 4). When the amino acid sequence of *TTC4* was analyzed using the coiled-coil prediction program in the ExPASY-Tools (NCBI), a coiled-coil domain that is situated between aa 170 and aa 230 (Fig. 4) was identified.

Homology searches using the SwissProt program revealed 26–41% identify with several genes in the database (Fig. 5). In all cases, the highest homology was within a TPR motif. Amino acid sequence alignment over the TPR domains reveals that *TTC4* has 16% amino acid identity with TPR1 (Murthy *et al.*, 1996), 26% with TPR2 (Murthy *et al.*, 1996), 19% with the transformation-sensitive protein, IEP SSP 3521 (Honore *et al.*, 1992), and 33% with the yeast STI1 protein (Wilson *et al.*, 1994). All of the aligned proteins

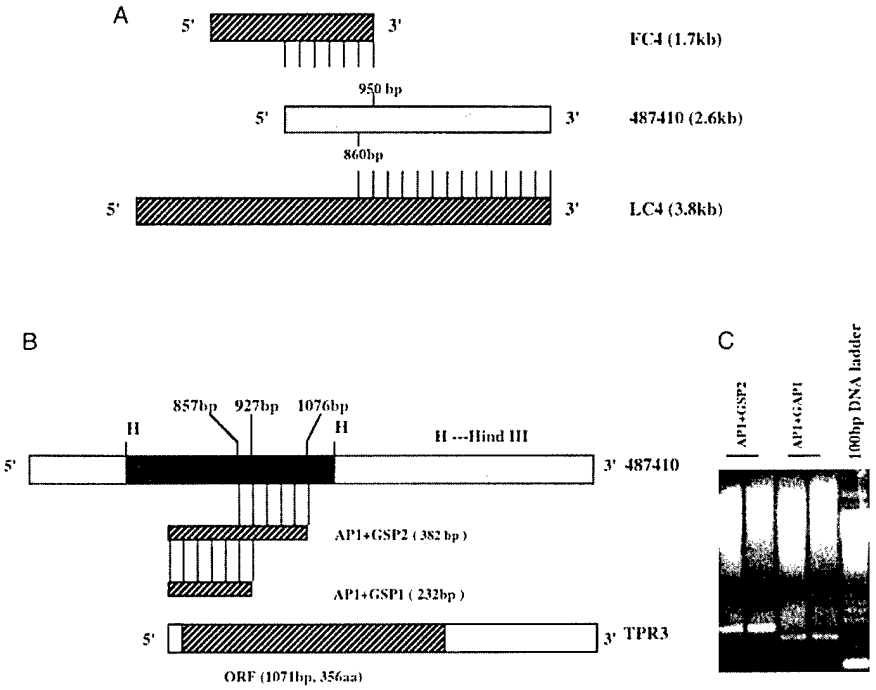


FIG. 3. Construction of the *TTC4* gene sequence. (**A**) The homologous regions between the original clone, 487410, and the newly isolated clones from fetal brain (FC4) and adult liver (LC4) are shown. (**B**) The relationship between the two RACE products (**C**) derived from 487410 that allowed the final gene structure to be established is shown.

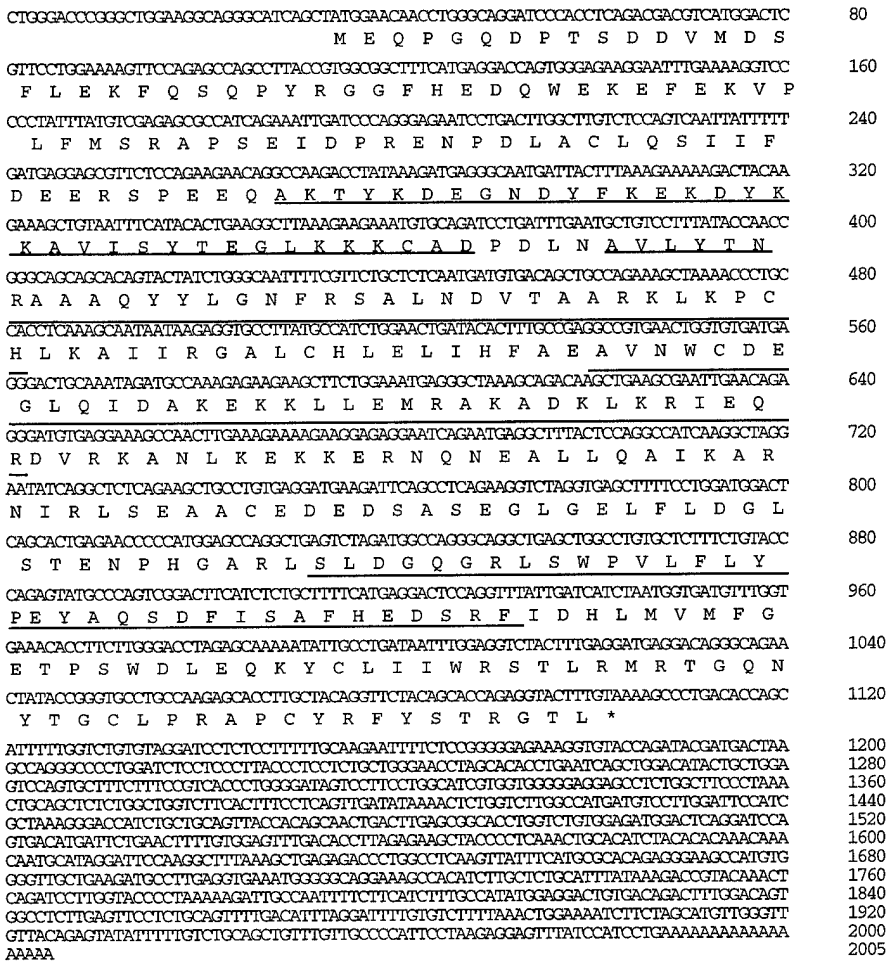


FIG. 4. Nucleotide sequence and the deduced amino acid sequence of the *TTC4* cDNA. The four potential TRP repeat motifs are underlined.

contain the degenerate 34-amino-acid motif that defines this structure. TPRs are quite often tandemly arrayed, and it is also common for single and double TPRs to be found separated from the arrays (Lamb *et al.,* 1995). The *TTC4* protein has four separated TPR motifs that are spread throughout the *TTC4* protein. Coiled-coil domains were also found in TPR1, TPR2, IEF SSP 3521, and the yeast STI1 protein (Fig. 4). The location of the coiled-coil domain varies within the proteins. In TPR1 and TPR2 they are found within the TPR domain, whereas in the other genes the coiled coil is located outside the TPR domain toward the carboxy terminal of the gene (Fig. 6). *TTC4* also shows identity with some other proteins in the database such as Lm STI1, serine/threonine protein phosphatase T, and the heat shock protein STI1 (data not shown).

Mutation Analysis in Breast Cancer Cell Lines

Northern blot analysis demonstrates that *TTC4* is expressed in normal breast tissue (data not shown). In a preliminary investigation of whether *TTC4* is mutated in breast cancer cells, we undertook a limited mutation screen in eight breast cancer cell lines (G1C1A, MCF7, MCF7ICI, MDS 435, MDS 453, MDS 468, SKBR3, and T47D) using SSCP. Because we do not have information about the exonic structure of *TTC4,* we analyzed RNA transcripts from these cells. The coding region of *TTC4* (nt 145–945) was amplified using RT-PCR. Because conformation changes are more readily detectable by SSCP in smaller fragments (Hogg *et al.,* 1992), the PCR products were digested with *Bsm*1 and *Bgl*II to generate fragment sizes below 300 bp. When the mobilities of the individual fragments on the SSCP gel were compared, we were unable to detect any differences suggesting mutations or polymorphisms. DNA samples from the same eight cell lines were also analyzed by Southern blotting, and no structural rearrangements of this gene were detected. Unfortunately, it was not possible to determine which of these tumors had undergone LOH at 1p31 because the corresponding normal tissue was not available. Given the overall frequency of LOH in breast tumors, however it might be expected on average that 50% of them had experienced LOH in 1p31.

DISCUSSION

The human TTC family of genes previously contained three members, *TTC1* from chromosome region

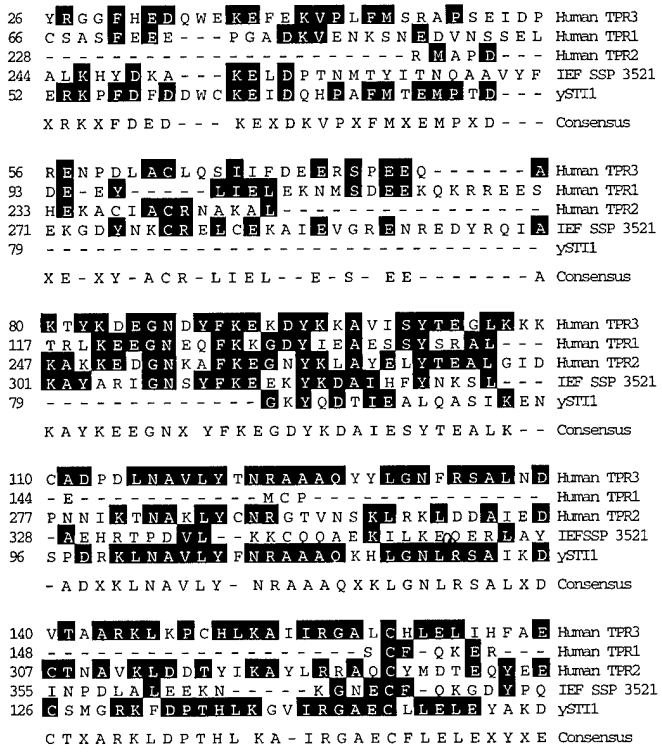


FIG. 5. Multiple sequence alignment with a partial amino acid sequence with those from human TPR1, TPR2, IEF SSP 1532, and yeast STI1.

5q32–q33, *TTC2* from chromosome region 17q11–q23, and *TTC3* from chromosome region 21q22 (Murthy *et al.,* 1996; Tsukahara *et al.,* 1996; Ohira *et al.,* 1996). We have now identified the fourth member of this family, *TTC4,* from chromosome region 1p31. *TTC4* contains four repeats of the motif and is apparently ubiquitously expressed, as are all of the other members of the family (Murthy *et al.,* 1996; Tsukahara *et al.,* 1996; Ohira *et al.,* 1996). A role for the TTC gene family in neoplasia has recently been demonstrated in liver tumorigenesis (Isfort *et al.,* 1997). The *Tg737* gene, which contains 10 tetratricopeptide repeats, was identified as the gene that was inactivated by insertional mutation in a transgenic mouse with a phenotype resembling autosomal recessive polycystic kidney disease (Moyer *et al.,* 1994). These mice also showed liver dysplasia. When Isfort *et al.* (1997) analyzed genetic changes in carcinogen-induced liver tumors in rats, they detected deletions and rearrangements of *Tg737* in 40% of cases. Similar changes were detected in human liver, pancreas, and kidney tumors but not in breast cancers. In cells lacking *Tg737,* overexpression of this gene resulted in a decrease in cell growth *in vitro* and a suppression of the ability of these cells to form tumors in nude mice. More circumstantial evidence that *TTC1* and *TTC2* play a role in tumorigenesis came from a yeast two-hybrid analysis that showed that they interact preferentially, via their tetratricopeptide domains, with a truncated form of the GAP-related domain of the *NF1* gene, which predisposes to the development of neurofibromatosis. Yeast two-hybrid associations do not necessarily reflect intracellular interactions, however, and to date, these interactions have not been confirmed by other means. Another member of the TTC family, p58, has been shown to inhibit the interferon-inducible *PKR* gene (Lee *et al.,* 1994). *PKR* has been suggested to be a tumor suppressor gene, and p58 has oncogenic potential. Thus, although the function of *TTC4* is not known, it might be predicted to have other interesting functions related to tumorigenesis because of its relatedness to other TTC-motif genes. In yeast, for example, some of the cell division cycle genes, which are essential for the progression of the cell cycle, contain TTC repeats (Boguski, 1990; Sikorski *et al.,* 1990). These genes are also homologous to fungal genes that are involved in mitosis (O'Donnell *et al.,* 1991). The human homolog of one of the yeast TPR-containing genes was shown to be upregulated in SV40-transformed human fibroblasts (Honore *et al.,* 1992) compared with their normal counterparts, again possibly suggesting a role in the transformation process. *TTC4* also shares homology to stress-inducible proteins such as STI1 in yeast, a member of a family of genes from various species (Blatch *et al.,* 1997) with a common ancestor. STI1 is important for growth under suboptimal conditions and can either transactivate other genes (Nicolet and Craig, 1989) or, as shown for the homologous *Ski* gene (Toh and Wickner, 1980), regulate RNA synthesis. All of these func-

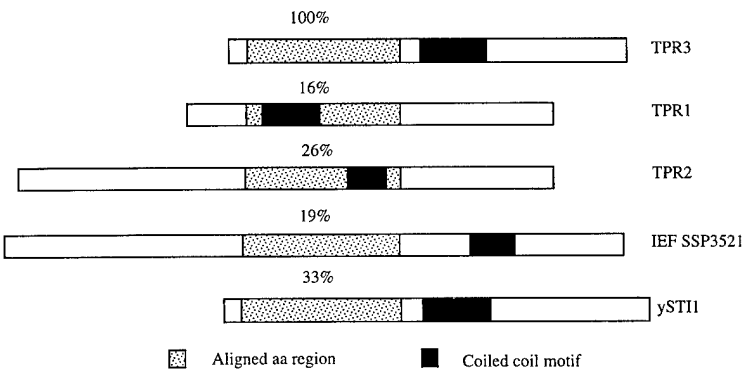


FIG. 6. Diagrammatic representation of the intragenic location and degree of homology of the TPR domains between *TTC4* and other TPR-containing genes. The relative locations of the predicted coiled-coil domains within these genes are also indicated.

tions are potentially relevant to cancer cell growth and survival.

The function of the *TTC3* gene, identified originally as the *TPRD* gene, which was isolated from the Down syndrome (DS) critical region in 21q22 (Tsukahara *et al.,* 1996; Ohira *et al.,* 1996), is also poorly understood, although it is speculated to be responsible for some of the developmental abnormalities seen in DS patients, based on its pattern of expression in the developing mouse (Tsukahara *et al.,* 1998).

The crystal structure of a TPR-containing gene, *PP5,* was recently reported (Das *et al.,* 1998). This study demonstrated that the repeat motif folds into a right-handed superhelical structure, creating a continuous helical groove that would be suitable for the recognition of target proteins. This structure represents a novel mechanism of protein recognition and has the capacity to interact with many proteins and form a multiprotein complex.

The demonstration that *TTC4* maps within the region of LOH in breast cancer makes it a potential candidate for a role in breast tumorigenesis. One unequivocal way of implicating a candidate gene in tumorigenesis is to identify mutations consistently in a specific tumor type. Because LOH is the observation for 1p31 in breast cancer, the prediction is that these would be inactivating mutations, although this is not always the case. The WT1 gene, for example, is clearly implicated in Wilms tumorigenesis, and, although located within the region of chromosome 11 frequently undergoing LOH in Wilms tumors, mutation analysis can identify mutations in this gene in only a handful of tumors, mostly from patients with specific subtypes of tumors and genetic predispositions (Cowell *et al.,* 1993). At present, we have not established the exon/intron structure of *TTC4,* which would allow mutation analysis at the genomic level. Consequently, in a preliminary screen, we only analyzed RNA from a series of commonly available breast cancer cell lines. The advantages of using these cells are that they are not contaminated by normal cells and that RNA can be prepared from them relatively easily. This is not always true for many of the clinical specimens obtained from breast tumors. Even though no mutations were detected in these experiments, a more extensive screen is still warranted because of the implications of the TTC genes in a variety of cellular functions potentially related to tumorigenesis. Once the structure of the *TTC4* gene has been characterized, an exon by exon analysis would clearly be more informative because this analysis would not be influenced by the consequences of the mutation on gene expression or stability (Kratzke *et al.,* 1994).

REFERENCES

Blatch, G. L., Lassle, M., Zetter, B. R., and Kundra, V. (1997). Isolation of a mouse cDNA encoding mSTI1, a stress-inducible protein containing the TPR motif. *Gene* **194:** 277–282.

Boguski, M. S. (1990). Expanding family. *Nature* **346:** 114.

Brocard, C., Kragler, F., Simon, M. M., Schuster, T., and Hartig, A. (1994). The tetratricopeptide repeat-domain of the PAS10 protein of *Saccharomyces cerevisiae* is essential for binding the peroxisomal targeting signal-SKL. *Biochem. Biophys. Res. Commun.* **204:** 1016–1022.

Cowell, J. K., Groves, N., and Baird, P. N. (1993). Loss of heterozygosity at 11p13 in Wilms' tumour does not necessarily involve mutations in the WT1 gene. *Br. J. Cancer* **67:** 1259–1261.

Das, A. K., Cohen, P. T. W., and Barford, D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: Implications for TPR-mediated protein–protein interactions. *EMBO J.* **17:** 1192–1199.

Fransen, M., Brees, C., Baumgart, E., Vanhooren, J. C., Baes, M., Mannaerts, G. P., and Van Veldhoven, P. P. (1995). Identification and characterization of the putative human peroxisomal C-terminal targeting signal import receptor. *J. Biol. Chem.* **270:** 7731–7736.

Goebl, M., and Yanagida, M. (1991). The TPR snap helix: A novel protein repeat motif from mitosis to transcription. *Trends Biochem. Sci.* **16:** 173–177.

Hirano, T., Kinoshita, N., Morikawa, K., and Yanagida, M. (1990). Snap helix with knob and hole: Essential repeats in *S. pombe* protein nuc2+. *Cell* **60:** 319–328.

Hoggard, N., Hey, Y., Brintell, B., James, L., Jones, D., Mitchell, E., Weissenbach, J., and Varley, J. M. (1995). Identification and cloning in yeast artificial chromosomes of a region of elevated loss of heterozygosity on chromosome 1p31.1 in human breast cancer. *Genomics* **30:** 233–243.

Hogg, A., Onadim, Z., Baird, P. N., and Cowell, J. K. (1992). Detection of heterozygous mutations in the RB1 gene in retinoblastoma patients using single-strand conformation polymorphism analysis and polymerase chain reaction sequencing. *Oncogene* **7:** 1444–1461.

Honore, B., Leffers, H., Madsen, P., Rasmussen, H. H., Vandekerckhove, J., and Celis, J. E. (1992). Molecular cloning and expression of a transformation-sensitive human protein containing the TPR motif and sharing identity to the stress-inducible yeast protein STI1. *J. Biol. Chem.* **267:** 8485–8491.

Kratzke, R. A., Hogg, A., Otterson, G. A., Coxon, A. B., Geradts, A., Cowell, J. K., and Kaye, F. J. (1994). Incomplete penetrance of familial retinoblastoma and benign retinal tumors associated with partial inactivation of the RB product. *Oncogene* **9:** 1321–1326.

Lamb, R. J., Tugendreich, S., and Hieter, P. (1995). Tetratricopeptide repeat interactions: To TPR or not to TPR? *Trends Biochem. Sci.* **20:** 257–259.

Lee, T. G., Tang, N., Thompson, S., Miller, J., and Katze, M. G. (1994). The 58,000-Da cellular inhibitor of the interferon-induced double stranded RNA-activated protein kinase (PKR) is a member of the tetratricopeptide repeat family of proteins. *Mol. Cell. Biol.* **14:** 2331–2342.

Mead, R. S., and Cowell, J. K. (1995). Molecular characterisation of a t(1;10)(p22;q21) constitutional translocation from a patient with neuroblastoma. *Cancer Genet. Cytogenet.* **81:** 151–157.

Murthy, A. E., Bernards, A., Church, D., Wasmuth, J., and Gusella, J. F. (1996). Identification and characterization of two novel tetratricopeptide repeat-containing genes. *DNA Cell Biol.* **15:** 727–735.

Nicolet, C. M., and Craig, E. A. (1989). Isolation and characterization of STI1, a stress inducible gene from *Saccharomyces cerevisiae. Mol. Cell. Biol.* **9:** 3638–3646.

O'Donnell, K. L., Osmani, A. H., Osmani, S. A., and Morris, N. R. (1991). bimA encodes a member of the tetratricopeptide repeat family of proteins and is required for the completion of mitosis in *Aspergillus nidulans. J. Cell Sci.* **99:** 711–719.

Ohira, M., Ootsuyama, A., Suzuki, E., Ichikawa, H., Seki, N., Nagase, T., Monura, N., and Ohki, M. (1996). Identification of a novel