Award Number: DAMD17-02-1-0384

TITLE: Cloning of a New Gene/s in Chromosome 17p3.2-p13.1 That Control Apoptosis

PRINCIPAL INVESTIGATOR: Jose Russo, M.D.

CONTRACTING ORGANIZATION: Fox Chase Cancer Center
Philadelphia, Pennsylvania 19111

REPORT DATE: April 2003

TYPE OF REPORT: Annual

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
Cloning of a New Gene/s in Chromosome 17p3.2-p13.1 That Control Apoptosis

Jose Russo, M.D.

Fox Chase Cancer Center
Philadelphia, Pennsylvania 19111
E-Mail: J_Russo@fccc.edu

Loss of genetic material (LOH) in the chromosome 17 p13.2 at the microsatellite marker D17S796 has been identified in atypical ductal hyperplasia and in situ ductal carcinoma of the breast. Our results shown LOH at the same region in MCF-10F cells treated with the chemical carcinogen benz(a)pyrene (BP). Moreover, microcell-mediated transfer of an intact copy of human chromosome 17 inhibited the tumorigenicity of BP1E and PCR-SSCP analyzes showed a restoration of the lost material in BP1E-17-neo. These experiments suggested the presence of a tumor suppressor gene in 17p13.2 near the marker D17S796. We have been able to clone a fragment of the genes that could represent the last exon of a bigger peptide. The presence of a 3' splicing site in the putative introns and the ATTTAAC region at the 3'end support this idea. The predicted amino acid sequence does not share significant homology with any known protein supporting the idea that this could be a novel protein. Further experiments will be done in order to clone the full-length cDNA and to study the regulation of the expression of this novel gene.
RUSSO, Jose

Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRONT COVER</td>
<td>1</td>
</tr>
<tr>
<td>STANDARD FORM 298</td>
<td>2</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>3</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>4</td>
</tr>
<tr>
<td>BODY</td>
<td>4</td>
</tr>
<tr>
<td>KEY RESEARCH ACCOMPLISHMENTS</td>
<td>16</td>
</tr>
<tr>
<td>REPORTABLE OUTCOMES</td>
<td>16</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>16</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>16</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>N/A</td>
</tr>
</tbody>
</table>
A-INTRODUCTION

Breast cancer is a hormone dependent malignancy whose incidence is steadily increasing in most western societies and in countries that are becoming industrialized (1-5). Despite significant advances in early detection, surgery, radiation, chemotherapeutic and endocrine therapy, the mortality caused by this disease, that in the United States is only second to lung cancer as a cause of cancer-related death, has remained almost unchanged for the past five decades (1). New studies of the biochemical mechanisms evoked by conventional treatments for neoplastic diseases point to apoptosis as a key process for elimination of unwanted cells (6). Impaired function of apoptosis-related genes is deeply involved in oncogenesis and the progression of cancers (6-9). Among the genes that control apoptosis is Fas (CD95/Apo-1) a cell membrane receptor that upon binding by its ligand (FasL) triggers a signal resulting in apoptotic cell death (8-10). Fas is a cell-surface receptor that exists in two forms, transmembrane and soluble. The former induces apoptosis by ligation of FasL or agonistic anti-Fas antibody, whereas the latter inhibits Fas-mediated apoptosis by neutralizing its ligand (11, 12). Harnessing the power of this complex molecule is expected to lead to the development of powerful chemotherapeutic advances. Our Laboratory has recently found a link between the Fas ligand-receptor complex in chemically transformed human breast epithelial cells and a gene/s located in chromosome 17p13.2p13.1 (13). Whereas, gene transfer of Fas ligand (CD95L) using adenoviral vectors has been shown to generate apoptotic responses and potent inflammatory reactions that can be used to induce the regression of malignancies in vivo, there are significant unwanted reactions such as hepatotoxicity that may limit their clinical utility (7,10). Therefore what is needed is to identify which is the gene or genes that may regulate Fas ligand-receptor complex (12,14-18). For this purpose we have proposed the following aims: 1) To isolate in the precise location in chromosome 17p13.2-p13.1 the gene(s) responsible for the control of Fas-receptor ligand complex function, 2) To test the cloned gene/s functionality in cell transformation by cDNA transfection studies in vitro, and 3) To determine the functional role of the isolated gene in the process of neoplastic progression in vivo.

B-BODY

B-i--The experimental system. We have developed an in vitro - in vivo system in which the environmental carcinogen benz(a)pyrene (BP) has been utilized for inducing in vitro the transformation of human breast epithelial cells (HBEC) (19-32). For developing this paradigm of human breast cancer we have capitalized in the availability of the mortal HBEC-MCF-10M, which without viral infection, cellular oncogene transfection, or exposure to carcinogens or radiation became spontaneously immortalized, originating the cell line MCF-10F (33,34). Treatment of MCF-10M cells with chemical carcinogens failed in inducing cell transformation, while MCF-10F cells responded to in vitro treatment with BP with the expression, in a progressive fashion, of all the phenotypes indicative of neoplastic transformation. BP-treated MCF-10F cells expressed increased survival and formation of colonies in agar methocel, loss of ductulogenic properties in collagen matrix, invasiveness in a Matrigel in vitro system (clones BP-l) and tumorigenesis in severe combined immunodeficient (SCID) mice (BPI-E) (19, 22, 27). We have demonstrated, by utilizing DNA amplification of microsatellite length polymorphism, allelic losses and microsatellite instability in different areas of the genome of the HBEC transformed by chemical carcinogens (24, 25, 34-37, 42, 43). MSI was detected in chromosomes 11 and 17 in the immortalized cells, and additional MSI in chromosomes 11, 13 and 17 in the transformed cells. LOH was detected only in one locus of chromosome 17p13.2-p13.1 in the transformed cells BPI-E using the marker D17S796 (13) (Figure 1). We have been unable to identify any other genetic alterations in more than 40 other markers, including p53, in the p or q arm of chromosome 17 (13). Because chromosomes 11 and 17 were involved in both the early and late stages of carcinogenesis we selected them for testing their functional roles in chemically transformed HBEC using a microcell-mediated chromosome transfer (MMCT) technique (35-38). Our study found that seven out of ten clones with chromosome 17 transferred in to
BP1E cells had reverted transformed phenotypes such as advantageous in growth, colony formation in agar-methocel, loss of ductulogenesis and resistant to Fas mediated apoptosis (13). Our molecular analysis of MCF10F and BP1E cells revealed that in the process of chemical transformation loss of genetic material (LOH) in the chromosome 17 p13.2 (D17S796), which has been identified by other investigators in atypical ductal hyperplasia and in situ ductal carcinoma of the breast (39), has been restored in the cells with abrogated transformation phenotype (Figure 1).

![Figure 1: LOH in BP1E using the marker D17S796 and restoration of this region in BP1E-17 neo. Single-stranded conformation polymorphism (SSCP) analyses were done using DNA from MCF-10F (line 1), BP1E (line 2) and BP1E-17 neo clone II-3 (line 3). PCR reactions were performed in a total volume of 10 µl containing: 1X buffer with MgCl₂ (Perkin Elmer), 400 µM of each dATP, dTTP, dCTP and dGTP, 200 µM of the primer D17S796 Reverse, 200 µM of the primer D17S796 Forward, 0.2 U Taq Polymerase (Perkin Elmer), 1 µCi [γ²P]-dATP and 120 ng total DNA. The thermal cycling consisted on: 94°C, 2 min for denaturation, and 35 cycles of 94°C, 45 sec; 56°C, 45 sec and 72°C, 45 sec followed by an extension at 72°C, 10 min. PCR products were run on 6% polyacrylamide gel and autoradiographed.]

This indicates that the LOH in this region may regulate the early event of breast cancer initiation and that the region of D17S796 (7586bp) might contain gene/s controlling the transformation phenotypes. More importantly it was observed that in those clones in which the reversions of the phenotype were present also the cells, loss the resistance to Fas mediated apoptosis. Transfer of chromosome 11 neither revert the transformation phenotypes or the resistance to the Fas mediated apoptosis. These data allowed us to postulate that chromosome 17 p13.2 (marker D17S96) might contain a gene or genes that are controlling this process.

**B-ii-Progress Report.**

In this report we cover the work performed from July 1, 2002 to March 30, 2003, covering the statement of work 1 that is to isolate in the precise location in chromosome 17p13.2 the gene(s) responsible for the control of Fas-receptor ligand complex function.

**B-ii-a- Methods and procedures.**

**DNA isolation:** DNA was prepared from MCF-10F, BP1E and BP1E-17 neo (clone II-3). All the cells cultures were treated with lysis buffer (100 mM NaCl, 20mM Tris-Cl pH 8.0, 25 mM EDTA pH 8.0, 0.5% SDS) with 200 µg/ml proteinase K and incubated at 65°C, 15 minutes with gentle agitation. The samples were cooled down on ice and treated with 100µg/ml RNase at 37°C, 30 minutes. The samples were purified by extracting with 1 volume phenol followed by an extraction with 1 volume of chloroform: isoamyl alcohol (24:1). The aqueous layer was adjusted to 0.75M with ammonium acetate and the DNA
was precipitated with 2 volumes of 100% ethanol. The samples were centrifuged, dried and dissolved in distilled water.

**PCR amplification:** PCR amplification was performed using the High Fidelity PCR system (Roche). The primers used to amplify were: Primer JR 853 Reverse 5' CTCTCTTAGGATTTACCTTCC 3' and Primer JR 864 Forward 5' TCTTTCTGATGGCTCTCC 3'. The genomic DNA from MCF-10F, BP1E and BP1E-17 neo (clone II-3) were use as templates in different PCR reactions. The PCR was performed in a 100µl volume containing 100 ng of genomic DNA, 1X buffer with MgCl₂, 200 µM of each dATP, dCTP, dGTP, and dTTP, 300 nM of each primer, and 5U Expand High Fidelity enzyme (Roche). The PCR conditions were: denaturalization at 94°C, 2 min followed by 30 cycles at 94°C for 15 sec, 60°C for 30 sec and 72°C for 50 sec, followed for 72°C, 10 minutes for extension. The 10 minutes extension is important to ensure that all PCR products are full length and 3'adenylated (A).

**Cloning of the PCR products:** The PCR products were cloned in the vector pCR2.1 using the TOPO TA cloning kit (Invitrogen). The linearized vector supplied has a single, overhanging 3' ends deoxythymidine (T) residues. This allows PCR inserts 3' adenylated to ligate efficiently with the vector. The ligations were used to transform *Escherichia coli* TOP10 (Invitrogen) using the chemical transformation method. The transformants were plated on LB agar with 50µg/ml kanamycin and X-gal and the plates were incubated at 37°C during 24h. The plasmid pCR 2.1 has a fragment of the lacZ gene and the cloning site is within this region. The peptide LacZ produces a blue color when the bacteria are plate on X-Gal. The LacZ peptide will not be present and the colonies remains white if the lacZ gene has been interrupted by the cloned insert.

**Plasmid preparation and Sequencing:** Mini prepreparations of the plasmid DNA were made using the alkaline lysis method (48). The pellet was resuspended in 100 µl of distilled water and treated with RNAase (20 µg/ml) during 30 minutes at 37°C. One volume of PEG/NaCl (13% 6-8K PEG, 1.6M NaCl) was added and incubated on ice 30 minutes. The samples were centrifuged 30 minutes at 4°C. The pellet was resuspended in 50 µl of distilled water. The sequencing reactions were done at the Fox Chase Cancer Center Facility. The M13 Reverse and M13 Forward primers were used for sequencing. The Sequencher and GCG programs were use for the sequences analysis.

**RT-PCR:** RNA was extracted from MCF-10F, BP1E and BP1E-17neo using Trizol reagent (Life Technologies, Inc.) and the pellet was resuspended in DEPC water. Approximately, 4 µg of RNA was treated with 4 U DNase I (Ambion) during 30min at 37°C followed by the addition of 5 µl of Dnase Inactivation reagent (Ambion) according to manufacture protocol. Another step of DNase I inactivation was made by incubation at 80°C, 10 minutes. The sample was divided in two tubes with 2 µg of RNA each. One of these RNA samples was used to prepared cDNA and the other was used to test contamination with DNA (negative control). For the cDNA synthesis, 2 µg of RNA was mix with 0.4 µg oligo (dT) 12-18 in a final volume of 9.6 µl and incubated at 70°C during 10 min and cool down on ice. The following reagents were added: 3.2 µl 5X First Strand buffer, 0.8 µl 10mM dNTP and 1.6 µl 0.1M DTT. After incubation during 5 min at 42°C, 160U SuperScript II Reverse Transcriptase (Invitrogen) was added. The reaction mix was incubated at 42°C, 1 h. The enzyme was inactivated at 70°C, 15 min. The reaction mix was cool down on ice and 1.6U of *E. coli* RNase H was added and incubated at 37°C, 15 min. After the cDNA synthesis, the PCR reactions were done.

The primers JR869 Reverse and JR870 Forward were used and a PCR fragment of 340 bp was expected. In other PCR reactions, the primers JR 867 Reverse and JR 868 Forward were used and the expected product was 273-285 bp. The conditions for both PCR were the same. The PCR was performed in a 50 µl volume containing 50 ng of DNA, 1X buffer with 1.5 Mm MgCl₂, 200 µM of each dATP, dCTP, dGTP, and dTTP, 200 nM of each primer, and 5 U of Platinum Taq (Invitrogen). The PCR conditions
were: denaturation at 94°C, 4 min and 35 cycles at 94°C for 45 sec, 56°C for 45 sec and 72°C for 45 sec, following an extension at 72°C, 10 minutes. The sequences of the primers are: Primer JR 869 Reverse: 5’ CTG CTC CAG GAT TTC AAG GAC C 3’, Primer JR870 Forward: 5’ GCT GAT GGC TCT CCT ATG ACC 3’, Primer 867 Reverse: 5’ CTT GGG GTA TAC ATG CAC CTG 3’ and Primer 868 Forward: 5’ TCT CCC TGT GAG CAT AGG GTT GAC 3’. RNA treated with DNase I (without reverse transcription) was used in the PCR reactions to test for DNA contamination (negative controls).

**B-ii-b- Results.**

**Cloning of D17S796 region.**

It has been reported that the frequent LOH of a specific chromosomal marker is indicative of a closely linked tumor suppressor gene (TSG) (49, 50). In previous experiments we have shown LOH in the locus 17p13.2 using the microsatellite marker D17S796 in BP1E cells (Figure 1), and a detailed analysis of this region using Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi) has shown the KIAA0523 gene of approximately 200 Kb upstream and the AIP1 (aryl hydrocarbon receptor interacting protein-like 1) at 78 Kb downstream of this marker. Therefore, we pursued the search for a novel putative TSG in this region, which is approximately 940 bp. We have amplified this region by PCR from MCF-10F, BP1E and BP1E-17 neo (clone II-3), using the primers JR 853 and JR864. The sequences of these primers were obtained based on the sequence from the Genebank AC005668 (Figure 2).

```
JR 853
126421 aatgtttttt gattcaaaa acctcttttt tcagccttga tgactactgt gtaactactc
126431 tcttagtttt tacctccccct gaggggccag ctggtctaaa agtcactcag tttgctactg
126541 gatgagggcg ccagaggttg atgtgagaa ta ccaaggtgg agtgagggca ggcaggtggcc
126601 acctctggcca ggtctgcttc ccaggtgcct ct acaagcag ggtgcagcag tgagggggcag
126661 atggagggcg cacactctgt gcacagcagcag ccttcagatcg agctgtgaggt ccagagctcag
126721 gctgtgttttt acaagacata tcccacaagga cctgcocctcc ccagocactct ttctctcccat
126781 atgtctacaac ctcactactgct tgcacgcgtgg atgcctcccttg tgctctcttg agctgtgatgt
126841 tcctctctcgac ctctctcttct tcagacacag ggtatgtgag cagctgtgatg caaggggccag
D17S796 Rev
126901 gagatgtggtcttc gacagcagca tgatcattccttc atgtgtgatag ctgtgttcgcc cctttggtcc
126961 gataatgcca ggtgtgtctct ctctcaacagt acgcctttatg ctctctcttgag tacctgggtg
127021 atacatcggactcgtgtggtggtggtggt gttgtgtgtgt gttgtgtgtgtg tttgtgtgttgta aaaaaagtgc
127081 ctctcaggatt ttcagcagcag catttggttct cattggttctttcttttattgttg gcaagctaac
D17S796 Forw
127141 aagaaaaagcatgagacag catggtcggtg catgtaaccagctgcctcctgtag gcattaccac
127201 cagaggagcg tgggcaacag cttctctccct gcagcggtcct ccagagttccag tgcgcctcaatg
127261 tccacccctct gcagcagctga agagtttggttttaaggttggta gtaaagcggtgaccccttttaa
127321 atatatgtggt aatatggtggtggtggtggtggtggtgtggtggt gocctttcaatgttggtaa gatgtggtctcagctgtggct
127381 tgtgtgttgtgc tgttggtgtgc atagagagac gtcgacgtcga gatttttac tccagctttgc
JR864
127441 atgtttggtac atatattaaat agctattaaag tagaatattag tttgtgcacag aaaggagagag
127501 atgtctccttg agatattttc cttttcagaa gaagagcagcag gcattctctat gatttggaag
127561 atatattttt taattggtgta tttccccagc gagagcatcg agcgcgctac cttcccccag
Figure 2: Nucleotide sequence of the region D17S796 from the GeneBank AC005668. Part of the nucleotide sequence of the GeneBank AC005668 (Homo sapiens chromosome 17, clone hRPK.243-K-12), bases 126421 to 127440, is shown. The position of the primers D17S796 Reverse and D17S796 Forward used in the SSCP analyses are underline. Also, the position of the primers JR853 and JR864 used in the PCR amplification are shown.

The 940bp region was amplified from using a high fidelity Taq polymerase (Figure 3). The PCR products were cloned in the vector pCR2.1 and the mix was used to transform E. coli TOP 10 (Invitrogen). The plates were incubated at 37°C during 24 h for detecting the growth of blue and white colonies (Figure 4). Ten white colonies were picked up and grown on LB with kanamycin.

Figure 3: PCR products obtained from MCF-10F, BP1E and BP1E-17 neo clone II-3. The primers JR853 and JR864 were used to amplified 940 bp region contained D17S796. The total DNA from MCF10-F, BP1E and BP1E-17 neo (clone II-3) were used as template in three different PCR reactions. Approximately, 10 μl of each PCR reaction was run on 1.5% agarose gel. The 940 PCR products are shown. PCR reactions without DNA were also made as control (negative control). A 100 bp ladder was run in both gels as molecular weigh marker.

Figure 4: White and blue colonies growth on LB agar with kanamycin and X-gal. The PCR products obtained were cloned in the vector pCR2.1. The plasmids were used to transform E. coli TOP 10. White and blue colonies growth after 24h incubation at 37°C. This plate shown the colonies obtained using the PCR product amplified from BP1E. White colonies harbored plasmids with the 940bp insert. Blue colonies harbored pCR2.1 without insert. Ten white colonies were picked up for further analyses.
Sequence analyses.

The plasmids were prepared using the alkaline lysis method followed by PEG/NaCl precipitation. The inserts were sequenced at the Fox Chase Cancer Facility using M13 Reverse and Forward primers. The sequences from the different clones were compared using Seqencher and GCG programs. We found differences between MCF-10F, BP1E-17 neo and BP1E in a region of 160-174 bp that corresponds to the marker D17S796 that contains a zone of TG repetition. Using the multiple sequence alignment, we found that the sequences from BP1E differ with MCF-10F and BP1E-17neo basically in the number of TG repetition that lies between the primers D17S796 Reverse and D17S796 Forward (Figure 5). In the Figure 5, the alignment between the sequences from BP1E and MCF-10F/BP1E-17neo are shown. Approximately, 530 bp that include D17S796 region is shown. The only missing bases were found in this region.

Approximately, 550bp nucleotide sequence from BP1E with 19 TG repetitions is shown on the top. The nucleotide sequence from MCF-10F and BP1E-17neo with 25 TG repetitions is shown at the bottom. The positions of the primers D17S796 Reverse and D17S796 Forward are underline.

Sequences with different number of TG repetitions were found in the same cell line. In MCF-10F, five alleles were found (Figure 6). One colony from 9 (1/9) shown 18 TG repetition; three colonies (3/9) shown 20 TG repetitions; one colony (1/9) shown 21 TG repetition; two colonies (2/9) shown 24 and two others (2/9) shown 25 TG repetitions (Figure 6). In BP1E, two different alleles were found (Figure 6). Five out of 7 colonies analyzed (5/7) shown 20 TG repetitions and two (2/7) shown 19 TG repetitions (Figure 6). In BP1E-17 neo, five alleles were found (Figure 6). Two out of 10 analyzed colonies (2/10) shown 19 TG repetitions and 2 (2/10) shown 20 TG repetitions. Three colonies out of 10 analyzed (3/10) showed 22 TG repetitions, one (1/10) shown 23 and two (2/10) shown 25 TG repetitions (Figure 6).

Our PCR-SSCP analysis of MCF10F and BP1E cells revealed that in the process of chemical transformation loss of genetic material (LOH) in the chromosome 17 p13.2-p13.1 (D17S796), has been restored in the cells with abrogated transformation phenotype. The LOH found in BP1E using PCR-SSCP (Figure 1) is explained by the loss of the bigger PCR products. The expected sizes of the PCR...
products using D17S796 Reverse and D17S796 Forward are indicated in each case based on the sequences obtained (Figure 6). According to our previous results, one of these fragments, present in MCF-10F and BP1E-17 neo and absent in BP1 could be responsible for the reversion of the transformed phenotype of BP1E-17 neo. We found that BP1E has one allele with (TG)$_9$ and the other allele with (TG)$_{20}$ (Figure 6). The same fragments were also found in BP1E-17 neo. From the others three alleles present in BP1E-17 neo, only the one with 25 TG repetitions was also found in MCF-10F. It is possible that this fragment was responsible for the reversion of the phenotype of BP1E-17 neo.

<table>
<thead>
<tr>
<th>MCF-10F:</th>
<th>Number of colonies</th>
<th>Expected size of the D17S796 fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td>......CACC (TG)$_{19}$ TT AAAAA........</td>
<td>1/9</td>
<td>169 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{20}$ TT AAAAA........</td>
<td>3/9</td>
<td>164 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{21}$ TT AAAAA........</td>
<td>1/9</td>
<td>166 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{22}$ TT AAAAA........</td>
<td>2/9</td>
<td>172 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{23}$ TT AAAAA........</td>
<td>2/9</td>
<td>174 bp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BP1E:</th>
<th>Number of colonies</th>
<th>Expected size of the D17S796 fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td>......CACC (TG)$_{19}$ TT AAAAA........</td>
<td>2/7</td>
<td>162 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{20}$ TT AAAAA........</td>
<td>5/7</td>
<td>164 bp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BPE1-17 neo:</th>
<th>Number of colonies</th>
<th>Expected size of the D17S796 fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td>......CACC (TG)$_{19}$ TT AAAAA........</td>
<td>2/10</td>
<td>162 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{20}$ TT AAAAA........</td>
<td>2/10</td>
<td>164 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{21}$ TT AAAAA........</td>
<td>3/10</td>
<td>168 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{22}$ TT AAAAA........</td>
<td>1/10</td>
<td>170 bp</td>
</tr>
<tr>
<td>......CACC (TG)$_{23}$ TT AAAAA........</td>
<td>2/10</td>
<td>174 bp</td>
</tr>
</tbody>
</table>

We have performed a detailed computational analysis of the fragment with 25 TG repetitions isolated from MCF-10F and BP1E-17 neo. Sequence translation shown a putative peptide of 131 amino acids (Figure 7). The predicted amino acid sequence does not seem to share significant homology with any known protein.

We have performed a detailed computational analysis of the fragment with 25 TG repetitions isolated from MCF-10F and BP1E-17 neo. Sequence translation shown a putative peptide of 131 amino acids (Figure 7). The predicted amino acid sequence does not seem to share significant homology with any known protein.

Comparison of the DNA sequences of many different genes revealed certain similarities at the intron-exon junctions (53). The base sequence of the intron usually ends with AG. The
RUSSO. Jose

3' splice site sequence of most introns is similar to the sequence (C/U)_{10}N(C/T)AG/G such that most introns end in AG after a long stretch of pyrimidines (53). These sites appear to be very important for introns processing. We have found that in the sequence of this novel protein, the bases AG are present at the 3' putative splice site (Figure 7). In mammalian genes, polyadenylation sites are usually preceded by AATAAA or ATTAAA ~20 bases before the cleavage site and followed by a more weakly conserved GT-based motif. A putative polyadenylation site ATTAAA and GT region was found at the 3' end. These data suggest that this peptide could be the last exon of a novel protein. This peptide forms part of a larger novel protein and it is possible that also others part of this protein would be mutated in BP1E.

Sequences analyzes from BP1E, showed that translation of the sequence with 19 TG repetitions will give a putative peptide of 127 amino acids similar to the one found in BP1E and BP1E-17 neo (Figure 8). These peptides differ in four amino acids: VCVC (V: valine; Cysteine). The translation of the sequence with 20 TG repetitions, also found in BP1E, have a non-sense codon (stop codon) and a shorter peptide of 61 aminoacids.

![Figure 8: Alignment between the putative peptides from BP1E and the putative peptide from MCF-10F and BP1E-17 neo. The peptide alignment has shown that four aminoacids are missing in BP1E.](image)

The putative peptide was compared with known proteins using BLASTP. The predicted peptide was compared with a known proteins using BLASTP. The Swiss-Prot/TrEMBL primary accession numbers are shown (http://us.expasy.org/sprot). The putative peptide of 131 amino acids showed very low homology with known proteins (Table 1). The nuclear transition protein 2 (DNA-binding protein) shown 22.1% similarity and 15.3% identity.

<table>
<thead>
<tr>
<th>Protein</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitional protein 2 (Nuclear transition protein 2)</td>
<td>15.3%</td>
<td>22.1%</td>
<td>Q9N1A7</td>
</tr>
<tr>
<td>Similar to expressed sequence AW049250</td>
<td>14.5%</td>
<td>22.1%</td>
<td>Q8K116</td>
</tr>
<tr>
<td>Histamine H3 receptor (HH3R)</td>
<td>14.5%</td>
<td>22.1%</td>
<td>P58406</td>
</tr>
<tr>
<td>Netrin 4 (β netrin)</td>
<td>11.5%</td>
<td>17.6%</td>
<td>Q9JI33</td>
</tr>
</tbody>
</table>

Study of the expression of the region that contains D17S79.
As our “in silico” analyzes has shown a putative novel peptide in the region D17S796. The next step was to use RT-PCR to determine if there is expression of this region in MCF-10F, BP1E and BP1E-17neo cells. For this purpose total RNA was extracted from these cells and cDNA was synthesized using reverse transcriptase and oligo(dT). The cDNA obtained from the different cell lines were used in PCR reactions with specific primers for the putative coding region. Primers with annealing sites within the putative coding region were used in the amplifications and PCR products would be expected only if there is mRNA corresponding to this region (Figure 9). A fragment of 340 bp was obtained in the three cell lines when the primers JR869 and JR 870 were used (Figure 10). A fragment of approximately 285 bp was obtained when the primers JR867 and JR868 were used (Figure 10). These results showed that the region that harbors D17S796 is transcribed to mRNA in MCF-10F, BP1E and BP1E-17 neo.

---

**Figure 9:** Annealing sites of the primers used in RT-PCR experiments. The sequences of AC005668 (upper sequence) and MCF-10F (bottom sequence) are shown with the corresponding sites for the primers used in the RT-PCR: JR869, JR870, JR867 and JR868. Also the site for JR864 is shown. The beginning of the putative exon, the stop codon and the ATTAAA at the 3' are shown.
Also we found that the expressed-sequence tag EST 3179739 matches a region that is located 120 bp downstream of the cloned region (Figure 1: bases 127541 to 127544). The EST 3179739 sequence comes from a cDNA library from lung (tissue type: carcinoid). Expressed sequence tags (ESTs) are short (200-500bp) DNA sequences generated from the 3' and 5' ends of randomly selected cDNA clones and the purpose of EST sequencing is to rapidly scan for all the protein coding genes and to provide a tag for each gene on the genome. A 99% identity was found between both sequences using Blast (www.ncbi.nlm.nih.gov/blast/Blast.cgi).

These results strongly suggest that we have cloned the last exon of a novel protein.

![Image](image-url)
B-ii-c-Next proposed plan of research

We are planning to continue with the aims proposed using the following strategies:

1-The full-length cDNA corresponding to this novel gene will be isolated using "rapid amplification of cDNA ends" (RACE). RACE is a procedure for obtaining full-length cDNA copies of low abundance mRNAs. This technique achieves amplification of the region between a single short sequence in a cDNA and its unknown 3' or 5' end (51). The 3'RACE takes advantage of the natural poly(A) tail found in mRNA as priming site for PCR. The mRNAs are converted into cDNA using reverse transcriptase (RT) and an oligo-dT adapter primer. Specific cDNA is then amplified by PCR using a gene-specific primer (GSP) that anneals to a region of known exon sequences and an adapter primer that targets the poly(A) tail region. This permits the capture of the 3'-mRNA sequences located between the exon and the poly(A) tail. The 5'RACE uses an antisense gene specific primer for the synthesis of specific cDNA by reverse transcriptase. Prior to PCR, a TdT-tailing step attaches an adapter sequence to the unknown 5'sequences of the cDNA. Specific cDNA is then amplified by PCR using a GSP that anneals in a region of known exon sequences and an adapter primer that targets the 5' terminus. Products generated by the 3' and 5' RACE procedures may be combined to generate full-length cDNAs. RACE has been used for amplification and cloning of rare mRNAs (51). The full-length cDNA will be cloned in stable expression vector such as pSV2neo and transfection experiments will be performed.

2-We have hypothesized that this novel protein could be one of the proteins in the caspase cascade of apoptosis through CD95/Fas signaling. Northern blot analysis will be performed to study if apoptotic agents like human chorionic gonadotropin (hCG) induces an elevation of the expression of this novel gene. Our previous findings showed that hCG induces an elevation of the expression of apoptotic genes like Trpm2, ICE, TGF-β, p53 and p21 in MCF-10F (52). Using Northern blot analysis, information such as the size of the mRNA and expression level will be obtained.

---

**Figure 11. Chromosomal localization of D17S796 marker in chromosome 17. The position of the D17S796 in the chromosome 17 is shown. Some of the genes near D17S796 are described like the arylhydrocarbon receptor interacting protein-like and the death effector filament-forming Ced-4-like apoptosis protein (DEFCAP).**

1. Arylhydrocarbon receptor interacting protein-like/ Hypotetical protein FLJ10156
2. LOC 255868
3. KIAA0523
4. Ced 4-like apoptosis protein

---

3-A detailed analyses will be done to study mutations in other known genes localized at 17p13.1-13.2 near the D17S796 marker. The known genes located in this region are DEFCAP, KIAA0523, Aryl hydrocarbon receptor interacting protein-like1 and the hypothetical protein FLJ10156 (Figure 11). Different primers will be used to analyze putative mutations in them using PCR- SSCP analyses (Table 2).
RUSSO, Jose

These studies will help to determine regions near D17S796 marker that could also be mutated during the transformation of BP1E. Genomic DNA from MCF-10F, BP1E and BP1E-17 neo cells will be used. The availability of different markers on chromosome 17 will allow us to perform a detailed analysis for locating regions that have been retained after transfer of this chromosome to the cells of interest. This analysis will define regions near D17S796 that could also change. The sequences and characteristics of microsatellite oligonucleotide primers were obtained from GDB database (http://www.gdb.org) and Cooperative Human Linkage Center Web site (http://www.chlc.org). These experiments will allow us to determine all the gene(s) involved in the reversion of the phenotype in BP1-17 neo.

Table 2: Details of the primers that will be used in LOH mapping at 17p13.1-13.2.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Position</th>
<th>Size range</th>
<th>Primer sequence 5'—3'</th>
<th>Locus description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RH121127</td>
<td>17p13</td>
<td>273 bp</td>
<td>Forw: AATGTCCACAGAGCCATGAGT&lt;br&gt;Rev: TCCATTGTAAGAAACCTGGCTGT</td>
<td>Death effector filament-forming Ced-4-like apoptosis protein (DEFCAP)</td>
</tr>
<tr>
<td>Cda1bh06</td>
<td>17p13.2</td>
<td>110 bp</td>
<td>Forw: TCTTTCTGCTCTCAACTCTCT&lt;br&gt;Rev: AGGTAGTATATTCTCCTGGTC</td>
<td>KIAA0523 protein</td>
</tr>
<tr>
<td>D17S938</td>
<td></td>
<td>164-182 bp</td>
<td>Forw: CCAGAATTGCTACACCTAAAT&lt;br&gt;Rev: AACAGTCTCTNCTGCAGCAG</td>
<td></td>
</tr>
<tr>
<td>D17S1579</td>
<td></td>
<td>195 bp</td>
<td>Forw: CGATCCAAACTCTCTCCAG&lt;br&gt;Rev: TCTCTCTGAAAGTACAGACTGC</td>
<td></td>
</tr>
<tr>
<td>G42173</td>
<td>17p13.1</td>
<td>104 bp</td>
<td>Forw: GAGAGCTCTAAGTGCTCCTGC&lt;br&gt;Rev: TTTCATGCTGTCACCTC</td>
<td>Aryl hydrocarbon receptor interacting protein-like1 (AIPL1)</td>
</tr>
<tr>
<td>RH104452</td>
<td>17p13.2</td>
<td>146 bp</td>
<td>Forw: AGGGCAGAAATTGCCCTTGA&lt;br&gt;Rev: TTTCCTACCAGGGACTG</td>
<td>Hypothetical protein FLJ10156</td>
</tr>
<tr>
<td>D17S260</td>
<td></td>
<td>147 bp</td>
<td>Forw: AATGGCTCTCAAAGGAGATATG&lt;br&gt;Rev: CTCCCCAAACATGCTTCCTGC</td>
<td></td>
</tr>
<tr>
<td>D17S2057</td>
<td></td>
<td>260-261 bp</td>
<td>Forw: TTGTCTTGCAAACACTCTG&lt;br&gt;Rev: TGCAAACTCAGAGGAGG</td>
<td></td>
</tr>
<tr>
<td>D17S600</td>
<td></td>
<td>100-990 bp</td>
<td>Forw: GCTTTCTATTACTGCCCAGT&lt;br&gt;Rev: CTTCATCAGTCTGCCAGATTATAC</td>
<td></td>
</tr>
<tr>
<td>WIAF-1598</td>
<td>17p13.2</td>
<td>186-189 bp</td>
<td>Forw: CCTATATTAATTCAAGTGACACAT&lt;br&gt;Rev: CTCTGAACTCAAAGAATTTTCAG</td>
<td>KIAA0753 gene product</td>
</tr>
<tr>
<td>D17S1881</td>
<td></td>
<td>216-230 bp</td>
<td>Forw: CCAGTTTAAGGAGTTTGGC&lt;br&gt;Rev: TAGGAGCTGCAGTCTCG</td>
<td></td>
</tr>
<tr>
<td>SHGC-32950</td>
<td>17p13.2</td>
<td>125 bp</td>
<td>Forw: CGTGCACTCCCTCCAGTAGC&lt;br&gt;Rev: GTAGGGCAGAAATTGGCTCTGC</td>
<td>Hypothetical protein MCG14353</td>
</tr>
</tbody>
</table>
C- KEY RESEARCH ACCOMPLISHMENTS:

- We found a novel gene in the region D17S796. This gene is located at approximately 200 Kb of the KIAA0523 gene and 78 Kb of AIPL1 (aryl hydrocarbon receptor interacting protein-like 1).

- We have cloned the last exon of this novel gene and the sequence of this fragment will be used to select the full-length cDNA.

- RT-PCR experiments shown that this gene is expressed in MCF-10F, BP1E and BP1E-17 neo and the EST 3179739 matched the 3' end of this gene.

- We found difference in the region D17S796 between MCF10F, BP1E-17 neo and BP1E. The difference is in the number of the TG repeats in this region. We found that MCF10F and BP1E-17 neo has one allele in which the D17S796 region is 174 bp although, BP1E shown a shorter region.

D-REPORTABLE OUTCOMES

No publications have been originated from these data.

E-CONCLUSIONS

Loss of genetic material (LOH) in the chromosome 17 p13.2 at the microsatellite marker D17S796 has been identified in hepatocellular carcinoma (54), atypical ductal hyperplasia and in situ ductal carcinoma of breast (39, 55). Our results shown LOH at the same region in MCF-10F cells treated with the chemical carcinogen benz(a)pyrene (BP). Moreover, microcell-mediated transfer of an intact copy of human chromosome 17 inhibits the transformation phenotype of BP1E cells and PCR-SSCP analyzes show a restoration of the lost material in BP1E-17 neo. These experiments suggest the presence of a TSG in 17p13.2 near the marker D17S796. For this purpose a 940 bp region has been amplified and cloned. The BP1E cell has lost 10-12 bases consisting in a TG repetition. RT-PCR experiments have shown that this region is expressed in MCF-10F, BP1E and BP1E-17 neo. Also we have found that the expressed-sequence tag EST 3179739 matches with the 3'end of this region. Computational analyses show that the cloned fragment could be the last exon of a bigger peptide. The presence of a 3' splicing site in the putative introns and the ATTAATA region at the 3'end support this idea. The predicted amino acid sequence does not share significant homology with any known protein supporting the idea that this could be a novel protein. Further experiments will be done in order to clone the full-length cDNA and to study the regulation of the expression of this novel gene.

F. REFERENCES


