

***MRP8*, A New Member of ABC Transporter Superfamily, Identified by EST Database Mining and Gene Prediction Program, Is Highly Expressed in Breast Cancer**

Tapan K. Bera, Sanghyuk Lee*, Giuliana Salvatore, Byungkook Lee and Ira Pastan

Laboratory of Molecular Biology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

Contributed by I. Pastan. Accepted May 22, 2001

Abstract

Background: With the completion of the human draft genome sequence, efforts are now devoted to identifying new genes. We have developed a computer-based strategy that utilizes the EST database to identify new genes that could be targets for the immunotherapy of cancer or could be involved in the multistep process of cancer.

Materials and Methods: Utilizing our computer-based screening strategy, we identified a cluster of expressed sequence tags (ESTs) that are highly expressed in breast cancer. Northern blot and reverse transcriptase polymerase chain reaction (RT-PCR) analyses demonstrated the tissue specificity of the computer-generated cluster and comparison with the human genome sequence assisted in isolating a full-length cDNA clone.

Results: We identified a new gene that is highly expressed in breast cancer. This gene is expressed at moderate levels in normal breast and testis and at very low levels in liver,

brain, and placenta. The gene has two major transcripts of 4.5 kb and 4.1 kb. The 4.5-kb transcript is very abundant in breast cancer, and has an open reading frame of 1382 amino acids. The predicted protein sequence of the 4.5-kb transcript reveals that it has high homology with *MRP5*, a member of multidrug resistant-associated protein family (MRP). There are seven reported members in the MRP family; we designate this gene as *MRP8* (ABCC11). The 4.5-kb *MRP8* transcript consists of 31 exons and is located in a genomic region of over 80.4 kb on chromosome 16q12.1. The smaller 4.1-kb transcript of *MRP8* is found in testis and may initiate within intron 6 of the gene.

Conclusion: The selective expression of *MRP8* (ABCC11), a new member of ATP-binding cassette transporter superfamily could be a molecular target for the treatment of breast cancer.

Introduction

Expressed sequence tags (EST) are sequences derived from randomly selected clones from various cDNA libraries. Because each cDNA clone is generated from a transcript, the frequency and distribution of the many different transcripts in any given tissue depend on the tissue-specific activity of the expressed genes. Therefore ESTs provide a valuable source of information about the expression patterns of certain genes in different types of tissues (1,2). With the completion of the human draft genome sequence, much effort is now devoted to identifying new genes for a variety of purposes. Our laboratory is interested in identifying new genes that could be targets for the immunotherapy of prostate cancer or could be involved in the multistep process of prostate cancer. To accomplish this we have developed a computer-based screening

strategy that uses the EST database and generates clusters of ESTs that are expressed in normal prostate and/or prostate cancer but not in essential human tissues (3–7). We have now utilized the same computer screening strategy to identify EST clusters that are specifically expressed in normal breast and/or breast cancer.

Breast cancer is the most common type of epithelial cancer among women in the United States. More than 180,000 women are diagnosed with breast cancer each year. About one in eight women in the United States (approximately 12.8%) will develop breast cancer during her lifetime. At present there are no curative therapies available for breast cancer that has metastasized from its site of origination and there is urgent need for developing new targets for breast cancer therapy.

In this study, using a cluster of ESTs as a lead to screen and search human cDNA libraries, we identified a new gene expressed in many breast cancer samples. The protein it encodes is a new member of ATP-binding cassette (ABC) transporter superfamily and is a member of ABCC subfamily (ABCC11). Our data show that a variant of this gene is highly expressed in breast cancer and could be a potential molecular target for the treatment of breast cancer.

Address correspondence and reprint requests to: Ira Pastan, Laboratory of Molecular Biology, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Building 37, Room 4E16, 37 Convent Dr MSC 4255, Bethesda, MD 20892-4255. Phone: (301) 496-4797; fax: (301) 402-1344; e-mail: pasta@helix.nih.gov

*Permanent address: Department of Chemistry, Ewha Womans University, Seoul 120-750, Korea.

Materials and Methods

EST Database Mining and Computer Analysis

The ESTs from human tissues and tumor libraries were downloaded from the National Center for Biotechnology Information dbEST database. As of February 2001, this database contains over 2.5 million human EST sequences from 4673 libraries, of which 119,163 ESTs are from human breast and breast cancer. The EST sequences were clustered and sorted as described previously (8). The corresponding Unigene cluster for the individual candidate was identified and analyzed. Alignment information of the individual EST with the genomic sequence was obtained from "Golden Path" browser (<http://genome.cse.ucsc.edu/goldenPath/hgTracks.html>). The cDNA sequences determined experimentally were aligned with the genomic sequences using Blat, ClustalW, and Blast programs. GenomeScan and GraileXP gene prediction programs were used to predict genes and the PHD procedure in the PredictProtein program package was used to identify putative transmembrane regions.

RNA Dot Blots and Northern Blot Hybridization

RNA hybridization was performed on multiple tissue northern blots (MTN, Clontech, Palo Alto, CA, USA) and a Human Multiple Tissue Expression Array (Clontech, Cat# 7775-1) containing mRNA from 76 human tissues in separate dots. The 400-bp polymerase chain reaction (PCR) fragment generated by primer T352 and T354 was used as a probe. The sequence of the primers used in this study is described in Table 1. For a breast-specific probe, a 480-bp PCR fragment was generated using primer pair T377 and T389 based on the 5' DNA sequence of the 4.5-kb MRP8 cDNA. For hybridization, membranes were blocked for more than 4 hr, hybridized for 15 hr,

rinsed in $2\times$ SSC/0.1% SDS at room temperature, and finally washed twice in $0.1\times$ SSC/0.1% Sodium Dodecyl Sulfate at 65°C.

PCR Rapid Scan Gene Expression Panel

A rapid scan gene expression panel, containing PCR-ready first-strand cDNA from 24 different tissues (Cat# HSCA-101; OriGene, Rockville, MD, USA) was used as a template for PCR with a primer pair (T352 and T354) that should give a 400-bp fragment. For expression analysis of MRP8 in normal breast and breast cancer, we have used a Human breast cancer rapid scan panel (Cat# TSCE-101; OriGene), which contains PCR-ready first-strand cDNA from 12 normal and 12 breast cancer tissues. PCR composition and conditions used are according to the supplier's instructions.

RACE PCR Analysis for Identification of Full-Length cDNA

Rapid amplification of cDNA ends (RACE) was performed on Marathon Ready normal breast and testis cDNA (Clontech). Gene-specific primers T360 and T359 were used for the 5' and 3' RACE, respectively. The RACE PCR product was gel purified (QIAquick gel extraction, Qiagen, Santa Clara, CA, USA) and cloned into the pCR2.1 TOPO vector (Invitrogen, Carlsbad, CA, USA). Different clones were analyzed by restriction digestion using EcoRI restriction enzyme. The longest clones were sequenced using Perkin-Elmer's dRhodamine terminator sequencing kit (Perkin-Elmer Applied System, Warrington, UK).

In Vitro Translation

The in vitro translation of the 4.5- and 4.0-kb variants of MRP8 cDNA was examined in an in vitro transcription coupled translation system using T7 RNA polymerase and rabbit reticulocyte extract (TNT, Promega, Madison, WI, USA). 35 S-Met (ICN, Costa Mesa, CA, USA) was incorporated in the reaction for visualization of translated products. The reaction mixture was heated at 37°C in the presence of 5 M of urea and then analyzed under reducing condition on a polyacrylamide gel (7.5% Tris/Glycine, Bio-Rad) together with a pre-stained marker (Bio-Rad, Richmond, CA). The gel was dried and subjected to autoradiography.

Results

EST Database Mining for the Identification of Breast-Specific EST Clusters

As part of our search for genes expressed in breast cancer, we identified a cluster of ESTs, referred to as Br001, which consists of six ESTs; three ESTs (AW372855, AW372856, AW372862) are from a breast cancer library, and three from a normal prostate (AI401832, AI676121, BF447217) library (Fig. 1). By assembling the six ESTs, we obtained a 369-nucleotide sequence that contains a polyadenylation site (TATAAA). The cluster contains no ESTs from

Table 1. Sequences of primers used in PCR fragment generation

Primer Name	Primer Sequence
T352	TAC AGA TGT GGT GGA AAA CGG TGG
T354	CTC CAT GAA GTC TCC ACA TCT CCT
T357	GGA GAA CCA CAT GAT CAC TGA ATC CCA TGT C
T359	CCT AGT ACT GAT CAC CTG CGC ATC GCT
T360	AGC GAT GCG CAG GTG ATC ACT AGG
T377	GTC ACC TGG ATG GTG TAA AGG AAA CAA GAG
T389	AAG CGA CTC CGT AAG CTT TGG ATC ATG AGC

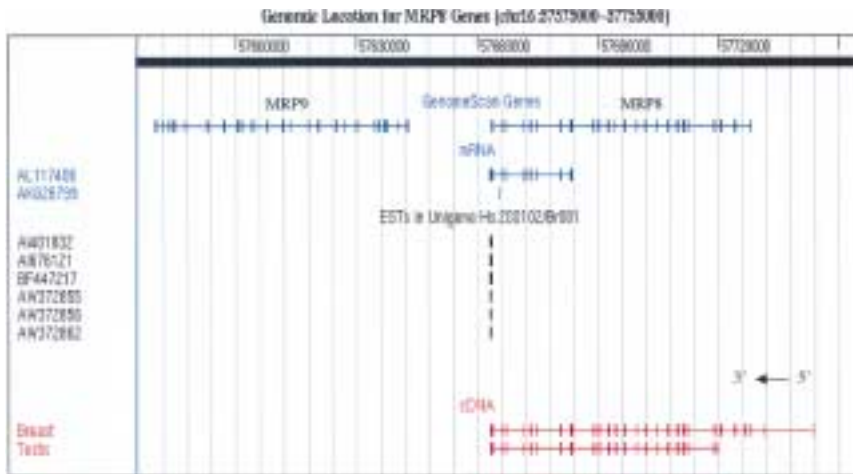


Fig. 1. Schematics showing the alignment of EST cluster *Br001* into chromosome 16: 57575000–57755000. GenomeScan gene finder program predicts two genes in this region, which we named as *MRP8* and *MRP9*. All ESTs from the cluster *Br001* aligned at the very 3' exon of *MRP8*. The alignments of the cDNA sequences derived from experimental clones are marked as breast and testis. The known genes reported in this region are the sequences from cDNA clone AL117406 and AK026799.

other tissues, which suggests that the *Br001* transcript might only be expressed in breast and prostate tissues.

Specificity Analysis by Dot Blot and Rapid Scan RT-PCR Analysis

To examine the computerized EST database analysis experimentally, we performed a mRNA dot blot (human multiple tissue expression array, Cat# 7775-1, Clontech) analysis using a PCR-generated probe (3' probe; Fig. 5) from the consensus EST sequence of cluster *Br001*. As shown in Figure 2A, among the 76 samples of normal and fetal tissues, *Br001* is detected strongly in human testis (F8), breast (F9), and weakly in liver (A9). A very weak signal is also detected in prostate (E8) and placenta (B8). To confirm

the dot blot result, we conducted a PCR using the human rapid-scan multi-tissue cDNA panel (Origene) with primers T352 and T354. This rapid scan panel consists of cDNAs derived from 24 different tissues and the cDNAs are serially diluted over a 4-log range and arrayed into a multi-well PCR plate. As shown in Figure 2B, the expected 400-bp PCR product was evident in testis and less intense in liver. Because this multi-tissue cDNA panel does not have any sample from human breast, we added corresponding amount of human breast cDNA (Clontech) in 1 of the 96 wells and ran the PCR at the same condition. As shown in Figure 2B, lane 13, a 400-bp PCR product is detected in the breast lane and the intensity of the band is comparable with the

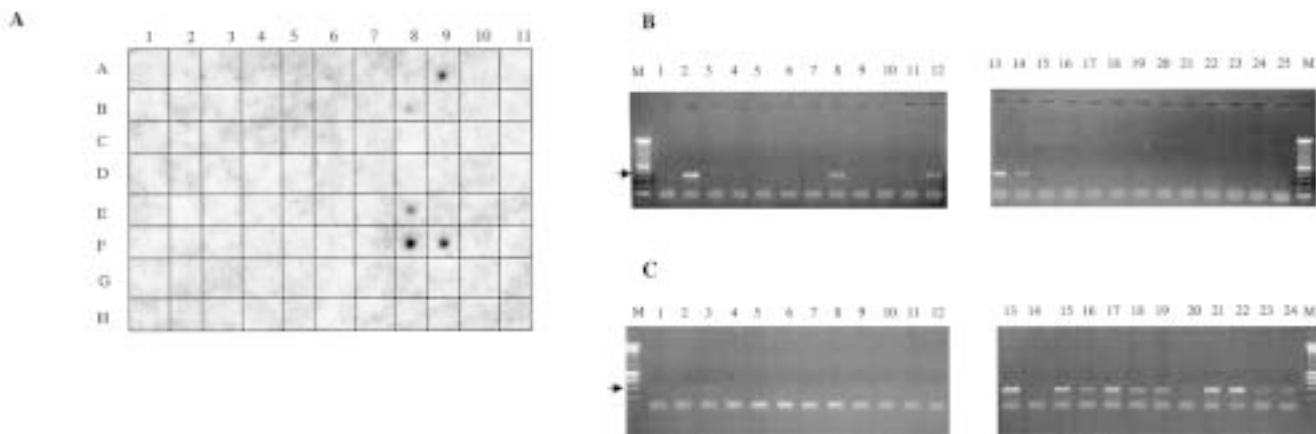


Fig. 2. Tissue distribution of *MRP8* mRNA expression. (A) RNA hybridization of a multiple tissue dot blot containing mRNA from 50 normal human cell types or tissues using a cDNA probe from the 3' end of the *MRP8* transcript. Strong expression in testis (F8) and breast (F9). There is a weak signal observed in liver (A9), prostate (E8), and placenta (B8). (B) PCR on cDNA from 24 different human tissues (Rapid Scan panel, Origene). The expected size of the *MRP8* PCR product is indicated by an arrow (400 bp). Lanes are 1, placenta; 2, testis; 3, stomach; 4, muscle; 5, small intestine; 6, lung; 7, colon; 8, liver; 9, spleen; 10, kidney; 11, heart; 12, brain; 13, breast; 14, fetal liver; 15, fetal brain; 16, bone marrow; 17, peripheral blood leukocytes; 18, skin; 19, prostate; 20, uterus; 21, ovary; 22, pancreas; 23, adrenal gland; 24, thyroid gland; and 25, salivary gland. (C) PCR on cDNA from 12 different normal breast (lanes 1–12) and 12 breast cancer (lanes 13–24) tissues.

band in the testis lane. This 400-bp product was also weakly detected in adult and fetal brain and liver as well as in placenta and prostate. The relative expression in prostate, placenta, brain, and liver is 10–100 times lower than in breast and testis (data not shown).

Analysis of the Genomic DNA to Identify Genes Around the Tissue-Specific EST Cluster

The EST cluster *Br001* maps on chromosome 16q12.1 and the accession number for the corresponding genomic contig is AC007600. In the “Golden Path” human genome browser, there is only one mRNA (accession number AL11706) reported as full-length, which is in fact a fully sequenced cDNA clone from a library made from human testis. There is no other known gene reported in this region (Fig. 1). We analyzed about 180 kb of genomic sequences around the cluster *Br001* for genes predicted to encode membrane proteins using different gene prediction programs. As shown in Figure 1, GenomeScan (9) predicted two large genes, both of which code for membrane proteins. One of the two genes includes the breast-specific cluster *Br001*. A BLAST search (10) with the deduced full amino acid sequence of the predicted gene against the GenBank nr database resulted in significant hits with ABC super family proteins (11), in particular with MRP5, which belongs to ABCC subfamily (12,13). Because *Br001* is a new member of ABCC subfamily and currently there are seven identified members of MRPs, we named this gene *MRP8*. Amino acid sequence analysis suggests that *MRP8* is a full transporter and has two conserved nucleotide binding domains and 12 putative transmembrane domains (12). Further analysis of the amino acid sequence reveals that *MRP8* has overall a 42% identity and 51% similarity with the *MRP5* sequence.

The other gene predicted by the GenomeScan program has homology with MRP groups of proteins and is located downstream of *MRP8*. We named this gene *MRP9* (ABCC12) and the characterization of this gene will be reported elsewhere.

Expression of *MRP8* in Normal Breast and Breast Cancer

To investigate whether *MRP8* is expressed in different normal breast and breast cancer samples, an RT-PCR analysis was carried out using the human breast cancer rapid-scan panel from Origene. In this panel, cDNAs from 12 different normal breast and 12 different breast cancer tissues are arrayed into a multi-well PCR plate. As shown in Figure 2C, using the PCR primer T352 and T354, the expected 400-bp PCR product was weakly detected in 8 out of 12 normal breast samples tested. Importantly, in 10 of the 12 breast cancer samples this 400-bp band was 10–100 times stronger than in normal breast.

Northern Analysis of *MRP8*

To determine the size of the *MRP8* transcript expressed in testis and breast tissue we used a multiple tissue northern blot (Clontech). As shown in Figure

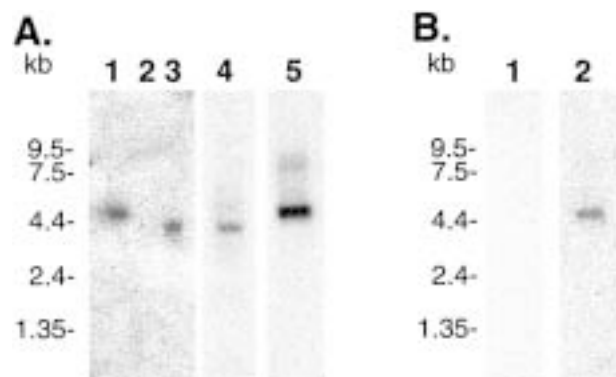


Fig. 3. Northern blot analysis showing expression and transcript sizes of *MRP8* in different normal tissues. PCR probe from the 3' end (A) and 5' end (B) of the *MRP8* cDNA was used for hybridization. Lanes are: for panel A, Breast, 1; Bone marrow, 2; Testis, 3; Liver, 4; and Breast cancer cell line HTB20, 5. For panel B, Liver, 1; and Breast cancer cell line HTB20, 2. The transcript expressed in Breast and HTB20 is 4.5 kb, whereas the transcript size in testis and liver is 4.1 kb.

3A, *MRP8* is detected in breast (lane 1), testis (lane 3), and liver (lane 4), and these differ in size. There is one distinct transcript of about 4.5 kb in the breast lane, whereas the transcript size in testis and liver is about 4.1 kb. In the rapid scan RT-PCR analysis, brain tissue was positive for *MRP8* expression. In the Northern blot analysis we are not able to detect any band in brain, probably because Northern analysis is much less sensitive than the RT-PCR. There is high expression of the *MRP8* transcript in the breast cancer cell line HTB20 (lane 5) and the transcript size is 4.5 kb, which is similar to the size observed in normal breast. The expression profile of *MRP8* in different tissues as determined by dot blot, Northern blot, and RT-PCR analysis is summarized in Table 2.

Full-Length cDNA Cloning and Analysis of the cDNA

To isolate the full-length *MRP8* cDNA, we performed a 5' and 3' RACE-PCR using normal breast cDNA (Marathon-Ready cDNA, Clontech) with the primers designed from the sequence predicted by the gene prediction program. The antisense primer T360 was used for 5' RACE and the sense primer T359 was used for 3' RACE reaction. Both primers contain the predicted sequence of exon five (exon 8 of the experimental clone, Fig. 4). The RACE-PCR analysis at the 5' end gave a 1-kb band, and the 3' RACE gave a 3.5-kb band. After subcloning and sequencing of the PCR products, we obtained the complete sequence of *MRP8* cDNA. Primers T377 and T357 were used to obtain a full-length *MRP8* clone from human breast cDNA. The full-length cDNA is 4538 bp long and has a long ORF of 1382 amino acids starting at bp 146 and ending at 4202. The calculated molecular weight of the protein encoded by the ORF is about 150 kDa. As shown in Figure 5, the full-length cDNA consists of 31 exons and is located in a genomic region of over 80.4 kb on chromosome 16q12.1. To determine if the exons and

Table 2. Summary of the expression of *MRP8* in different tissues

Tissue	RT-PCR (3' Primers)	Dot Blot (3' Probe)	Northern (3' Probe)	Northern (5' Probe)	EST (Number)
Cancer					
Breast	++++	ND	++++	+++	3
Normal					
Breast	++	+++	+++	++	0
Prostate	+	+	—	—	3
Testis	+++	++++	+++	—	0
Liver	++	++	+	—	0
Brain	+	—	—	—	0
Placenta	+	+	—	—	0

Probes: 3' probe can hybridize to both 4.5-kb and 4.1-kb variants of *MRP8*, whereas 5' probe is specific for 4.5-kb transcript. Signal intensity: +++++, high; +++, medium; ++, low; +, very low; —, undetectable; ND, not determined.

introns predicted by the GenomeScan program are correct, we compared the nucleotide sequences of the *MRP8* cDNA clone with the predicted exons. Out of 31 exons determined experimentally, 26 were predicted by the GenomeScan program, of which 23 were exactly the same, but 3 had incorrect boundaries (donor/acceptor site). There are 5 exons (1, 2, 4, 8, and 14) in the experimental clones (Fig. 4) that are not predicted by the program.

Long MRP8 Transcript Is Specifically Expressed in Breast Tissue

Several attempts to clone a 4.5-kb *MRP8* cDNA from liver and brain cDNA libraries failed, even though the

MRP8 cDNA is detected by RT-PCR analysis (Fig. 2A). To determine if the *MRP8* transcript is unique to the breast tissue at the 5' end, we performed Northern analysis using a radiolabeled DNA probe designed from the very 5' end (bp 1–480) of the *MRP8* cDNA. In this analysis a multi-tissue Northern blot containing 2 µg of mRNAs from eight essential human tissues including liver was used. For a breast sample, 2.5 µg of mRNA from breast cancer cell line HTB20 was analyzed. As shown in Figure 3B and summarized in Table 2, a 4.5-kb *MRP8* transcript is detected in the HTB20 (lane 2), but not in liver (lane 1). These data indicate that the 4.5-kb transcript of the *MRP8* is specifically expressed in the breast tissue.

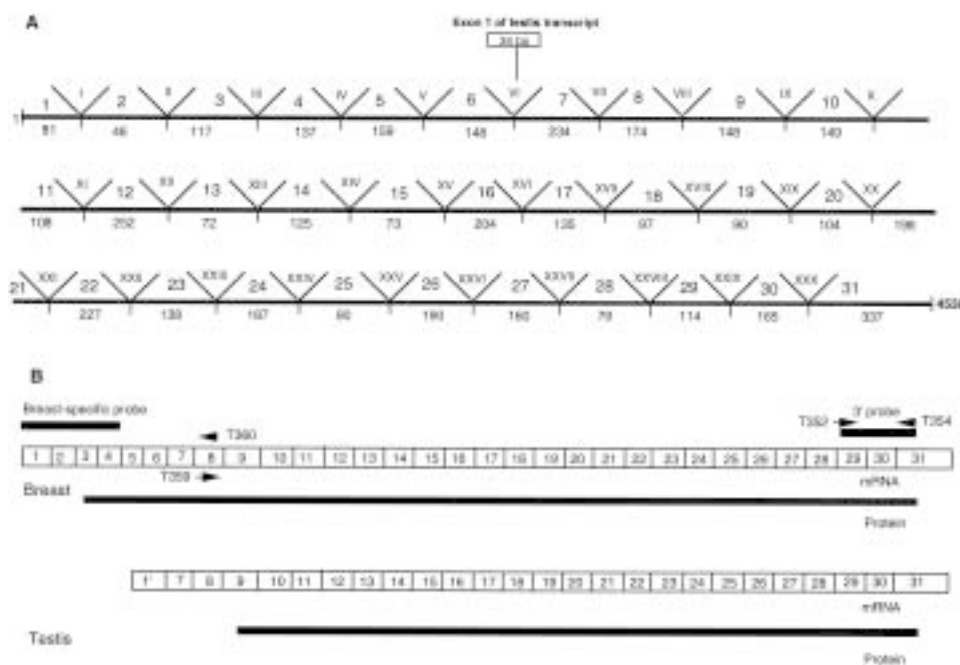


Fig. 4. Genomic organization and variants of *MRP8* transcript. (A) Exons and introns of 4538-bp *MRP8* transcript are arranged. Exons are numbered 1–31 and the length of each exon is indicated below in base pairs. The introns are labeled I-XXX. The start of the testis transcript within intron number VI is also shown. (B) Variants of *MRP8* transcript and predicted open reading frames. The predicted open reading frame for breast variant starts at exon 3 and ends at exon 31, whereas the open reading frame for testis variant starts at exon 9 and ends at exon 31. Both transcripts use the same reading frame. The region of cDNA sequences used for Northern hybridization are indicated as 3' probe and breast-specific probe. The primers T352 and T354 are used for rapid scan RTPCR analysis, and T359 and T360 are used for 3' and 5'RACE analysis, respectively.

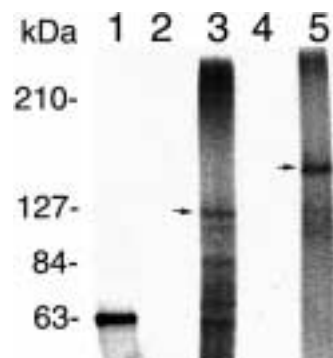


Fig. 5. Analysis of translated products of *MRP8* cDNA. Two variants of *MRP8* cDNAs were transcribed with T7-RNA polymerase and coupled translated with rabbit reticulocyte lysate in the presence of ^{35}S -methionine. The translated products were analyzed by SDS-PAGE and fluorography. Lane 1, positive control (Luciferase); lanes 2 and 4, vector control; lane 3, *MRP8* testis variant; lane 5, *MRP8* breast variant. The correct size protein product is indicated by an arrow.

Analysis of the MRP8 Transcript Expressed in Testis

Northern blot analysis of *MRP8* detects a 4.1-kb band in the testis lane, which is smaller than the band in the breast lane (Fig. 3). The difference in transcript size in two different tissues could be the result of either alternate splicing or a different transcription start site. To analyze the testis-specific transcript, we isolated the *MRP8* cDNA by two different approaches. First, we performed a RACE-PCR analysis using testis cDNA (Marathon-Ready cDNA, Clontech) with the primers designed from the sequence predicted by the gene prediction program as we have done earlier with breast cDNA. Second, we utilized a "High-throughput Longest-clone cDNA library screening" service from Origene to obtain an *MRP8* clone from the testis cDNA library. From both sources, we isolated an *MRP8* clone that is about 4.1 kb in size and matches with the transcript size we observed in the Northern blot analysis. Complete nucleotide sequence analysis of the 4.1-kb cDNA indicates that it has an ORF of 1064 amino acids starting at bp 457 and ending at bp 3648. As shown in Figures 4A and 4B, the cDNA consists of 26 exons, and exons 2–26 match exactly with the cDNA from the human breast. The first exon for the testis variant starts from a region of intron 6 of the breast variant (Fig. 4A). As a result, the translated product for the testis variant lacks the first 319 amino acids that are present in the *MRP8* clone from breast. This finding indicates that the smaller transcript size in testis is probably due to an alternate transcriptional start site and not due to an alternate splicing event.

In Vitro Transcription and Translation of the cDNA

The *MRP8* cDNA isolated from breast and testis has a predicted open reading frame of 1382 amino acids and

1063 amino acids, respectively. To determine the actual size of the protein encoded by the two different *MRP8* transcripts isolated from breast and testis, *in vitro* transcription and translation was performed using rabbit reticulocyte lysate system. SDS-PAGE analysis and fluorography of the translated product showed (Fig. 5) that the 4.5-kb cDNA from breast encodes a protein product of about 150 kDa in size (lane 5), whereas the protein product for the 4.1-kb transcript from testis is about 120 kDa in size (lane 3). The sizes of the protein products from the *in vitro* transcription and translation experiment agree with the predicted open reading frames of the cDNAs.

Discussion

We used a functional genomics approach to identify and clone a new member of the ABC superfamily of proteins and demonstrated by several experimental techniques that this newly identified gene is highly expressed in breast cancer, moderately expressed in breast and testis, and expressed at a very low levels in liver and brain. The deduced amino acid sequence of this gene has high homology with *MRP5* and we named the gene *MRP8* (ABCC11).

MRP8 Is a Member of ABC Superfamily and Belongs to ABCC Subfamily

The multidrug resistance/ATP-binding cassette (MDR/ABC) superfamily of membrane transporters is involved in energy-dependent transport of a variety of substrates across the membrane. In humans there are seven subfamilies (ABC-A to -F) based primarily on the sequence similarities. The sequence of *MRP8* is closely related to *MRP5* (Fig. 6), which belongs to the ABCC subfamily. It has an overall 42% identity and 51% similarity with the *MRP5* sequence. *MRP8* has 12 distinct membrane-spanning regions and two conserved nucleotide binding domains (NBD), indicating that *MRP8* is a full transporter. Although there are seven reported members of the MRP family, very little is known about the physiologic functions of these proteins. *MRP1* and *MRP2* are the most extensively studied members of this group. The tissue distribution of MRPs is different: *MRP1* and *MRP5* are ubiquitously expressed whereas *MRP2* and *MRP6* are predominantly expressed in liver and kidney (12, 13). Our studies indicate that the *MRP8* has restricted tissue distribution and is predominantly expressed in testis and breast. In addition, a shorter variant of *MRP8* is expressed in testis and liver. The testis variant of *MRP8* was characterized and found to lack the first four membrane spanning regions. The physiological functions of *MRP8* and its variant are yet to be determined. The region of the chromosome 16q12, where both *MRP8* and *MRP9* are located, has been implicated for the potential candidate gene(s) for paroxysmal kinesigenic choreoathetosis (PKC; 14) and infantile convulsions with paroxysmal choreoathetosis (ICCA; 15).

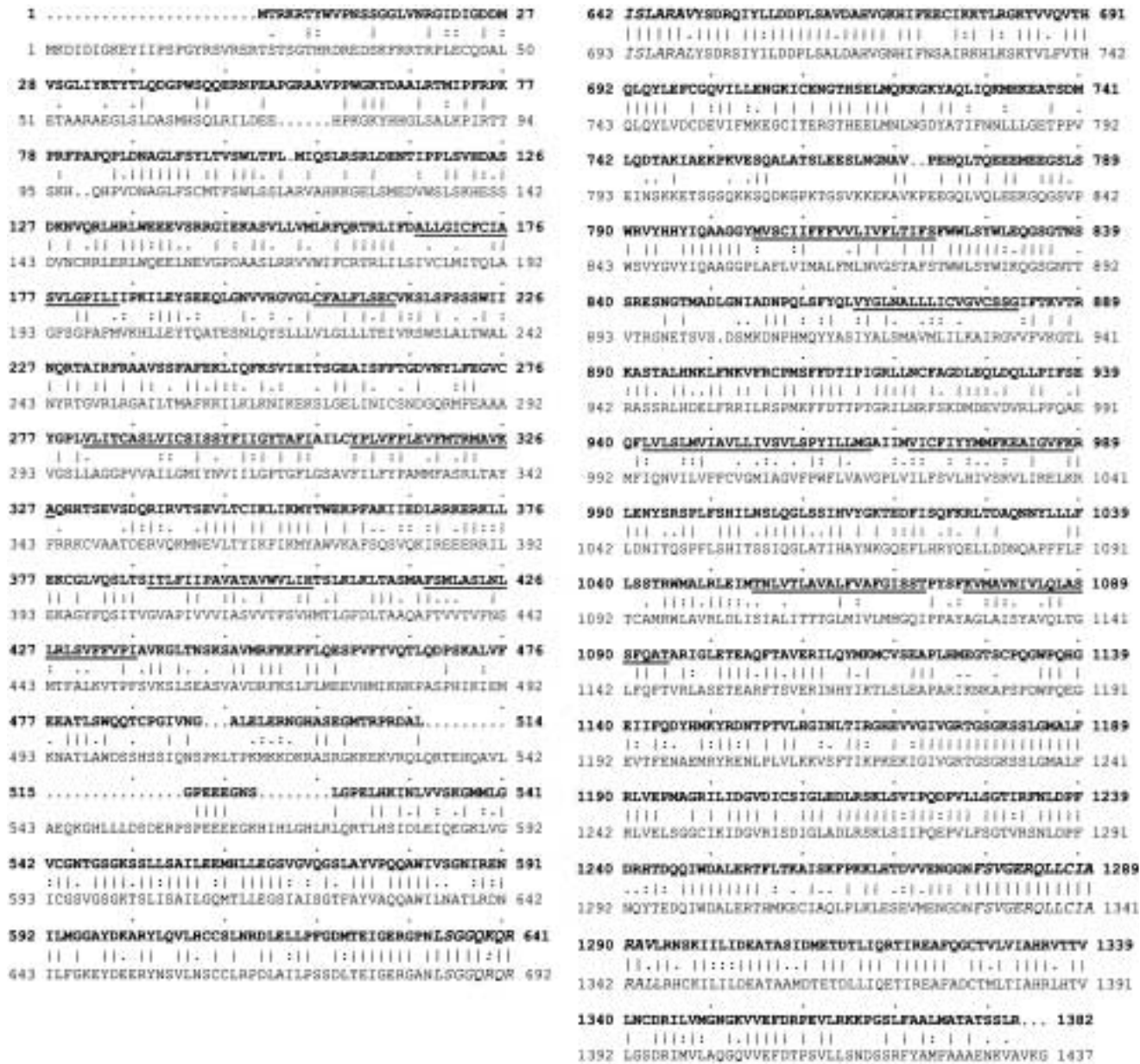


Fig. 6. Alignment of the MRP8 and MRP5 amino acid sequences. In the sequence MRP8 is shown in bold letters and the predicted transmembrane domains are underlined. Two conserved ABC signature motifs at amino acid positions 634 and 1278 of MRP8 are shown in italics.

Acknowledgments

We thank Drs. S. Ambudkar, J. Batra, K. Eglund, K. Santora, C. Wolfgang, and M. Gallo for their comments and R. Mann for editorial assistance. S. Lee is thankful for partial financial support from the Center for Cell Signaling Research at Ewha Womans University, Seoul 120-750, Korea.

References

1. Okubo K, Matsubara K. (1997) Complementary DNA sequence (EST) collections and the expression information of the human genome. *FEBS Lett.* 40: 225-229.

2. Burke J, Wang H, Hide W, Davison DB. (1998) Alternative gene form discovery and candidate gene selection from gene indexing projects. *Genome Res.* 8: 276-290.
3. Brinkmann U, Vasmatzis G, Lee B, Yerushalmi N, Essand M, Pastan I. (1998) PAGE-1, an X chromosome-linked GAGE-like gene that is expressed in normal and neoplastic prostate, testis, and uterus. *Proc. Natl. Acad. Sci. U.S.A.* 95: 10757-10762.
4. Brinkmann U, Vasmatzis G, Lee B, Pastan I. (1999) Novel genes in PAGE and GAGE family of tumor antigens found by homology walking in the dbEST database. *Cancer Res.* 59: 1445-1448.
5. Essand M, Vasmatzis G, Brinkmann U, Duray P, Lee B, Pastan I. (1999) High expression of a specific T-cell receptor gamma transcript in epithelial cells of the prostate. *Proc. Natl. Acad. Sci. U.S.A.* 96: 9287-9292.

6. Liu XF, Helman LJ, Yeung C, Bera TK, Lee B, Pastan I. (1999) *XAGE-1*, a new gene that is frequently expressed in Ewing's sarcoma. *Cancer Res.* **60**: 4752–4755.
7. Wolfgang CD, Essand M, Vincent JJ, Lee B, Pastan I. (2000) TARP: a novel protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the TCR γ locus. *Proc. Natl. Acad. Sci. U.S.A.* **97**: 9437–9442.
8. Vasmatazis G, Essand M, Brinkmann U, Lee B, Pastan I. (1998) Discovery of three genes specifically expressed in human prostate by expressed sequence tag database analysis. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 300–304.
9. Yeh R-F, Lim LP, Burge CB. (2001) Computational inference of homologous gene structures in the human genome. *Genome Res.* **11**: 803–816.
10. Altschul SS, Gish W, Miller W, Myers EW, and Lipman DJ. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
11. Allikmets R, Gerrard B, Hutchinson A, Dean M. (1996) Characterization of the human ABC superfamily: Isolation and mapping of 21 new genes using the expressed sequence tags database. *Hum. Mol. Genet.* **5**: 1649–1655.
12. Borst P, Evers R, Kool M, Wijnholds J. (2000) A family of drug transporters: The multidrug resistance-associated proteins. *J. Natl. Cancer Inst.* **92**: 1295–1302.
13. McAleer MA, Breen MA, White NL, Matthews N. (1999) PABC11 (also known as MOAT-C and MRP5), a member of the ABC family of proteins, has anion transporter activity but does not confer multidrug resistance when overexpressed in human embryonic kidney 293 cells. *J. Biol. Chem.* **274**: 23541–23548.
14. Tomita H, Nagamitsu S, Wakui K, Fukushima Y, Yamada K, Sadamatsu M, Masui A, Konishi T, Matsuishi T, Aihara M, Shimizu K, Hashimoto K, Mineta M, Matsushima M, Tsujita T, Saito M, Tanaka H, Tsuji S, Takagi T, Nakamura Y, Nanko S, Kato N, Nakane Y, Niikawa N (1999) Paroxysmal kinesigenic choreoathetosis locus maps to chromosome 16p11.2-q12.1. *American Journal of Human Genetics* **65**: (6) 1688–1697.
15. Lee WL, Tay A, Ong HT, Goh LM, Monaco AP, Szepietowski P (1998) Association of infantile convulsions with paroxysmal dyskinesias (ICCA syndrome): confirmation of linkage to human chromosome 16p12-q12 in a Chinese family. *Human Genetics* **103**: (5) 608–612.