POTE Paralogs Are Induced and Differentially Expressed in Many Cancers

Tapan K. Bera,[1] Ashley Saint Fleur,[1] Yoomi Lee,[1] Andre Kydd,[1] Yoonsoo Hahn,[1] Nicholas C. Popescu,[2] Drazen B. Zimonjic,[2] Byungkook Lee,[1] and Ira Pastan[1]

[1]Laboratories of Molecular Biology and [2]Experimental Carcinogenesis, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland

Abstract

To identify new antigens that are targets for the immunotherapy of prostate and breast cancer, we used expressed sequence tag and genomic databases and discovered POTE, a new primate-specific gene family. Each POTE gene encodes a protein that contains three domains, although the proteins vary greatly in size. The NH$_2$-terminal domain is novel and has properties of an extracellular domain but does not contain a signal sequence. The second and third domains are rich in ankyrin repeats and spectrin-like helices, respectively. The protein encoded by POTE-21, the first family member discovered, is localized on the plasma membrane of the cell. In humans, 13 highly homologous paralogs are dispersed among eight chromosomes. The expression of POTE genes in normal tissues is restricted to prostate, ovary, testis, and placenta. A survey of several cancer samples showed that POTE was expressed in 6 of 6 prostate, 12 of 13 breast, 5 of 5 colon, 5 of 6 lung, and 4 of 5 ovarian cancers. To determine the relative expression of each POTE paralog in cancer and normal samples, we employed a PCR-based cloning and analysis method. We found that POTE-2α, POTE-2β, POTE-2γ, and POTE-22 are predominantly expressed in cancers whereas POTE expression in normal tissues is somewhat more diverse. Because POTE is primate specific and is expressed in testis and many cancers but only in a few normal tissues, we conclude POTE is a new primate-specific member of the cancer-testis antigen family. It is likely that POTE has a unique role in primate biology. (Cancer Res 2006; 66(1): 52-6)

Introduction

As part of an effort to identify new targets for the immunotherapy of prostate cancer, we have identified several new genes that are expressed in prostate cancer but not in essential normal tissues (1–5). One of these genes is termed POTE because it is expressed in normal prostate, ovary, testis, and placenta as well as in prostate cancers. Subsequent studies showed that the POTE family of genes is found only in primates. At least 13 closely related paralogs have now been identified and these are located on eight different human chromosomes (6–8). Analysis of the synonymous and nonsynonymous mutations in the POTE genes indicates that the POTE gene amplification process began about 20 to 40 million years ago when the first POTE family member entered the primate genome (6).

The proteins encoded by the POTE genes all have an NH$_2$-terminal cysteine-rich domain followed by a series of ankyrin repeats and spectrin-like helices. The various POTE mRNAs are spliced in different ways so that the proteins encoded by these RNAs vary in size from 32 to 80 kDa (7). The POTE genes are named based on the chromosome on which they are located and if more than one gene is present on a chromosome, a Greek letter is added. POTE-21 was the first family member identified. The others are named POTE-2α, POTE-2β, POTE-2β′, POTE-2γ, and POTE-2δ, which are on chromosome 2, and POTE-8, POTE-13, POTE-14α, POTE-14β, POTE-15, POTE-18, and POTE-22 (7, 8). Examination of cells transfected with a POTE-21 cDNA tagged with green fluorescence protein at the COOH terminus showed that the POTE-21 protein was associated with the plasma membrane of the cell (7). Subsequent studies with other POTE paralogs have shown that they are also associated with the plasma membrane.[3] Initial studies on POTE expression did not distinguish among family members but clearly showed that POTE mRNA is found in a very limited number of normal tissues of the adult (6). These are prostate and testis in the male and ovary and placenta in the female. *In situ* hybridization showed that the RNA was found in prostate epithelial cells and spermatogonia in the testis (6).

In the current study, we show that POTE genes are expressed in a wide variety of human cancers (colon, lung, breast, ovary, and pancreas). To determine which paralog(s) is present in different tissues and cancers, we developed a method that combines reverse transcription-PCR (RT-PCR) with DNA sequencing. We find that specific paralogs are preferentially expressed in certain cancers and normal tissues. The absence of POTE expression in essential tissues in the adult makes it an attractive target for the immunotherapy of cancer. Moreover, the finding that the POTE gene has undergone rapid expansion in the primate genome combined with its expression in many cancers indicates that it is a useful marker for certain cancers and probably has a role in the special properties of these cells in humans and other primates.

Materials and Methods

Primers. Primers used in this study are T444, CAATGCCAGGAAGAT-GAATGTGCG, and T445, TCTCTGGCCGTCTGTCCAGATAGAT. The primers were synthesized by Lofstrand Laboratories (Gaithersburg, MD).

Cell lines and tissue samples. LNCaP and PC3 cells were acquired from the American Type Culture Collection (Manassas, VA) and cultured in our laboratory according to the instructions from the supplier. Cancer tissues were obtained from the cooperative Human Tissue Network, Southern Division (Philadelphia, PA).
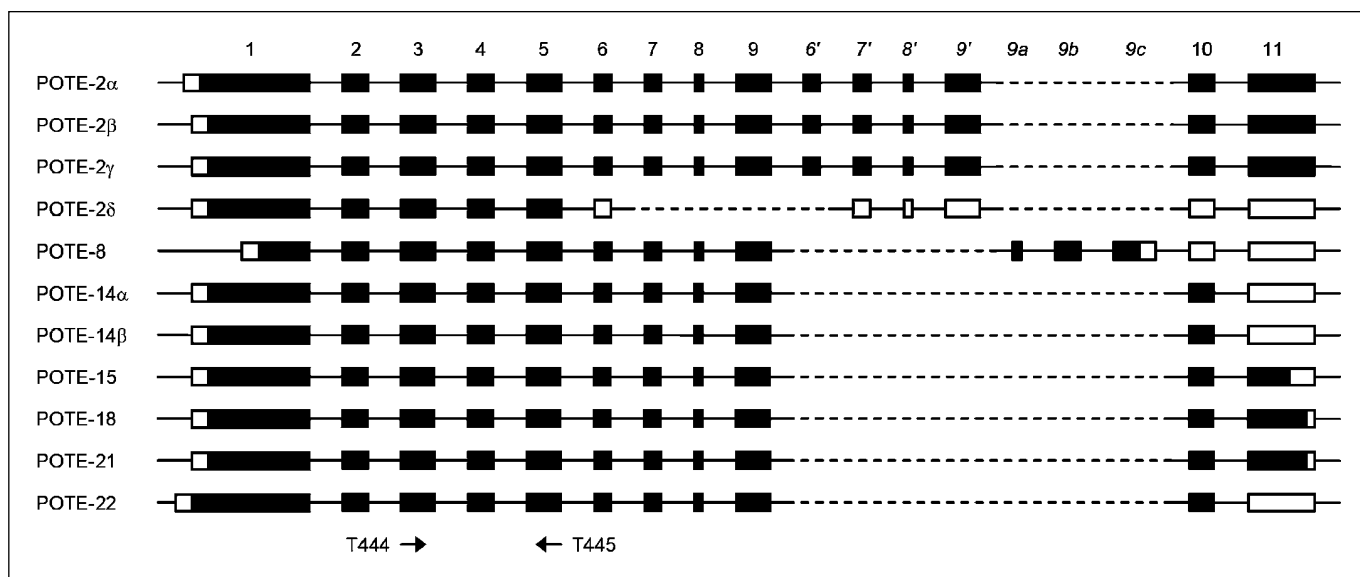
---

[3] Our unpublished data.

Figure 1. Schematic representation of genomic organizations of POTE family genes. The size of each box is roughly proportional to the actual size of the corresponding exon. Extra exons specific to some paralogs are italicized. The size difference among exon 1 is due to the different number of NH$_2$-terminal repeats. Introns are not to scale. *Dotted lines,* intron sequences that are not conserved in all paralogs. *Black boxes,* predicted open reading frames (ORF). The ORFs of POTE-2α, 2β, and 2γ are still open. Paralogs would produce proteins with different COOH termini owing to heterogeneity in the 3′ end of exon 11 and/or to mutations that lead to a frameshift or a premature termination. Alternatively spliced transcript variants would encode more isoforms. Locations of PCR primers, T444 and T445, are marked.

Expression of POTE by RT-PCR analysis. Total RNAs from different cancer cell lines and cancer tissues were isolated using TRIzol reagents (Invitrogen, Carlsbad, CA). First-strand cDNA was prepared from the isolated RNA using first-strand cDNA synthesis kit (Amersham, Piscataway, NJ) following the instruction of the manufacturer. Marathon-ready cDNAs from normal prostate, testis, ovary, and placenta were purchased from Clontech (Palo Alto, CA). PCR-ready cDNAs from breast and colon cancers were purchased from OriGene (Gaithersburg, MD) and BioChain Institute, Inc. (Hayward, CA), respectively. PCR was done on cDNA from different normal and cancer tissues following the instructions of the manufacturer using error-proof polymerase, Pfu Turbo (Stratagene, La Jolla, CA). The PCR conditions used are initial denaturation at 94°C for 1 minute, 30 cycles of denaturation at 94°C for 1 minute, annealing at 65°C for 1 minute, and elongation at 72°C for 2 minutes. The PCR primers used were T444 and T445 that should give a 386-bp fragment.

Cloning and analysis of PCR product. The PCR product obtained from the RT-PCR reaction was purified from low melting point agarose gel using a gel extraction kit (Qiagen, Chatsworth, CA) and cloned into pCR4Blunt-TOPO vector (Invitrogen). Positive colonies from each plate were selected randomly and plasmids were isolated from each colony. Plasmids with correct insert were sequenced and analyzed.

Computational discrimination of POTE paralogs. The initial attempt to identify clones by BLAT search of the human genome sequence gave spurious alignments, making it difficult to distinguish paralogs. To overcome this problem, we devised an in-house BLAST web server specialized for identification of POTE paralogs. The RT-PCR target sequences, excluding primer binding regions, were collected for each paralog and converted to a BLASTable database. The identification of a clone is based on its BLAST score against this database. The paralog giving the best score is assigned to the clone. The raw sequence data can be used without vector trimming or additional cleaning step.

Fluorescence *in situ* hybridization and spectral karyotyping. Labeled POTE probe (6) was hybridized on metaphase chromosome spreads prepared from exponentially growing LNCaP and PC3 prostate cancer cells. Hybridization, detection, primary imaging, rehybridization for spectral karyotyping, secondary imaging, and image analysis were done as previously described (6, 9).

Results

Relative expression of POTE paralogs in normal tissues. We reported earlier that at least five POTE paralogs and their splice variants are expressed in normal prostate and that the protein encoded by these cDNAs varies in size ranging from 32 to 80 kDa (7). To determine the relative expression of POTE paralogs in normal tissues and cancers, we designed a PCR primer pair from a specific region of POTE cDNA that gives a 386-bp product for POTE-21; these primers are located in exons 3 and 5 (Fig. 1). In the portion between the primers, there is a 1% to 5% sequence difference among the major splice forms of different paralogs although we encountered some minor splice variants with truncated or extended exon 3 or with an additional exon derived from an Alu in intron 3. We were able to distinguish paralogs based on the sequence difference between these primers. Figure 2 shows the specific nucleotide differences among the sequences of various paralogs that enable us to assign expression to a particular paralog. We PCR amplified that region using mRNAs from many different sources; TA cloned the product and sequenced 10 to 52 randomly selected cDNA clones from each sample. Nucleotide sequences from individual clones were analyzed using the Human Genome Browser (http://www.genome.ucsc.edu) and our custom POTE paralog identification server (see Materials and Methods). When the sequence shows the top-scoring match to a particular paralog, we conclude that the cDNA arises from the transcript of that specific paralog.

As described in Table 1, we tested 105 clones from normal prostate, 35 each from ovary and testis, and 42 from placenta. The pattern of expression varies markedly among these four normal tissues. The predominant paralogs expressed are, in prostate, POTE-2β, POTE-2γ, and POTE-22; in ovary, POTE-2α and POTE-2γ; in testis, POTE-15; and in placenta, POTE-2γ. This result

indicates POTE paralog expression is controlled in a tissue-specific manner.

POTE expression in cancer. We initially used RT-PCR to examine if members of the POTE family were expressed in cancer specimens. POTE expression was very frequent in many different cancer types. An initial experiment shown in Fig. 3 showed that POTE is expressed in 3 of 3 prostate cancers, 2 of 2 lung cancers, 1 breast cancer, 1 pancreatic cancer, and 2 of 2 colon cancers examined. Further analyses using cDNAs from many cancer samples show that POTE is frequently expressed in breast (12 of 13), prostate (6 of 6), lung (5 of 6), colon (5 of 5), and ovarian (4 of 5) cancer (data not shown). To identify which POTE paralogs were being expressed in cancers, we carried out the same analysis that was used in normal tissue (Table 1). In colon, breast, pancreatic, and lung cancers, the major paralogs expressed are on chromosome 2. POTE-2γ is expressed at the highest level with lower expression of POTE-2α in all of these cells. In addition, POTE-2β is expressed in lung and POTE-2δ in colon. The pattern of expression in ovarian cancer is different and there is expression of many paralogs (POTE-2α, POTE-2β, POTE-2γ, POTE-14α, POTE-14β, and POTE-22).

We also examined several normal and cancer cell lines for POTE expression by RT-PCR. POTE mRNA was detected in LNCaP, MCF7, HTB19, HTB20, and ZR-75-1, but not in PC3, MDA-MB-231, and two primary human fibroblast cell lines. This result indicates that the POTE expression in cell lines is not a result of the growth of cell lines in culture.

Fluorescent *in situ* hybridization analysis on prostate cancer cell lines LNCaP and PC3. The RT-PCR analysis shows that POTE is expressed in prostate cancer cell line LNCaP, but not in PC3. The relative expression analysis of POTE paralogs in LNCaP cell RNA suggests that the major paralogs are expressed from chromosomes 2 and 22. To investigate if there are deletions in those POTE loci in PC3 cells, we did fluorescent in situ hybridization (FISH) analysis on cell lines PC3 and LNCaP. FISH with the POTE probe on metaphase chromosome spreads from LNCaP cells produced on average 25 of 30 fluorescent signals per metaphase (data not shown). Spectral karyotyping of labeled metaphases enabled unequivocal identification of chromosomes with signals. The sites of POTE paralogs were identified on apparently normal chromosomes (chromosomes 2, 8, 13, 14, 15, 18, 21, and 22) as well as on three derivative chromosomes [del(2), del(13), and two copies of der(15)t(1;15)]. Hybridization of POTE probe on chromosomes from PC3 cell resulted, on average, in 26 to 29 detectable fluorescent signals (data not shown). As in the case of LNCaP cells, the number of signals reflected a ploidy
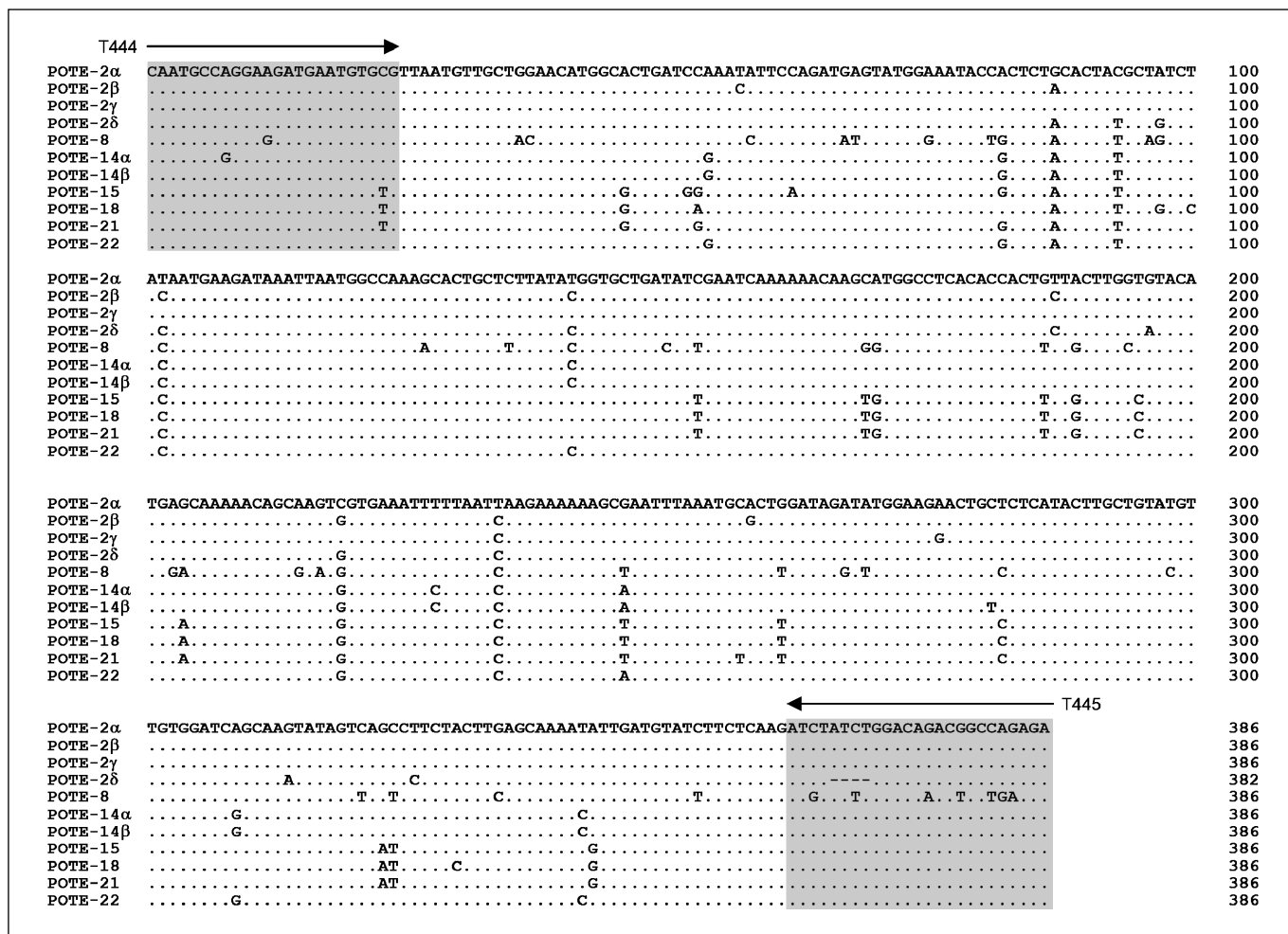
Figure 2. Multiple sequence alignment of the PCR amplified sequences from POTE family genes. Sequences to be amplified by RT-PCR of POTE paralogs are multiply aligned. Only the different nucleotides relative to the first sequence are shown. *Dashes,* missing residues. *Arrows,* PCR primer binding regions.

Table 1. Relative expression of POTE paralogs in target tissues

Tissue	Clones analyzed	POTE-2α (%)	POTE-2β (%)	POTE-2γ (%)	POTE-2δ (%)	POTE-8 (%)	POTE-14α (%)	POTE-14β (%)	POTE-15 (%)	POTE-18 (%)	POTE-21 (%)	POTE-22 (%)
Normal												
Prostate	105	1	28	28			1					42
Ovary	35	60	9	22								9
Testis	35	6	11	11					40	14	14	4
Placenta	42	17		64	14					5		
Cancer												
Prostate	36		42	28								30
Ovary	61	21	27	21			17	2				12
Lung	46	15	25	58								2
Pancreatic	15	13		87								
Breast	13	23		77								
Colon	38	13		55	32							
Cancer cell line												
LNCaP	61	22	24	46								8

status of this cell line. Seven of eight paralogs were found either exclusively on normal copies of their respective chromosomes or on both normal copies and derivatives of their respective carriers, suggesting that there are no obvious deletions of POTE loci in PC3 cell line.

Discussion

We have found that the POTE genes are expressed in a variety of human cancers (colon, breast, lung, ovary, pancreas, and prostate) but only in a limited number of normal tissues, one of which is testis. This indicates that POTE can be considered as a member of the cancer-testis antigen class (10).

Examination of the POTE paralogs expressed in normal tissues shows that there are clear patterns of expression with POTE-22 predominating in normal prostate, POTE-2α in ovary, POTE-15 in testis, and POTE-2γ in placenta. However, both POTE-2α and POTE-2γ were expressed at some level in all these tissues.

In contrast to limited expression in normal tissues, members of the POTE family are expressed in many of the commonly occurring cancers. We originally found that POTE was expressed in prostate cancer but we now have found that POTE is expressed in many of the most common and important cancers including lung, colon, breast, pancreas, and ovary. In these cancers, the POTE paralogs on chromosome 2 are most frequently expressed, with POTE-2γ and POTE-2α being most commonly found. This indicates that POTE expression in cancers is highly regulated.

Cancer-testis antigens are a distinct class of antigens that have restricted expression in essential tissues and aberrant expression in many cancers (10). Cancer-testis antigens are often expressed at higher levels in testis and placenta, which are known to express only low amounts of MHC class I molecules. Thus, expression of cancer-testis antigens in these normal tissues should not lead to T-cell activation and this makes these antigens attractive candidates for cancer vaccines (11).

Although cancer-testis antigens have been known for over a decade, no function has been described for them in the literature. Frequently, the genes for these proteins are located on the X chromosome. There is speculation that some cancer-testis gene products are transcriptional factors but there is no direct experimental evidence supporting that concept. Two of the previously reported cancer-testis antigens, SCP-1 and CT9, have small basic domains and several conserved motifs, which are characteristics of DNA-binding proteins (12, 13). POTE genes are different; there are many paralogs which reside in several chromosomes and they encode proteins with ankyrin and spectrin motifs.

One mechanism of gene activation is dependent on gene rearrangements placing new promoters close to protein-encoding regions of genes. Because POTE is expressed in the prostate cancer cell line LNCaP but not in PC3 and because PC3 cells are known to have many chromosomal abnormalities, we decided to analyze the POTE loci in this cell line. However, we were unable to find any gene rearrangements involving POTE, indicating that POTE activation in cancer is likely due to another mechanism. On the other hand, we cannot rule out the possibility that micro-deletions of the size below the sensitivity of spectral karyotyping may have eliminated a fraction of copies of some of the POTE paralogs. It is possible that some of the changes involving chromosomal regions other than those harboring the POTE gene family may have affected yet unknown functional elements driving and/or controlling expression of some or all POTE paralogs.
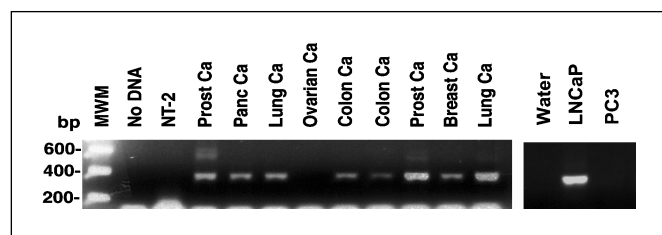

Figure 3. RT-PCR analysis on RNA. RT-PCR analysis on RNA from different cancer tissues and cell lines. The expected size of the PCR product is 386 bp.

Acknowledgments

References

1. Brinkmann U, Vasmatzis G, Lee B, Pastan I. 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 1998;95: 10757–62.

2. Essand M, Vasmatzis G, Brinkmann U, Duray P, Lee B, Pastan I. High expression of a specific T-cell receptor γ transcript in epithelial cells of the prostate. Proc Natl Acad Sci U S A 1999;96:9287–92.

3. Wolfgang CD, Essand M, Vincent JJ, Lee B, Pastan I. TARP: a novel protein expressed in prostate and breast cancer cells derived from an alternate reading frame of the TCRγ locus. Proc Natl Acad Sci U S A 2000;97:9437–42.

4. Olsson P, Bera TK, Essand M, et al. GDEP, a new gene differentially expressed in normal prostate and prostate cancer. Prostate 2001;48:231–41.

5. Bera TK, Maitra R, Iavarone C, et al. PATE, a gene expressed in prostate cancer, normal prostate, and testis, identified by a functional genomic approach. Proc Natl Acad Sci U S A 2002;99:3058–63.

6. Bera TK, Zimonjic DB, Popescu NC, et al. POTE a highly homologous gene family located on numerous chromosomes and expressed in prostate, ovary, testis, placenta and prostate cancer. Proc Natl Acad Sci U S A 2002;99:16975–80.

7. Bera TK, Huynh N, Maeda H, Sathyanarayana B, Lee B, Pastan I. Five POTE paralogs and their splice variants are expressed in human prostate and encode proteins of different lengths. Gene 2004;337: 45–53.

8. Hahn Y, Bera TK, Pastan IH, Lee B. Duplication and extensive remodeling shaped POTE family genes encoding proteins containing ankyrin repeat and coiled coil domains. Gene. In press 2005.

9. Zimonjic DB, Rezanka L, DiPaolo JA, Popescu NC. Refined localization of the erbB-3 proto-oncogene by direct visualization of FISH signals on LUT-inverted and contrast-enhanced digital images of DAPI-banded chromosomes. Cancer Genet Cytogenet 1995;80:100–2.

10. Ono T, Kurashige T, Harada N, et al. Identification of proacrosin binding protein sp32 precursor as a human cancer/testis antigen. Proc Natl Acad Sci U S A 2001;98: 3282–7.

11. Moingeon P. Cancer vaccines. Vaccine 2001;19: 1305–26.

12. Jones MH, Numata N, Shimane M. Identification and characterization of BRDT: a testis-specific gene related to the bromodomain genes RING3 and *Drosophila fsh*. Genomics 1997;45:529–34.

13. Meuwissen RL, Offenberg HH, Dietrich AJ, Riesewijk A, van Iesel M, Heytin C. A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. EMBO J 1992;11:5091–100.