Identification of a Pseudogene That Can Masquerade as a Mutant Allele of the PTEN/MMAC1 Tumor Suppressor Gene

Young E. Whang, Xinyi Wu, Charles L. Sawyers*

Alterations of the candidate tumor suppressor phosphatase gene PTEN/MMAC1 on chromosome 10q23 have been implicated in a wide range of cancers, including glioblastoma, melanoma, and cancers of the breast, kidney, and endometrium (1–4). In addition, germ-line mutations in PTEN/MMAC1 have been described in cancer-prone conditions such as Cowden’s syndrome (5). In the midst of efforts to determine the frequency of PTEN/MMAC1 abnormalities in prostate cancer, we discovered a pseudogene closely related to PTEN/MMAC1 that can be mistakenly identified as a mutant form of PTEN/MMAC1.

Oligonucleotide primers specific for the initiating ATG codon in exon 1 (open reading frame [ORF]-5’) and the stop codon 404 in exon 9 (ORF-3’) of PTEN/MMAC1 were designed to amplify the full-length coding region from RNA extracted from human prostate cancer cell lines and xenografts by reverse transcription–polymerase chain reaction. Surprisingly, products of the appropriate size (1226 base pairs [bp]) were amplified from all cell lines tested, including those known to have homozgyous deletions of the entire PTEN/MMAC1 locus such as in the PC3 cell line (1). To investigate this anomaly, we performed polymerase chain reactions (PCRs) in the absence of reverse transcription and used RNA from the PC3 cell line as well as from DU145, a cell line known to express PTEN/MMAC1 messenger RNA (mRNA) (1). A PTEN/MMAC1 PCR product was obtained in both cases, but not if the RNA was pre-treated with deoxyribonuclease (DNase) I (Fig. 1, A). DNase I treatment, however, did not abrogate the ability to amplify the fragment of correct size after reverse transcription reaction in DU145 cells, indicating that PTEN/MMAC1 mRNA is expressed. These results could be explained by contamination of the RNA preparations with genomic DNA that contains a sequence homologous to PTEN/MMAC1. Indeed, the same PCR product was obtained by use of genomic DNA as a template from cells with an intact PTEN/MMAC1 gene (DU145 and LAPC-4) and those with homozygous deletions of PTEN/MMAC1 (PC3 and LAPC-9) (Fig. 1, A). Because the PTEN/MMAC1 sequence corresponding to these primers is separated by approximately 200 kilobases (kb) in the PTEN/MMAC1 genomic locus (1,2), we suspected that a highly processed PTEN/MMAC1 homologue or pseudogene exists.

The identity of the PTEN/MMAC1-related gene was determined by subcloning and partial sequencing of PCR products amplified from several cell lines. In all cases, an identical sequence was obtained that matched the coding region of PTEN/MMAC1 with the exception of eight nucleotide changes in the first 450 nucleotides (see Fig. 2). A search for sequences related to the putative PTEN/MMAC1 pseudogene in the GenBank database revealed a perfect match with a region of genomic sequence from the human T-cell receptor beta genomic locus duplicated into chromosome 9 (GenBank accession AF029308) (6). An identical sequence named PTEN2 has also been recently submitted (GenBank accession AF017999). Upon further inspection, the genomic sequence contains a 1.2-kb region with 98% homology to the entire coding region of PTEN/MMAC1. Of note, a missense mutation is present that changes the initiator methionine ATG to AGG. A restriction map of this locus revealed two HindIII sites and two EcoRI sites flanking the region of PTEN/MMAC1 homology; the HindIII sites and the EcoRI sites are separated by 3.9 kb and 11.7 kb, respectively. This is consistent with Southern blot analysis of HindIII- or EcoRI-digested genomic DNA that demonstrated bands of 3.9 kb or 11.7 kb, respectively, in all tumor cell line samples tested, including those with complete deletion of the PTEN/MMAC1 genomic locus (PC3) (Fig. 1, B). A 4.7-kb band (EcoRI) or a 2.2-kb band (HindIII) corresponding to PTEN/MMAC1 is present only in tumors with an intact PTEN/MMAC1 gene (DU145).

Pseudogenes are nonexpressed, intronless mRNA-like sequences of genomic DNA derived from the processed RNA of an expressed gene, and they frequently contain a number of conserved point mutations (7). The fact that this PTEN/MMAC1-related sequence displays all of these characteristics firmly establishes that it encodes a pseudogene. Methods to characterize the status of the PTEN/MMAC1 gene in tumor samples by reverse transcription–PCR and sequencing or by protein truncation in vitro transcription/translation assays (1) may give misleading results because small amounts of genomic DNA are generally present in RNA prepared by standard extraction methods. In our experience, the PTEN/MMAC1 pseudogene is routinely amplified from such samples with the use of ORF-5’ and ORF-3’ primers in the absence of reverse transcription. Indeed, using exon-based primers, Rhee et al (8) recently reported missense mutations in the PTEN/MMAC1 mRNA in breast cancer. Of the 11 variant sequences, 10 are identical to substitutions in the pseudogene and most likely represent the pseudogene rather than mutations present in the PTEN/MMAC1 mRNA. Kim et al. (9) also isolated by reverse transcription–PCR a complementary DNA clone with a sequence identical to that of the PTEN/MMAC1 pseudogene reported...
here and concluded that this sequence represents a novel homologue. However, our data call their conclusion into question.

In summary, studies that rely on reverse transcription–PCR of mRNA for a gene of interest can give misleading results unless rigorous controls, such as no reverse transcription or DNase treatment, are performed to exclude the possibility of amplifying a pseudogene.

Notes

Supported by grants from the James S. McDonnell Foundation, the Margaret Early Trust, and CaPCURE. C. L. Sawyers is a Scholar of the Leukemia Society of America. Y. E. Whang is supported by an American Society of Hematology Scholar Award.

Manuscript received December 1, 1997; revised March 3, 1998; accepted March 23, 1998.