BMC Cancer

BioMed Central

Research article

Open Access

Characterization of human mesothelin transcripts in ovarian and pancreatic cancer

Zhanat E Muminova, Theresa V Strong and Denise R Shaw*

Address: Division of Hematology and Oncology, Department of Medicine, and the Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, 35294-3300, USA

Email: Zhanat E Muminova - zhana@uab.edu; Theresa V Strong - tvstrong@uab.edu; Denise R Shaw* - dshaw@uab.edu

* Corresponding author

Published: 12 May 2004

BMC Cancer 2004, **4**:19

Received: 03 February 2004

Accepted: 12 May 2004

This article is available from: http://www.biomedcentral.com/1471-2407/4/19

© 2004 Muminova et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Mesothelin is an attractive target for cancer immunotherapy due to its restricted expression in normal tissues and high level expression in several tumor types including ovarian and pancreatic adenocarcinomas. Three mesothelin transcript variants have been reported, but their relative expression in normal tissues and tumors has been poorly characterized. The goal of the present study was to clarify which mesothelin transcript variants are commonly expressed in human tumors.

Methods: Human genomic and EST nucleotide sequences in the public databases were used to evaluate sequences reported for the three mesothelin transcript variants *in silico*. Subsequently, RNA samples from normal ovary, ovarian and pancreatic carcinoma cell lines, and primary ovarian tumors were analyzed by reverse transcription-polymerase chain reaction (RT-PCR) and nucleotide sequencing to directly identify expressed transcripts.

Results: *In silico* comparisons of genomic DNA sequences with available EST sequences supported expression of mesothelin transcript variants 1 and 3, but there were no sequence matches for transcript variant 2. Newly-derived nucleotide sequences of RT-PCR products from tissues and cell lines corresponded to mesothelin transcript variant 1. Mesothelin transcript variant 2 was not detected. Transcript variant 3 was observed as a small percentage of total mesothelin amplification products from all studied cell lines and tissues. Fractionation of nuclear and cytoplasmic RNA indicated that variant 3 was present primarily in the nuclear fraction. Thus, mesothelin transcript variant 3 may represent incompletely processed hnRNA.

Conclusion: Mesothelin transcript variant 1 represents the predominant mature mRNA species expressed by both normal and tumor cells. This conclusion should be important for future development of cancer immunotherapies, diagnostic tests, and gene microarray studies targeting mesothelin.

Background

Mesothelin is a glycosylphosphatidylinositol (GPI)-anchored cell-surface glycoprotein expressed at low levels by a restricted set of normal adult tissues but aberrantly expressed by ~70% of human ovarian epithelial tumors including up to 100% of serous papillary ovarian cancers, as well as by significant proportions of pancreatic adenocarcinomas, endometrioid uterine adenocarcinomas,

mesotheliomas, and squamous cell carcinomas of the esophagus, lung and cervix [1-7]. Full length mesothelin (~69 kD) can be proteolytically cleaved to release a ~33 kD soluble protein corresponding to megakaryocyte potentiating factor (MPF) [8,9]. The biologic functions of mesothelin and MPF remain speculative. Mutant mice with targeted mesothelin gene inactivation are normal, exhibiting no apparent anatomic, hematologic or reproductive abnormalities [10]. Analysis of the mesothelin protein sequence yields no strong homologies to known protein functional domains, beyond a C-terminal GPI-anchor motif. Mesothelin has been very recently reported to bind to CA125/MUC16 also commonly expressed on the surface of ovarian tumor cells [11], suggesting that mesothelin might play a role in heterotypic cell adhesion and metastatic spread of ovarian cancer, but data to support this idea are lacking for most human tumors.

Despite limited understanding of mesothelin's biological function, its restricted expression by normal tissues combined with frequent abundant expression by tumors has suggested several applications in clinical oncology. Circulating mesothelin/MPF may have diagnostic potential in mesothelin-positive malignancies [12], and tumor cell-associated mesothelin is being used as a target in ongoing clinical trials of passive immunotherapy using immunotoxins [13,14]. Mesothelin may also be a viable antigen for tumor vaccine therapies (unpublished data presented by E. Jaffee, 10th Annual SPORE Investigators Workshop, July 15, 2002). Further, mesothelin appears to be an important target in the evolving arena of cancer genetic profiling for several major tumor types [4,5,15-17]. Rational design and development of these potential clinical applications would be facilitated by a clear understanding of mesothelin gene and protein expression by both normal cells and tumors.

To date, three variants of human mesothelin transcripts have been reported: variant 1 (GenBank accession NM_005823) encoding MPF [8,9]; variant 2 (NM_013404) encoding mesothelin [18]; and variant 3 (AF180951), a partial alternatively spliced cDNA with a disrupted GPI-anchor motif [19]. Of note, each transcript variant has regions of unique oligonucleotide sequences that could result in differential results in genetic-based expression studies and further encode proteins with unique peptides. It is of particular relevance for the present study that all three mesothelin transcript variants have been reported to be expressed by human cancer cells. Nevertheless, there is a remarkable lack of studies directly investigating relative expression levels of mesothelin transcript variants in normal and malignant tissues, even though multiple gene expression studies of tumor expression profiles conducted to date have included putative mesothelin-specific cDNA or oligonucleotide probes [4,5,15-17]. The present study provides evidence for expression of human mesothelin transcript variants 1 and 3, but not variant 2, in a variety of human tissues and cell lines. Moreover, variant 1 appears to be the predominant mature mesothelin transcript in all studied specimens, including ovarian and pancreatic adenocarcinomas.

Methods

Cell lines

The human ovarian (SKOV-3, Caov-4, OVCAR-3), pancreatic (PANC-1, AsPC-1) and cervical (HeLa) tumor cell lines were obtained from the American Type Culture Collection (Rockville, MD). UCI 101 human ovarian adenocarcinoma cell line was kindly provided by Dr. David T. Curiel (Gene Therapy Center, University of Alabama at Birmingham). SKOV-3 cells were grown in McCoy's 5A medium with 2 mM L-glutamine and 10% FBS. Caov-4, PANC-1 and HeLa were grown in DMEM with 4.5 g/L glucose, 2 mM L-glutamine and 10% FBS. OVCAR-3 were grown in RPMI-1640 medium with 2 mM L-glutamine, 1.5 g/L sodium bicarbonate, 4.5 g/L glucose, 10 mM HEPES, 1 mM sodium pyruvate, 0.01 mg/ml bovine insulin and 20% FBS. AsPC-1 and UCI 101 were grown in RPMI-1640 medium with 2 mM L-glutamine, 1.5 g/L sodium bicarbonate, 10 mM HEPES, 1 mM sodium pyruvate and 10% FBS. Certified FBS was from HyClone (Logan, UT), insulin was from Sigma (St. Louis, MO), and all other cell culture media components were from Mediatech (Herndon, VA).

Tumor specimens

Ovarian surgical specimens were obtained after informed consent under protocols approved by the Institutional Review Board for Human Experimentation of the University of Alabama at Birmingham. Normal ovary was collected at the time of hysterectomy and bilateral salpingo-oophorectomy. All tissues were collected under sterile conditions and immediately frozen at -70°C. The 7 ovarian tumor samples used in the present study included 6 designated as endometroid/papillary serous and 1 as papillary serous.

RNA isolation

Total RNA was isolated from frozen tissues and cultured cell lines by RNA STAT-60 reagent (TEL-Test, Friendswood, TX) according to the manufacturer's protocol. Contaminating DNA was removed with RQ RNase-free DNase I (Promega, Madison, WI) and RNA concentration measured spectrophotometrically (GeneQuant II RNA/DNA Calculator, Pharmacia Biotech Ltd., Cambridge, United Kingdom). RNA integrity and quality was analyzed by electrophoresis on a 1.2% denaturing agarose gel. For some experiments, nuclear and cytoplasmic RNA were isolated separately. Briefly, cultured cells were suspended in TKM buffer (10 mM Tris-HCl pH 7.4, 1 mM KCl, 1 mM

Table 1: RT-PCR primer pairs used for mesothelin transcript amplification

amplification target	sequence, 5' – 3'	location	expected product size
full length	ACCCACGGTGCCTCCCTCCC	exon 1	1906 bp (tr1)
(tr1 & 2)	TCAGGCCAGGGTGGAGGCTAG	exon 17	1924 bp (tr 2)
24 bp insert	CTACCTCTTCCTCAAGATGAG	exon 12	177 bp (tr1)
region (tr1 & 2)	GTGTCTAGGGTGTCTTTGTCT	exon 13	201 bp (tr2)
82 bp insert	ATGAAGCTGCGGACGGATGCG	exon 15	262 bp (tr 1)
region (tr1 & 3)	TGAGAACAGGTCCAGGTCCTA	exon 17	344 bp (tr 3)
tr3-specific	ACCCACGGTGCCTCCCTCCC	exon 1	1873 bp
	ACCTCCACGCCCCCAGCTCTG	intron 16	

$MgCl_2$). After 5 min on ice, 10% Triton-X 100 was added and the suspension mixed gently. Nuclear and cytoplasmic fractions were separated by centrifugation. The nuclear pellet was lysed by suspension in RNA STAT-60 and processed as above. Cytoplasmic RNA was isolated from the supernatant fraction by phenol/chloroform extraction and ethanol precipitation.

Oligonucleotides and RT-PCR

RT-PCR was performed using the Gene Amp PCR kit (Roche Molecular Systems, Branchburg, NJ) as described in the manufacturer's protocol, using 1 μg of total RNA as template with random hexamers for RT priming. Negative controls without the addition of RT enzyme were performed for each RNA samples. For PCR, oligonucleotide primers were added at a concentration of 0.3 μM each and amplification carried out for 35 cycles with an annealing temperature of 65°C. Negative controls without template were performed for each PCR amplification. RT-PCR products were analyzed by electrophoresis on a 1.5% NuSieve GTG (Cambrex Bio Science Rockland, Rockland, ME) plus 1% agarose gel, and products visualized by ethidium bromide staining. PCR primers (Table 1) were synthesized by Invitrogen Life Technologies (Carlsbad, CA).

DNA sequencing

Amplified PCR products were either directly sequenced or first cloned into the pCR4-TOPO vector (Invitrogen) and subjected to automated dideoxy nucleotide sequencing in the UAB Center for AIDS Research Sequencing Core Facility, using fluoresence-based cycle sequencing and the ABI BigDye Terminator V 3.1 Cycle Sequencing kit with the ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, California, USA). T3 and T7 primers were used for sequencing genes cloned into pCR4-TOPO: 5'-ATTAACCCTCACTAAAGGGA-3' and 5'-TAATACGACTCACTATAGGG-3'. Sequence data were analyzed by BLAST nucleotide-nucleotide search (National Center Biotechnology Information, http://www.ncbi.nlm.nih.gov/) against non-redundant and expressed sequence tag (EST) databases.

Results

In silico *analysis of mesothelin transcripts*

The NCBI database contains three reported transcripts for mesothelin/MPF: variant 1 (GenBank accession NM_005823) encoding MPF/mesothelin [8,9], variant 2 (NM_013404) encoding mesothelin [18], and variant 3 (AF180951) partial cDNA encoding mesothelin with an alternatively spliced C-terminus [19] (Figure 1). Transcript variant 1 was originally reported as MPF derived from a human pancreatic tumor cell line; the published sequence is 2085 bp with an opening reading frame of 1869 bp [8,9]. Transcript variant 2 was isolated from a HeLa cDNA library [18], and the reported sequence is 2107 bp long with an open reading frame of 1887 bp. The sequences of mesothelin variants 1 and 2 are 98% identical at the nucleotide level and 95% identical at the amino acid level. Differences include several amino acid variations near the N-terminus (amino acids 4–56) and a small insert in variant 2 compared to variant 1 after amino acid 410 of MPF (NP_005814). Full-length mesothelin/MPF protein is predicted to be encoded by 17 exons, occupying approximately 8 kb of human chromosome 16p13.3.

In silico comparison of the reported mesothelin transcripts with genomic sequences (contig clone NT_037887.3 and genomic clone AL031258) suggests that the basis for differences in the reported transcript 1 and 2 N-terminal region may have resulted from sequencing errors. Specifically, differences between amino acids 4–56 are most readily explained by incorrect sequences generated in GC rich areas (Figure 2A). The available genomic sequence matches variant 1, suggesting this variant is correct. Consistent with this, 25 of 25 EST sequences obtained by BLAST search matched variant 1 in this region, whereas no EST sequences were found to be identical to variant 2 in this region.

```
tr1   MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISS   60
tr2   ...QRLDPCW-...D-RP.............H.A........T.S...G...TT.H....   58

tr1   LSPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPL  120
tr2   ..........................................................  118

tr1   DLLLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEA  180
tr2   ..........................................................  178

tr1   DVRALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTW  240
tr2   ..........................................................  238

tr1   SVSTMDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKT  300
tr2   ..........................................................  298
tr3                                                          ....    4

tr1   ACPSGKKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELY  360
tr2   ..........................................................  358
tr3   ..........................................................   64

tr1   PQGYPESVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQ--------VA  412
tr2   .....................................D........APRRPLPQ..  418
tr3   .....................................N........--------..  116

tr1   TLIDRFVKGRGQLDKDTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQ  472
tr2   ..........................................................  478
tr3   ..........................................................  176

tr1   LDVLYPKARLAFQNMNGSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVL  532
tr2   ..........................................................  538
tr3   ..........................................................  236

tr1   PLTVAEVQKLLGPHVEGLKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLS  592
tr2   ..........................................................  598
tr3   ..........................................................  296

tr1   VQEALSGTPCLLGPGPVLTVLALLLASTLA  622
tr2   ...T.........................  628
tr3   ..GGRG.QARAG.RA.G.EVGALSHPSLCRGPLGDALPPRTWTCSHRPGTAPSLHPGLRA 356


tr3   PLPCWPQPCWGSPPGQEQARVIPVPPQENSRSVNGNMPPADT  398
```

Figure 1
Amino acid sequences of mesothelin transcript variants 1, 2 and 3. Amino acids in single-letter code are displayed for maximal sequence alignment of translated mesothelin variant 1 (tr1, GenBank accession NM_005823), variant 2 (tr2, NM_013404) and variant 3 (tr3, AF180951). (.) indicates amino acid identity, and (-) represents gaps for optimal sequence alignment. Note that tr3 is a partial cDNA lacking N-terminal sequence.

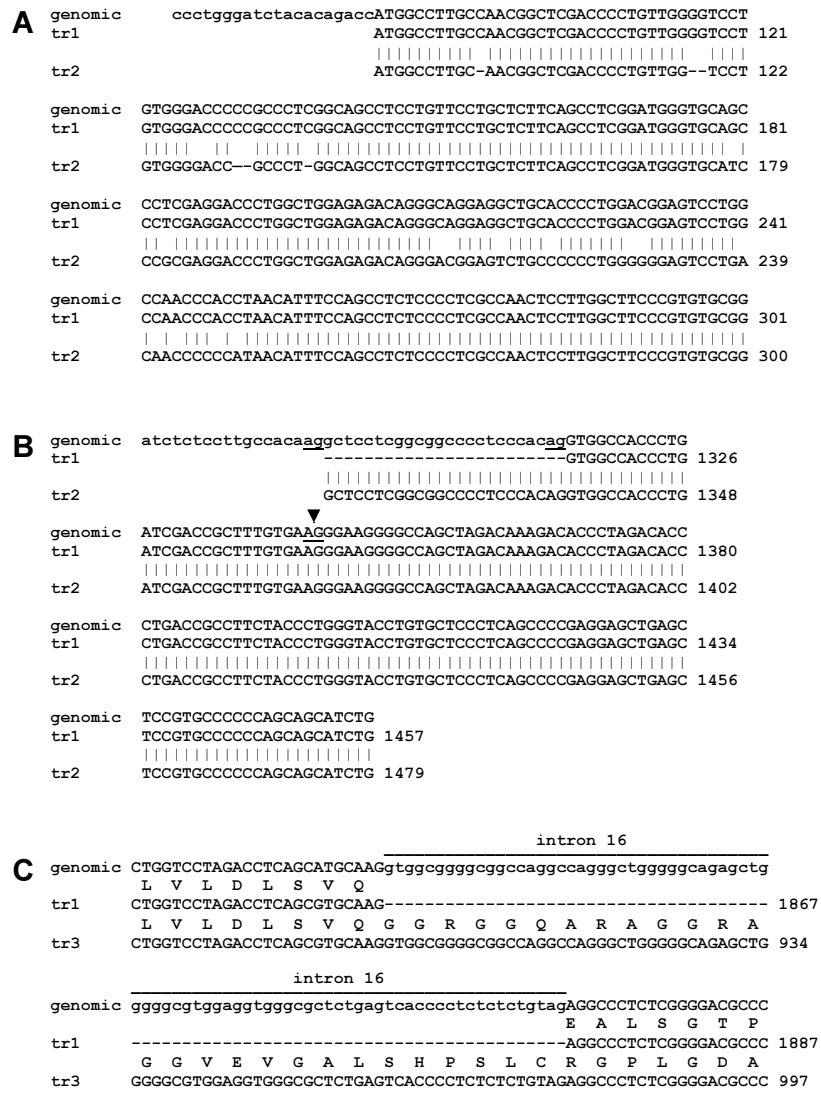

Figure 2

Human genomic DNA sequence alignment with mesothelin transcript variants. Selected portions of human genome sequences corresponding to the mesothelin gene (AL031258) were aligned with corresponding sequences of reported mesothelin transcript variants. (-) represents sequence gaps for optimal alignment. **A.** The 5'-ends of mesothelin transcript variants 1 and 2 (tr1, NM_005823; tr2, NM_013404) aligned with the genomic DNA sequence. Lower case letters represent presumed untranslated nucleotide sequences. **B.** Nucleotide sequence at the exon 13 5'-boundary using the same reference database sequences as in (A). Lower case letters in the genomic sequence indicate presumed introns for transcript variant 1. Splice acceptor sites potentially utilized by transcript variants 1 and 2 are underlined, and the arrowhead indicates a third splice acceptor detected by sequence analysis in this study. **C.** Genomic sequence for the intron 16 region are aligned with sequences for transcript variants 1 (tr1, NM_005823) and 3 (tr3, AF180951). Lower case letters represent nucleotides presumed to be untranslated for tr1. Expected amino acid sequences are shown above each transcript, highlighting the intron 16 insert and reading frame shift predicted for tr3.

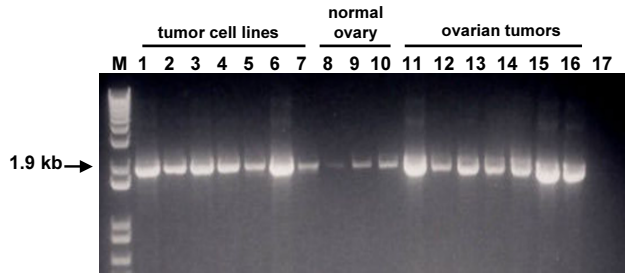

Figure 3
RT-PCR analysis of mesothelin transcripts. Primers designed to amplify full length mesothelin from transcript variants 1 and 2 (Table 1, line 1) were used in RT-PCR analyses of total RNA from cell lines and tissues, and analyzed by agarose gel electrophoresis. M, molecular size markers; Lane 1, HeLa; Lane 2, OVCAR-3; Lane 3, Caov-4; Lane 4, SKOV-3; Lane 5, UCI 101; Lane 6, AsPC-1; Lane7, PANC-1; Lanes 8–10, three individual normal ovary tissues; Lanes 11–16, six individual primary ovarian tumors; Lane 17, no template PCR control.

The other difference between variants 1 and 2 is an eight amino acid insertion following amino acid 410 (NP_005814). Analysis of the genomic sequence is consistent with the use of an alternative splice acceptor site in exon 13 (Figure 2B), resulting in an additional 24 bp insert in variant 2. Based on BLAST searches of expressed sequences, the 24 bp insert does not appear to be highly expressed in cancer cells, as only one perfect match was found in the dbEST database (CB266931); this contrasts with the adjacent sequence in exon 13 which was identical to more than 50 ESTs. Thus, analysis of existing sequence data is consistent with the existence of variant 1, and with the very rare use of an alternative splice acceptor site in exon 13 that corresponds to variant 2.

The third reported variant of mesothelin is a partial cDNA that includes an 82 bp insert near the C-terminus (AF180951). This sequence, which has not yet been reported as a full length transcript, is predicted to encode an alternative C-terminus with a hydrophilic tail that may produce a secreted form of mesothelin [19]. Genomic sequence data suggest that this 82 bp insert represents an unspliced intron 16 (Figure 2C). BLAST search of the dbEST database with the 82 bp insert sequence indicates that it is rarely expressed, because no ESTs were found that were identical to the entire 82 bp sequence.

RT-PCR analysis of mesothelin transcripts

Mesothelin/MPF transcripts expressed in human tissues and cell lines were analyzed by RT-PCR. Initial studies used PCR primers corresponding to the 5'-untranslated region and the 3' stop codon region of both transcript variants 1 and 2 (Table 1, full length), and a single PCR amplification product was detected in all studied cell lines and tissue samples (Figure 3). Complete nucleotide sequences were obtained for amplification products from the OVCAR-3 cell line, and partial nucleotide sequences obtained for the 5' regions (≥ 650 nucleotides) of products from HeLa, SKOV-3 and AsPC-1 cells, one normal ovary and two ovarian tumor samples. These demonstrated sequence identity to transcript variant 1 (NM_005823), with the exception of a single nucleotide change (A to G at nucleotide 222, conservative for amino acid) in HeLa, SKOV-3 and one ovarian tumor sample. In addition, for HeLa cells and one of the ovarian tumor samples, nucleotide sequences were also obtained on amplification products generated using a different set of mesothelin-specific primers, with identical results. None of the amplification products corresponded to transcript 2, consistent with the *in silico* analysis.

To further investigate possible use of the alternative splice acceptor site in exon 13 of mesothelin, as reported in variant 2, we designed primers in exons 12 and 13 (Table 1, 24 bp insert) and amplified RNA from various cell lines and tissues. Amplification of a transcript that utilizes the alternative splice acceptor site (variant 2) should result in a PCR product of 201 bp. However, only a 177 bp amplification product consistent with the size expected from variant 1 was detected (Figure 4B). Nucleotide sequence analysis of amplification products confirmed use of the exon 13 splice acceptor site corresponding to transcript variant 1. We conclude that utilization of the alternative splice donor site, as reported for variant 2, is rare in both normal ovary and in pancreatic and ovarian cancer. Interestingly, we also detected a minor amplification product that was smaller than expected (147 bp, Figure 4B). Cloning and sequence analysis demonstrated that this transcript, estimated to comprise less than 5% of the total mesothelin message, utilized an internal splice acceptor site in exon 13 (indicated by the arrowhead in Figure 2B), predicted to encode a protein with an in-frame deletion of 10 amino acids (VATLIDRFVK).

We next investigated expression of mesothelin transcript variant 3 (AF180951), predicted to encode a form of mesothelin with a disrupted GPI-anchor region (exons 16 and 17). PCR primers corresponding to sequences in exons 15 and 17 (Table 1, 82 bp insert) should generate an amplification product of 262 bp if intron 16 is spliced from the final transcript, whereas a 344 bp fragment would be expected if intron 16 is retained (Figure 5A). Amplification of cDNA from cell lines and tumors demonstrated a predominant product of 262 bp (Figure 5B). Additional minor products of 344 bp and of a size

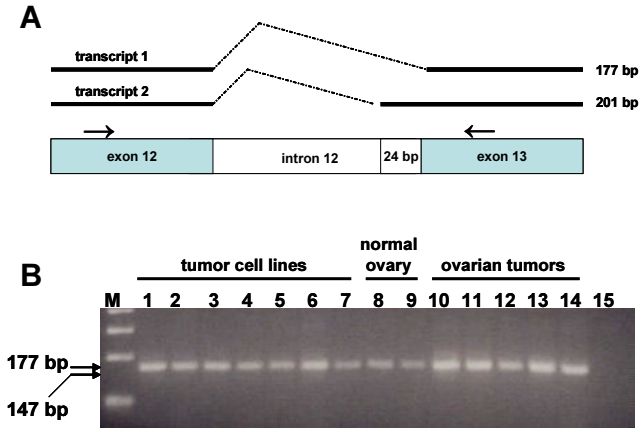

Figure 4
RT-PCR analysis of potential mesothelin exon 13 splicing differences between transcript variants 1 and 2. **A.** Diagram of the PCR strategy, showing predicted size of amplified products from transcript variants 1 and 2. PCR primers (Table 1, line 2) are indicated by arrows. **B.** Agarose gel analysis of representative RT-PCR results. Lane M, molecular size marker; Lane 1, HeLa; Lane 2, OVCAR-3; Lane 3, Caov-4; Lane 4, SKOV-3; Lane 5, UCI 101; Lane 6, AsPC-1; Lane7, PANC-1; Lanes 8–9, two individual normal human ovaries; Lanes 10–14, five individual primary ovarian tumors; Lane 15, no template PCR control.

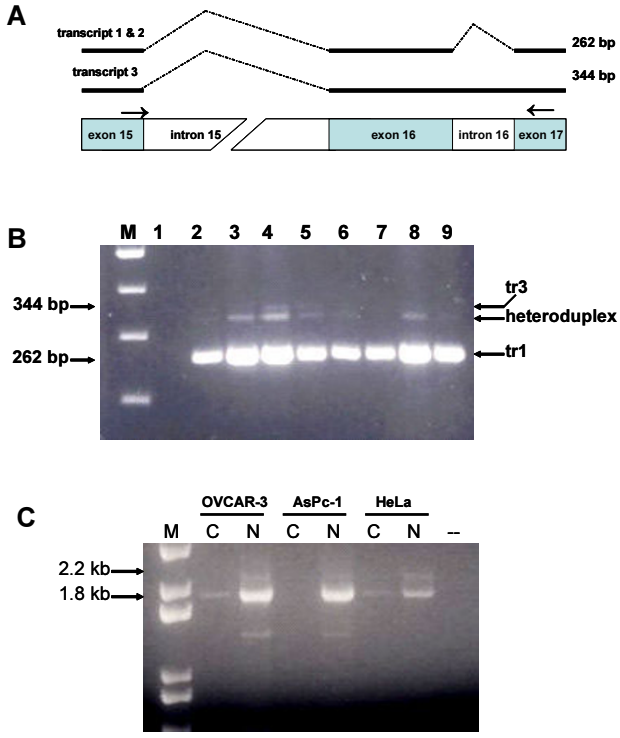

Figure 5
RT-PCR analysis of potential mesothelin exon 16–17 splicing differences between transcript variants 1 and 3. **A.** Diagram of the PCR strategy, showing predicted size of amplified products from transcript variants 1 and 3. PCR primers (Table 1, line 3) are indicated by arrows. **B.** Agarose gel analysis of representative RT-PCR results. Lane M, molecular size markers; Lane 1, no template PCR control; Lane 2, PANC-1; Lane 3, AsPc-1; Lane 4, HeLa; Lane 5, OVCAR-3; Lane 5, normal ovary tissue sample; Lane 6–8, three individual primary ovarian tumors. The product with a size between 262 and 344 bp represents heteroduplex formation between transcript variants 1 and 3 (tr1, tr3). **C.** RT-PCR and agarose gel analysis of full length mesothelin (transcript 1 versus 3) in cytoplasmic (C) versus nuclear (N) RNA fractions as a template, using primers in Table 1, line 4. Lane M, molecular size marker; Lane 1, OVCAR-3 cytoplasmic fraction; Lane 2, OVCAR-3 nuclear fraction; Lane 3, AsPC-1 cytoplasmic fraction; Lane 4, AsPC-1 nuclear fraction; Lane 5, HeLa cytoplasmic fraction; Lane 6, HeLa nuclear fraction; (--), no template PCR control.

between 262 and 344 bp were visualized in all samples. Sequence analysis demonstrated that the amplified products corresponded to transcript variant 1 (262 bp) and variant 3 (344 bp), with the middle band representing heteroduplexes between the two. Relative intensity of the amplified bands suggested that transcript variant 3 represents less than 10% of total mesothelin transcripts in the studied samples.

Because only a partial cDNA sequence for mesothelin variant transcript 3 has been submitted to the public database, RT-PCR analysis was performed using a 3' primer specific for the 82 bp insert of transcript 3 (intron 16) and a 5' primer located immediately before the start codon of transcript 1 (Table 1, tr3-specific). This primer pair should allow amplification of full length transcripts only from variant 3. Results using RNA from Caov-4, OVCAR-3, PANC-1, AsPC-1, HeLa, and 5 ovarian tumor samples yielded several amplification products including one of the size expected for full length mesothelin variant 3 (1873 bp, data not shown). Cloning and sequence analysis of individual amplification products demonstrated that in addition to the predominant product corresponding to variant 3 with an unspliced intron 16, additional products included unspliced introns 8 and 9, or partial sequences of other introns. These results raised the possibility that such transcripts might represent hnRNA which has not been completely processed.

To test this possibility, cultured cell lines were fractionated to obtain cytoplasmic RNA (representing fully processed mature transcripts) and nuclear RNA (which should include both mature and incompletely processed transcripts). RT-PCR amplification of using the above primers for mesothelin transcripts containing intron 16 (variant 3) with three cell lines indicated that the intron 16 containing transcripts were primarily amplified from the nuclear pool of RNA, and only small amounts of product were recovered from the cytoplasmic RNA (Figure 5C). Thus, although transcript variant 3 appears to be present in low levels in both normal and tumor cells, it seems to be very rare among the cytoplamic fraction of mature, completely processed RNA.

Discussion

Mesothelin is a promising target for cancer diagnostics and therapy [12-14,19]. In normal cells, expression is primarily restricted to mesothelial cells in the pleural, pericardial, and peritoneal membranes. Limited mesothelin immunoreactivity has also been reported in trachea, tonsil, fallopian tube and kidney [1]. In contrast, mesothelin is abundantly expressed in certain tumors, including up to 100% of serous papillary ovarian tumors [5] and up to 100% of pancreatic tumors [4]. Frequent and high level expression by tumors suggests that mesothelin might play a role in tumorigenesis. Immunohistochemical studies demonstrate that mesothelin expression is directly correlated with progressive development of metastatic pancreatic cancer [20]. Recent evidence that mesothelin may promote cell-cell adhesion in ovarian cancer through heterotypic binding to CA125/MUC16 further supports a role in tumor progression [11]. Circulating mesothelin may serve as a serum marker for early detection and monitoring of mesothelin-positive tumors [12]. However, in order to fully exploit mesothelin as a target for clinical applications, it is important to understand mesothelin expression with regard to the transcript variants that have been reported to date.

Various investigators have referred to this gene and its products as mesothelin [18], MPF [8,9], soluble mesothelin [19], mesothelin family members or soluble mesothelin related proteins (SMR) [12]. The *in silico* and *in vitro* studies reported here support mesothelin variant 1 as the major transcript expressed by both normal and malignant cells. Mesothelin variant 2 (NM_013404) was not detected. While it remains possible that nucleotide differences in variant 2 are the result of allelic variation, alignment of the transcript variant 2 sequence with available genomic sequences suggests that the differences in amino acids 4–56 between variants 1 and 2 may be due to sequencing errors. Another feature of variant 2, a 24 bp insertion after base pair 1313 (transcript 2 NM_013404), appears to arise from use of an alternative splice acceptor site upstream of the usual splice acceptor site, and was very rare in our studies. The third reported mesothelin transcript variant, encoding an alternative C-terminus, appears to be largely restricted to the nuclear fraction of RNA and to be infrequent in the cytoplasmic pool of mature mRNA, suggesting that this variant may represent incompletely processed hnRNA. Although definitive studies to detect the soluble isoform of mesothelin protein predicted to be encoded by variant 3 will be needed to rule out the possibility of low level expression, data presented here suggests that variant 1 represents the predominant form of mesothelin expressed by both normal and tumor cells.

Despite the presence of a single major transcript for mesothelin in the samples evaluated here, immunohistochemical studies of mesothelin protein have reported distinct expression patterns in tumor samples. In human pancreatic tumors, some samples exhibited focal staining, and mesothelin was commonly detected in single tumor cells surrounded on all sides by stroma; malignant glands were generally less well stained, but expression was accentuated at luminal borders and luminal contents were frequently positive, suggesting secretion or shedding of mesothelin from tumor cells [4]. Other analyses of a variety of tumors, including ovarian cancer, report mesothelin staining at the cell membrane, in the cytoplasm, or both, depending on the sample; staining was sometimes accentuated at apical cell surfaces and in intracytoplasmic lumens, with extracellular luminal contents again noted to be frequently labeled [5,6,19]. Our own immunohistochemical studies of ovarian tumors (unpublished data) are consistent with the available literature and are notable in that the pattern of localization of mesothelin is consistent within a tumor specimen, with some showing apical localization and apparent secretion of mesothelin, while cytoplasmic staining is prominent in others.

Given that mesothelin transcript variant 1 is widely and highly expressed, the basis for differential localization of mesothelin protein in different tumor samples is most likely a consequence of post-translational processing. Mesothelin variant 1 encodes a GPI anchor motif, and proteins attached to cell plasma membranes by GPI anchors can be readily released *in vivo*. A relevant example is carcinoembryonic antigen (CEA), a tumor-associated GPI-anchored glycoprotein that is commonly shed by CEA-positive malignant cells [21]. A furin-like protease cleavage site within mesothelin that is probably responsible for the release of the N-terminal MPF portion [9] could conceivably contribute to mesothelin-immunoreactive extracellular contents observed in tumor tissue samples and in serum from tumor patients, but evidence to date suggests that mesothelin-specific antibodies used for both types of studies react with epitopes in the mem-

brane-proximal C-terminal portion of mesothelin. Additional studies delineating the post-translational processing and secretion of mesothelin are warranted.

The basis of the reported cross-reactivity of mesothelin antibodies with other proteins in serum or ascites [19] is not clear, but is not likely to be due to cross-reactivity with endogenous homologous proteins. Current databases contain only one vertebrate gene with significant homology to mesothelin, otoancorin (NM_170664), which has been implicated in autosomal recessive deafness and has 21–22% amino acid identity to mesothelin in the C-terminal region [22]. Interestingly, otoancorin is also located on human chromosome 16p, but its specific function is unknown and its expression is apparently limited to the inner ear.

Conclusions

Mesothelin appears to be predominantly encoded by a single transcript corresponding the previously reported mesothelin/MPF variant 1 [8,9]. Although alternative splicing of the transcript may occur at low levels, the vast majority of message in primary ovarian tumors, ovarian and pancreatic tumor cell lines, and normal ovary appears to be represented by variant 1. Targeting mesothelin transcript variant 1 products for future clinical applications is predicted to be the most valid approach for developing novel therapies and diagnostic assays for mesothelin-expressing malignancies.

List of abbreviations

BLAST, basic local alignment search tool; CEA, carcinoembryonic antigen; dbEST, database of EST; EST, expressed sequence tag; FBS, fetal bovine serum; GPI, glycosylphosphatidylinositol; hnRNA, heterogenous nuclear RNA; MPF, megakaryocyte potentiating factor; NCBI, National Center for Biotechnology Information; PCR, polymerase chain reaction; RT, reverse transcriptase; tr, transcript; UAB, University of Alabama at Birmingham.

Competing interests

None declared.

Authors' contributions

ZEM performed the experiments, interpreted the data, and contributed to writing the manuscript. TVS and DRS conceived the study, designed the experiments, supervised the work, interpreted the data, and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank J. Michael Straughn for assisting in acquisition of ovarian tissue samples, Saengsoon Trirekapan for preparation of ovarian tissue RNA, and David T. Curiel for providing the UCI 101 ovarian carcinoma cell line. This research was supported by Developmental Research Projects awarded to DRS and TVS by the University of Alabama at Birmingham SPORE in Ovarian Cancer (NIH NCI 1 P50 CA83591).

References

1. Chang K, Pastan I, Willingham MC: **Isolation and characterization of a monoclonal antibody, K1, reactive with ovarian cancers and normal mesothelium.** *Int J Cancer* 1992, **50:**373-381.
2. Chang K, Pai LH, Pass H, Pogrebniak HW, Tsao MS, Pastan I, Willingham MC: **Monoclonal antibody K1 reacts with epithelial mesothelioma but not with lung adenocarcinoma.** *Am J Surg Pathol* 1992, **16:**259-268.
3. Chang K, Pastan I: **Molecular cloning and expression of a cDNA encoding a protein detected by the K1 antibody from an ovarian carcinoma (OVCAR-3) cell line.** *Int J Cancer* 1994, **57:**90-97.
4. Argani P, Iacobuzio-Donahue C, Ryu B, Rosty C, Goggins M, Wilentz RE, Murugesan SR, Leach SD, Jaffee E, Yeo CJ, Cameron JL, Kern SE, Hruban RH: **Mesothelin is overexpressed in the vast majority of ductal adenocarcinomas of the pancreas: identification of a new pancreatic cancer marker by serial analysis of gene expression (SAGE).** *Clin Cancer Res* 2001, **7:**3862-3868.
5. Frierson H. F., Jr., Moskaluk CA, Powell SM, Zhang H, Cerilli LA, Stoler MH, Cathro H, Hampton GM: **Large-scale molecular and tissue microarray analysis of mesothelin expression in common human carcinomas.** *Hum Pathol* 2003, **34:**605-609.
6. Ordonez NG: **Value of mesothelin immunostaining in the diagnosis of mesothelioma.** *Mod Pathol* 2003, **16:**192-197.
7. Ordonez NG: **Application of mesothelin immunostaining in tumor diagnosis.** *Am J Surg Pathol* 2003, **27:**1418-1428.
8. Yamaguchi N, Hattori K, Oh-eda M, Kojima T, Imai N, Ochi N: **A novel cytokine exhibiting megakaryocyte potentiating activity from a human pancreatic tumor cell line HPC-Y5.** *J Biol Chem* 1994, **269:**805-808.
9. Kojima T, Oh-eda M, Hattori K, Taniguchi Y, Tamura M, Ochi N, Yamaguchi N: **Molecular cloning and expression of megakaryocyte potentiating factor cDNA.** *J Biol Chem* 1995, **270:**21984-21990.
10. Bera TK, Pastan I: **Mesothelin is not required for normal mouse development or reproduction.** *Mol Cell Biol* 2000, **20:**2902-2906.
11. Rump A, Morikawa Y, Tanaka M, Minami S, Umesaki N, Takeuchi M, Miyajima A: **Binding of ovarian cancer antigen CA125/MUC16 to mesothelin mediates cell adhesion.** *J Biol Chem* 2004, **279:**9190-8.
12. Robinson BW, Creaney J, Lake R, Nowak A, Musk AW, de Klerk N, Winzell P, Hellstrom KE, Hellstrom I: **Mesothelin-family proteins and diagnosis of mesothelioma.** *Lancet* 2003, **362:**1612-1616.
13. Hassan R, Viner JL, Wang QC, Margulies I, Kreitman RJ, Pastan I: **Anti-tumor activity of K1-LysPE38QQR, an immunotoxin targeting mesothelin, a cell-surface antigen overexpressed in ovarian cancer and malignant mesothelioma.** *J Immunother* 2000, **23:**473-479.
14. Hassan R, Lerner MR, Benbrook D, Lightfoot SA, Brackett DJ, Wang QC, Pastan I: **Antitumor activity of SS(dsFv)PE38 and SS1(dsFv)PE38, recombinant antimesothelin immunotoxins against human gynecologic cancers grown in organotypic culture in vitro.** *Clin Cancer Res* 2002, **8:**3520-3526.
15. Schaner ME, Ross DT, Ciaravino G, Sorlie T, Troyanskaya O, Diehn M, Wang YC, Duran GE, Sikic TL, Caldeira S, Skomedal H, Tu IP, Hernandez-Boussard T, Johnson SW, O'Dwyer PJ, Fero MJ, Kristensen GB, Borresen-Dale AL, Hastie T, Tibshirani R, van de Rijn M, Teng NN, Longacre TA, Botstein D, Brown PO, Sikic BI: **Gene expression patterns in ovarian carcinomas.** *Mol Biol Cell* 2003, **14:**4376-4386.
16. Iacobuzio-Donahue CA, Ashfaq R, Maitra A, Adsay NV, Shen-Ong GL, Berg K, Hollingsworth MA, Cameron JL, Yeo CJ, Kern SE, Goggins M, Hruban RH: **Highly expressed genes in pancreatic ductal adenocarcinomas: a comprehensive characterization and comparison of the transcription profiles obtained from three major technologies.** *Cancer Res* 2003, **63:**8614-8622.
17. Wang K, Gan L, Jeffery E, Gayle M, Gown AM, Skelly M, Nelson PS, Ng WV, Schummer M, Hood L, Mulligan J: **Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray.** *Gene* 1999, **229:**101-108.

18. Chang K, Pastan I: **Molecular cloning of mesothelin, a differentiation antigen present on mesothelium, mesotheliomas, and ovarian cancers.** *Proc Natl Acad Sci U S A* 1996, **93:**136-140.
19. Scholler N, Fu N, Yang Y, Ye Z, Goodman GE, Hellstrom KE, Hellstrom I: **Soluble member(s) of the mesothelin/megakaryocyte potentiating factor family are detectable in sera from patients with ovarian carcinoma.** *Proc Natl Acad Sci U S A* 1999, **96:**11531-11536.
20. Maitra A, Adsay NV, Argani P, Iacobuzio-Donahue C, De Marzo A, Cameron JL, Yeo CJ, Hruban RH: **Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray.** *Mod Pathol* 2003, **16:**902-912.
21. Horig H, Medina FA, Conkright WA, Kaufman HL: **Strategies for cancer therapy using carcinoembryonic antigen vaccines.** *Expert Rev Mol Med* 2000, **2000:**1-24.
22. Zwaenepoel I, Mustapha M, Leibovici M, Verpy E, Goodyear R, Liu XZ, Nouaille S, Nance WE, Kanaan M, Avraham KB, Tekaia F, Loiselet J, Lathrop M, Richardson G, Petit C: **Otoancorin, an inner ear protein restricted to the interface between the apical surface of sensory epithelia and their overlying acellular gels, is defective in autosomal recessive deafness DFNB22.** *Proc Natl Acad Sci U S A* 2002, **99:**6240-6245.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2407/4/19/prepub

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral