

(21)(A1) 2,329,527

(86) 1999/05/28

(87) 1999/12/09

(72) STARZINSKI-POWITZ, ANNA, DE

(72) KOTZIAN, SILVIA, DE

(72) HANDROW-METZMACHER, HEIKE, DE

(71) STARZINSKI-POWITZ, ANNA, DE

(51) Int.Cl.<sup>6</sup> C12N 15/12, A61K 31/70, C12Q 1/68, G01N 33/68,  
C12N 15/63, A61K 38/17, C12N 5/10, G01N 33/574, C07K 14/47,  
A61K 39/395, C07K 16/18

(30) 1998/05/29 (198 24 230.1) DE

(54) GENE ASSOCIE A L'ENDOMETRIOSE

(54) ENDOMETRIOSIS-ASSOCIATED GENE

(57) La présente invention concerne un gène associé à des processus invasifs, par exemple à l'endométriose, un polypeptide codé par ce gène, un anticorps dirigé contre ce polypeptide, ainsi que l'utilisation pharmaceutique de l'acide nucléique correspondant, dudit polypeptide et dudit anticorps.

(57) The invention relates to a gene associated with invasive processes, e.g. endometriosis, to a polypeptide coded by said gene, to an antibody directed against the polypeptide, and to the pharmaceutical application of the nucleic acid, the polypeptide and the antibody.



PCT

WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales BüroINTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation <sup>6</sup> :		A1	(11) Internationale Veröffentlichungsnummer: <b>WO 99/63079</b>
<b>C12N 15/12, C07K 14/47, 16/18, C12N 5/10, 15/63, G01N 33/68, 33/574, C12Q 1/68, A61K 38/17, 39/395, 31/70</b>			(43) Internationales Veröffentlichungsdatum: 9. Dezember 1999 (09.12.99)
(21) Internationales Aktenzeichen: PCT/EP99/03716		(81) Bestimmungsstaaten: AU, CA, CN, IL, JP, NZ, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) Internationales Anmeldedatum: 28. Mai 1999 (28.05.99)			
(30) Prioritätsdaten: 198 24 230.1	29. Mai 1998 (29.05.98)	DE	Veröffentlicht <i>Mit internationalem Recherchenbericht. Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist; Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i>
(71)(72) Anmelder und Erfinder: STARZINSKI-POWITZ, Anna [DE/DE]; Zeisselstrasse 9, D-60318 Frankfurt am Main (DE).			
(72) Erfinder; und			
(75) Erfinder/Anmelder (nur für US): KOTZIAN, Silvia [DE/DE]; Im Höhlchen 1, D-65795 Hattersheim am Main (DE).			
HANDROW-METZMACHER, Heike [DE/DE]; Alt Eschersheim 45, D-60433 Frankfurt am Main (DE).			
(74) Anwälte: WEICKMANN, H. usw.; Kopernikusstrasse 9, D-81679 München (DE).			

(54) Title: ENDOMETRIOSIS-ASSOCIATED GENE

(54) Bezeichnung: ENDOMETRIOSE-ASSOZIIERTES GEN

(57) Abstract

The invention relates to a gene associated with invasive processes, e.g. endometriosis, to a polypeptide coded by said gene, to an antibody directed against the polypeptide, and to the pharmaceutical application of the nucleic acid, the polypeptide and the antibody.

(57) Zusammenfassung

Die vorliegende Erfindung betrifft ein mit invasiven Prozessen, z.B. Endometriose assoziiertes Gen, ein davon kodiertes Polypeptid, einen gegen das Polypeptid gerichteten Antikörper sowie die pharmazeutische Anwendung der Nukleinsäure, des Polypeptids und des Antikörpers.

**Novel endometriosis-associated gene****Description**

5 The present invention relates to a gene associated with invasive processes, for example endometriosis, to a polypeptide encoded by it, to an antibody directed against the polypeptide, and to the pharmaceutical application of the nucleic acid, the polypeptide and  
10 the antibody.

Endometriosis is the second most common disease in women and is defined as the occurrence of endometrial cells outside the womb. Endometriosis affects about one  
15 in five women of reproductive age, and as many as one in two women with fertility problems.

In normal circumstances the endometrium is only found in the womb. In endometriosis, tissue with a  
20 histological appearance resembling the endometrium is found outside the womb, for example externally on the womb, on the intestine or even in the pancreas or the lung. Although these endometriotic foci are located outside the womb, they also bleed during menstruation,  
25 thus they are influenced by hormones of the female cycle. Since endometriotic foci like the endometrium go through volume changes during the cycle, these changes may cause pain depending on location. Moreover, the body reacts to endometriotic cells with an inflammatory  
30 response which again causes pain. Furthermore, inflammation leads to adhesions in the area of the ovaries and fallopian tubes and, as a result of these, is responsible for a so-called mechanical sterility of affected women. Apparently however, in endometriosis  
35 messengers are released as well (e.g. cytokines, prostaglandins) which can reduce the fertility of affected women even in the absence of adhesions.

In view of their pathobiological properties,

- 2 -

endometriotic cells could be classified as being between normal cells and tumor cells: on the one hand they show no neoplastic behavior, on the other hand, however, they are, like metastasizing tumor cells, 5 capable of moving across organ boundaries in the organism and of growing into other organs, i.e. they show invasive behavior. For this reason endometriotic cells are defined as "benign tumor cells" in the literature, although up until now no tumor-specific 10 mutations in proto-oncogenes have been found in cells of this type.

Since the pathogenesis of endometriosis is still not clarified completely, there are as yet no effective 15 options for the therapy or prevention of endometriosis-associated diseases.

It was the object of the invention to identify novel genes which play a role in invasive processes and which 20 may be associated with the pathophysiological phenotype of endometriosis.

This object is achieved according to the invention by identifying, cloning and characterizing a gene which is 25 called an endometriosis-associated gene and which codes for a polypeptide. This gene sequence was discovered with the aid of differential display RT-PCR (Liang and Pardee, Science 257 (1992), 967-971). For this, invasive and noninvasive variants of an endometriotic 30 cell line were compared with each other. In the process a cDNA sequence was found which is specific for the invasive variant of endometriotic cells. An associated RNA of 4 kb in length was found. A corresponding cDNA isolated from a cDNA phage bank has an open reading 35 frame (ORF) of 302 amino acids.

The present invention relates to a nucleic acid which comprises

(a) the nucleotide sequences depicted in SEQ ID NO. 1,

- 3 -

3 or/and 5, a combination or a protein-encoding segment thereof,

(b) a nucleotide sequence corresponding to the sequence in (a) within the scope of the degeneracy of the genetic code or

5 (c) a nucleotide sequence hybridizing with the sequences in (a) and/or (b) under stringent conditions.

10 The nucleic acids preferably code for a polypeptide associated with invasive processes or a segment thereof.

The following nucleotide sequences have been deposited in the EMBL EST database with the following accession numbers: Z98886, AC003017, AL023586, Aa52993, Aa452856. These sequences do not represent nucleic acids according to the invention. The first two of these sequences are DNAs which were isolated from human brain and show over 90% identical bases to SEQ. ID NO. 1 in the segments from nucleotide 970 to about 2000 and from 760 to about 1450, respectively, or in the segments from nucleotide 1054 to 2084 and from 844 to about 1534 in relation to SEQ ID NO. 3 which has 84 additional bases at the 5' end. AL023586 is also a human sequence which is very similar to Z98885 and also has homology with SEQ ID NO. 1 in the region from 970 to about 2000.

Sequences Aa452993 and Aa452856 originate from mouse embryos and show base identity with the nucleotides (nt) from about 1060 to about 1450 and from about 24 to 440, respectively, of SEQ. ID NO. 1, or from about 1144 to about 1534 and from about 108 to about 524, respectively, according to the nucleotide positions in SEQ. ID NO. 3. Up until now no reading frame or function has been assigned to any of these 4 sequences.

The nucleotide sequence depicted in SEQ. ID NO. 1 contains an open reading frame which corresponds to a polypeptide having a length of 302 amino acids. This

- 4 -

polypeptide is indicated in the amino acid sequence depicted SEQ. ID NO. 2. SEQ. ID NO. 3 shows a nucleotide sequence as in SEQ. ID NO. 1, but it has 84 additional nucleotides at the 5' end. As a result, the 5 positions of the nucleotides corresponding to each other shift by 84 nucleotides in each case. The polypeptide encoded by SEQ. ID NO. 3 therefore has 28 additional amino acids at the N terminus and is depicted in SEQ. ID NO. 4 with its total of 330 amino 10 acids. SEQ. ID NO. 2 and 4 depict a C-terminal segment of the native polypeptide.

For illustration purposes reference is made to figure 1 which shows a diagrammatic representation of the cDNA 15 of the endometriosis-associated gene according to the invention. Five exons, E1 to E5, and the position of fragment 1 (394 nt) used as a probe in DDRT-PCR are shown. The positions of the PCR primers (see example 4, table 1) used for RT-PCR are also shown.

20

Not shown in figure 1 is a further exon 4a whose nucleotide sequence is shown in SEQ. ID NO. 5. This exon 4a may be present. If it is present, it is found between exon 4 and exon 5. This corresponds to the 25 position between nt1054 and nt1055 in SEQ. ID NO. 3. A combination of the sequences SEQ. ID NO. 1/3 with SEQ. ID NO. 5 is accordingly, for example, a sequence which contains the sequence of the exon 4a at said position.

30 Besides the nucleotide sequences shown in SEQ. ID NO. 1, 3 and 5 and combinations thereof such as the sequence of SEQ. ID NO. 3, which has the sequence of SEQ. ID NO. 5 between nt1054 and 1055 and to a nucleotide sequences which corresponds to the sequences 35 within the scope of the degeneracy of the genetic code, the present invention also includes nucleotide sequences which hybridize with one of the sequences mentioned before. The term "hybridization

- 5 -

according to the present invention is used by Sambrook et al. (Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989), 1.101-1.104). Preferably a hybridization is called stringent if a 5 positive hybridization signal is still observed after washing for one hour with 1 X SSC and 0.1% SDS at 50°C, preferably at 55°C, particularly preferably at 62°C and most preferably at 68°C, in particular for 1 h in 0.2 X SSC and 0.1% SDS at 55°C, preferably at 55°C, 10 particularly preferably at 62°C and most preferably at 68°C. A nucleotide sequence hybridizing under these washing conditions with one or more of the nucleotide sequences depicted in SEQ ID NO. 1, 3 and 5, or with a nucleic sequence corresponding to these sequences 15 within the scope of the degeneracy of the genetic code, is a nucleotide sequence according to the invention.

The nucleotide sequence according to the invention is preferably a DNA. However, it can also include an RNA or a nucleic acid analog such as a 20 peptidic nucleic acid, for example. Particularly preferably the nucleic acid according to the invention includes a protein-encoding segment of the nucleotide sequences depicted in SEQ ID NO. 1, 3 and/or 5 or a sequence having a homology of more than 80%, preferably 25 more than 90% and particularly preferably more than 95% to the nucleotide sequences depicted in SEQ ID NO. 1, 3 or 5 or a segment of preferably at least 20 nucleotides (nt) and particularly preferably at least 50 nt thereof. The same also holds for nucleic acids which 30 have, as described above, the sequence of SEQ. ID NO. 5 in addition to those of SEQ ID NO. 1 or 3. The homology is given in percent identical positions when two nucleic acids (or peptide chains) are compared, where a 35 100% homology means complete identity of the compared chain molecules (Herder: Lexikon der Biochemie und Molekularbiologie [Dictionary of biochemistry and molecular biology], Spektrum Akademischer Verlag 1995).

Nucleic acids according to the invention are preferably

obtainable from mammals and in particular from humans. They may be isolated according to known techniques by using short segments of the nucleotide sequences shown in SEQ. ID NO. 1, 3 or/and 5 as hybridization probes 5 and/or as amplification primers. Furthermore, the nucleic acids according to the invention may also be prepared by chemical synthesis, it being possible to employ modified nucleotide building blocks, for example 2'-O-alkylated nucleotide building blocks, where 10 appropriate, instead of conventional nucleotide building blocks.

The nucleic acids according to the invention or segments thereof may therefore be used for preparing 15 primers and probes which preferably contain markers or labeling groups. Preference is also given to intron-bridging oligonucleotide primers which are particularly suitable for identifying different mRNA species.

20 The present invention further relates to polypeptides encoded by the nucleic acids defined as above. These polypeptides preferably comprise  
(a) the amino acid sequence depicted in SEQ ID NO. 2 or 4 or  
25 (b) a homology of more than 70%, preferably of more than 80% and particularly preferably of more than 90% to the amino acid sequence according to (a).

Besides the polypeptides depicted in SEQ ID NO. 2 or 4, 30 the invention also relates to muteins, variants and fragments thereof. These are sequences which differ from the amino acid sequences depicted in SEQ ID NO. 2 or 4 by substitution, deletion and/or insertion of single amino acids or of short amino acid segments.

35 The term "variant" includes both naturally occurring allelic variations or splicing variations of the endometriotic protein, and proteins generated by recombinant DNA technology (in particular *in vitro*

- 7 -

mutagenesis with the aid of chemically synthesised oligonucleotides) which correspond substantially to the proteins depicted in SEQ ID NO. 2 or 4 with respect to their biological and/or immunological activity. This 5 term also includes chemically modified polypeptides. Polypeptides which are modified at the termini and/or in the reactive amino acid side groups by acylation, for example acetylation or amidation belong to this group. Polypeptide fragments (peptides) representing a 10 segment of at least 10 amino acids of the amino acid sequence shown in SEQ ID NO. 2 or 4 also belong to the amino acid sequences according to the invention.

The present invention further relates to a vector 15 containing at least one copy of a nucleic acid according to the invention. This vector may be any prokaryotic or eukaryotic vector on which the DNA sequence according to the invention, preferably linked to expression signals such as promoter, operator, 20 enhancer etc., is located. Examples of prokaryotic vectors are chromosomal vectors such as bacteriophages and extrachromosomal vectors such as plasmids, with circular plasmid vectors being particularly preferred. Suitable prokaryotic vectors are described, for 25 example, in Sambrook et al., *supra*, Chapters 1-4. Particularly preferred is the vector according to the invention, a eukaryotic vector, e.g. a yeast vector, or a vector suitable for higher cells, e.g. plasmid vector, viral vector or plant vector. Vectors of this 30 type are well known to the skilled worker in the field of molecular biology so that there is no need for further explanation here. In particular, reference is made in this connection to Sambrook et al., *supra*, Chapter 16.

35

The invention also relates to a vector which contains a segment of at least 21 nucleotides in length of the sequences depicted in SEQ ID NO. 1, 3 or/and 5 or a combination thereof. Preferably this segment has a

nucleotide sequence which originates from the protein-encoding region of said sequences or from a region essential for the expression of the protein or polypeptide. These nucleic acids are particularly 5 suitable for preparing therapeutically employable antisense nucleic acids preferably of up to 50 nucleotides in length.

The present invention further relates to a cell 10 transformed with a nucleic acid according to the invention or a vector according to the invention. The cell can be both a eukaryotic and a prokaryotic cell. Methods for transforming cells with nucleic acids are general prior art and therefore need no further 15 explanation. Examples of preferred cells are eukaryotic cells, in particular animal and particularly preferably mammalian cells.

The present invention further relates to an antibody or 20 a fragment of such an antibody against the polypeptide(s) encoded by the endometriosis gene or against variants thereof. Antibodies of this type are particularly preferably directed against complete polypeptides encoded by it or against a peptide 25 sequence corresponding to amino acids 1-330 of the amino acid sequence depicted in SEQ ID NO. 4.

Identification, isolation and expression of a gene according to the invention which is specifically 30 associated with invasive processes and in particular with endometriosis provide the requirements for diagnosis, therapy and prevention of diseases based on those disorders mentioned above.

35 It becomes possible with the aid of a polypeptide according to the invention or fragments of this polypeptide as immunogen to prepare antibodies against those polypeptides. Preparation of antibodies may be carried out in the usual way by immunizing experimental

animals with the complete polypeptide or fragments thereof and subsequently obtaining the resulting polyclonal antisera. According to the method of Köhler and Milstein and its developments monoclonal antibodies 5 can be obtained from the antibody-producing cells of the experimental animals by cell fusion in the known manner. In the same way, human monoclonal antibodies can be produced according to known methods. Antibodies of this type could then be used both for diagnostic 10 tests, in particular of endometriotic cell tissue, or else for the therapy.

For example, samples such as body fluids, in particular human body fluids (e.g. blood, lymph or CSF) may be 15 tested with the aid of the ELISA technique on the one hand for the presence of a polypeptide encoded by the endometriosis gene, on the other hand for the presence of autoantibodies against such a polypeptide. Polypeptides encoded by the endometriosis gene or 20 fragments thereof can then be detected in such samples with the aid of a specific antibody, for example of an antibody according to the invention. For detecting autoantibodies it is preferably possible to employ recombinant fusion proteins which contain a part or a 25 domain or even the complete polypeptide encoded by the endometriosis gene and which are fused to a protein domain which facilitates detection, for example maltose-binding protein (MBP).

30 Diagnostic tests may also be carried out with the aid of specific nucleic acid probes for detecting at the nucleic acid level, for example at the gene or transcript level.

35 Provision of the nucleotide and amino acid sequences and antibodies according to the invention further facilitates a targeted search for effectors of the polypeptides/proteins. Effectors are agents which act in an inhibitory or activating manner on the

- 10 -

polypeptide according to the invention and which are capable of selectively influencing cell functions controlled by the polypeptides. These may then be employed in the therapy of appropriate pathologies, 5 such as those based on invasive processes. The invention therefore also relates to a method for identifying effectors of endometriotic proteins where cells expressing the protein are brought into contact with various potential effector substances, for example 10 low molecular weight agents, and the cells are analyzed for modifications, for example cell-activating, cell-inhibiting, cell-proliferative and/or cell-genetic modifications. In this way it is also possible to identify binding targets of endometriotic proteins.

15

Since many neoplastic diseases are accompanied by invasive processes, the discovery of the gene according to the invention additionally provides possibilities for the diagnosis, prevention and therapy of cancerous 20 diseases.

The discovery of a gene involved in the responsibility for invasive processes not only opens up possibilities for the treatment of diseases based on cellular 25 modifications of this type, but the sequences according to the invention may also be used in order to make such processes usable. This can be of importance, for example, for the implantation of embryos.

30 The present invention therefore also relates to a pharmaceutical composition which includes as active components nucleic acids, vectors, cells, polypeptides, peptides and/or antibodies, as mentioned before.

35 The pharmaceutical composition according to the invention may further contain pharmaceutically conventional carriers, excipients and/or additives and, where appropriate, further active components. The pharmaceutical composition may be employed in

- 11 -

particular for the diagnosis, therapy or prevention of diseases associated with invasive processes. Furthermore the composition according to the invention may also be employed for diagnosing a predisposition 5 for such diseases, in particular for diagnosing an endometriosis risk.

The invention is illustrated in more detail by the following figures, sequence listings and examples.

10

Figure 1 shows a diagrammatic representation of the cDNA of the endometriosis-associated gene where only exons E1 to E5 are shown.

15 SEQ ID NO. 1 represents a nucleotide sequence which contains genetic information coding for the endometriosis-associated gene, where an open reading frame extends from nucleotide 3 to 911, and

20

SEQ ID NO. 2 represents the amino acid sequence of the open reading frame of the nucleotide sequence shown in SEQ ID NO. 1, where the amino acid sequence of the open reading frame extends from amino acid 1 25 to 302.

30

SEQ ID NO. 3 represents a nucleotide sequence like that of SEQ ID NO. 1 but it contains an additional 84 nucleotides at the 5' end, the open reading frame extends from nucleotide 3 to 995.

35

SEQ ID NO. 4 represents the amino acid sequence of the open reading frame of the nucleotide sequence shown in SEQ ID NO. 3, where this amino acid sequence has 320 amino acids of which the C-terminal 302 are identical to those in SEQ ID NO. 2.

SEQ ID NO. 5 represents of the nucleotide sequence of the possibly present additional exon 4a consisting of the 218 nt shown, where 5 exon 4a, if it is present, is located between nucleotide 1054 and 1055 (in relation to SEQ ID NO. 3).

#### EXAMPLES

10

##### **Example 1 Cell culturing**

To identify an endometriosis-associated gene, invasive and noninvasive cells of the epithelial endometriotic 15 cell line EEC145T<sup>+</sup> were used. The cells were cultured in Dulbecco's medium (DMEM) with 10% fetal calf serum and diluted 1 : 5 2x per week (passage). For comparison of the expression patterns by means of DDRT-PCR (see below) invasive cells of passage 17 and noninvasive 20 cells of passage 33 were used. The cells were transformed with SV40 and analyzed by differential display reverse transcription polymerase chain reaction (DDRT-PCR).

##### **25 Example 2 DDRT-PCR**

This method developed by Liang and Pardee is a method for distinguishing expression patterns of different cell types or the alteration in the expression pattern 30 of one cell type under different living conditions or during altering stages of development (Liang and Pardee (1992), Science 257, 967-971). The basis of the DDRT-PCR technique is based on the idea that in each cell about 15,000 genes are expressed and that in principle 35 each individual mRNA molecule can be prepared by means of reverse transcription and amplification with random primers.

In this example the cellular polyA<sup>+</sup> RNA was initially

- 13 -

transcribed into cDNA with the aid of several different dT<sub>11</sub>VX primers (downstream primers, anchor primers). The resulting cDNA populations were then PCR-amplified using 4 downstream and 20 upstream primers from the RNA 5 Map™ Kit from Genhunter, Nashville (1994), with the addition of a radiolabeled nucleotide. After the amplification the reaction mixtures were concentrated in vacuo and the obtained cDNA fragments were fractionated in a six-percent native PAA (polyacrylamide) gel. DNA detection was carried out by 10 autoradiography. PCR mixtures showing distinct differences in the band pattern for the two cell variants to be studied were repeated twice in order to 15 test reproducibility. If the previously found differences were confirmed, the bands were eluted from the gel according to known methods, reamplified, cloned and sequenced.

By this method a 394 bp fragment (fragment 1, 20 nucleotides 1235 to 1628 of the nucleic acid sequence depicted in SEQ ID NO. 1, see also Figure 1) was found which was specific for the invasive cell variant. This fragment 1 was used as a probe in Northern blot analysis (see below).

25

**Example 3 Analysis of the fragment 1 expression profile  
in human Northern blot analyses**

To test the expression pattern for DDRT-PCR fragment 1, 30 Northern blot analyses were carried out. For this 20 µg of total RNA or 4 µg of polyA+ RNA were fractionated in 1% denaturating agarose gels and transferred onto a nylon membrane overnight. The RNA was fixed to the membrane by irradiation with UV light. Hybridization 35 with <sup>32</sup>P-labeled probes (labeling by means of RPL kit from Amersham) took place overnight in a formamide-containing hybridization solution at 42°C. Subsequently the membrane was washed under increasing stringency until the spots of radioactive emission were of

measurable intensity. The hybridization pattern was visualized by putting on an X-ray film (NEF-NEN, DuPont) and exposing over several days. To determine the expression pattern for DDRT-PCR fragment 1, Northern blot analyses were carried out using RNA from the following cells or tissues:

- invasive cells of the epithelial endometriotic cell line EEC145T<sup>+</sup> (passage 17)
- noninvasive cells of the epithelial endometriotic cell line EEC145T<sup>+</sup> (passage 33)
- cells of the peritoneal cell line EEC143T<sup>+</sup>
- endometrial tissue
- cells of the invasive human bladder carcinoma cell line EJ28
- cells of the noninvasive human bladder carcinoma cell line RT112

After hybridization with the probe for DDRT-PCR fragment 1 an mRNA of about 4 kb was detectable, and it was exclusively detectable in the invasive variant of the endometriotic cell line EEC145T<sup>+</sup>.

Further human tissues were tested. In the spleen an mRNA of 4 kb in length was found which hybridized unambiguously with fragment 1, and in brain mRNAs of 4 kb and > 9 kb in length, respectively, were found.

Northern blot analyses were carried out according to the manufacturer's protocol using two human multiple tissue Northern (MTN) blots from Clontech. Expression was tested in the following tissues: colon, small intestine, heart, brain, testicles, liver, lung, spleen, kidney, ovaries, pancreas, peripheral blood leukocytes, placenta, prostate, skeletal muscle, thymus. The expression pattern obtained using the radiolabeled 3' probe "DDRT-PCR fragment 1" appears as follows:

4 kb mRNA (expected size): brain,  
spleen,

- 15 -

9.5 kb mRNA:	pancreas
	brain

In the remaining tissues no specific hybridization was  
5 detectable.

#### In-situ hybridization

To elucidate the cellular expression pattern, mRNA in-  
10 situ hybridizations were carried out on 10  $\mu\text{m}$  paraffin  
sections of different tissues. For this the "DDRT-PCR  
fragment 1" was employed as digoxigenin-labeled RNA  
probe. The detection reaction was carried out by means  
of a digoxigenin-specific antibody coupled to alkaline  
15 phosphatase (A). BM Purple served as a substrate for AP  
and forms a blue precipitate after dephosphorylation.  
The results are listed in the following table and show  
predominant expression in invasive/migrating cells.

Strong expression	Weak, not quite unambiguous expression
epithelial cells from endometriotic lesions	skeletal muscle
carcinomas	heart
lymphatic infiltrates	sarcomas
thymus	
germinal centers of lymph follicles (spleen)	
<b>somewhat weaker:</b>	
epithelial cells of the endometrium	
angiogenetic endothelial cells	
migrating nerve cells	

- 16 -

**Example 4 RT-PCR**

RT-PCR (reverse transcription PCR) provides a sensitive method for testing the expression pattern.

5

For this, 1  $\mu$ g of the appropriate polyA<sup>+</sup> RNA was transcribed into cDNA with the aid of 400 U of M-MLV reverse transcriptase (Gibco-BRL) in a total volume of 30  $\mu$ l. 1  $\mu$ l of this was employed for the subsequent PCR 10 with different primer combinations.

The PCR primers P1 to P7 used are depicted in table 1 (see figure 1).

15 Table 1

Number	Sequence (nucleotide position in relation to SEQ ID NO. 1)
P1	5' -CCAGCTGCTGCCAAATCC-3' (36-53)
P2	5' -CATCATGGTCATAGCTGC-3' (545-562)
P3	5' -AGCGTCTCATCGGTGTAC-3' (793-776, reverse primer)
P4	5' -AACAGAAGTGGTAGGTG-3' (1080-1063, reverse primer)
P5	5' -AAAGGGACGGGAGGAAGC-3' (1243-1260)
P6	5' -CCAAAGTAGAAAACACTG-3' (1612-1595, reverse primer)
P7	5' -GCTTGTATGACACACACG-3' (2150-2133, reverse primer)

RT-PCR experiments were carried out using polyA<sup>+</sup> RNA from different cell lines and tissues and using 20 different primer combinations. The results are depicted in table 2.

Table 2

PC	P17	P33	Per	EM	EJ28	RT112	E	EE	PEE
P1-P4	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
P2-P6	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
P5+P7	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
P5+P6	+	-	-	+	-	-	+	+	+
P1+P3	+	-	-	+	-	-	+	+	+

PC = primer combination  
 5 P17 = endometriotic cell line EEC145T, passage 17,  
       invasive  
 P33 = endometriotic cell line EEC145T, passage 33,  
       noninvasive  
 Per = peritoneal cell line Per143T  
 10 EM = endometrial tissue  
 EJ28 = invasive bladder carcinoma cell line  
 RT112 = noninvasive bladder carcinoma cell line  
 E = endometrial tissue  
 EE = endometrial tissue of an endometriosis  
 15 patient  
 PEE = peritoneal endometriosis biopsy  
 n.d. = not determined

The RT-PCR results confirmed the fragment 1-specific expression in the early passages (passage 17, passage 20) of the endometriotic cell line EEC145T<sup>+</sup>. As a deviation from the Northern blot analyses it was possible to show in addition a weak expression in the endometrium.

25 **RT-PCR analyses using intron-bridging primers**

To test possible alternative exons, RT-PCR experiments using intron-bridging primers were carried out. In this 30 connection it was possible to show at least one further mRNA species which exists alongside the mRNA described and which contains a further exon (4a) of 218 bp in length between the 4th and 5th exons. This exon is

- 18 -

located in the 3'-UTR (untranslated region), that is to say after the coding region. The sequence of exon 4a is listed below.

5

gcggttgtcc ggaatgccag tggctcctgg gcagatgtgc accccagatt  
cagcctttgt gatagattcc aacacgttct ggctcagac cacctttgtg  
gtggggccag actgctctgg gcaaagtgaa gctggcctt atgctccaaag  
gaagggggccc tcgagagcag gcctgcattg gctctcggac taattcgca  
tcatctttca tacagcag

Nucleotide sequence of the alternative exon 4a

10

Example 5 Preparation of the cDNA phage bank EEC14  
The cDNA phage bank EEC14 was prepared according to the method of Short, J.M. et al. (1988) Nucleic Acids Res. 16: 7583-7600.

15

Initially, reverse transcription of polyA<sup>+</sup> RNA from invasive cells (passage 17) of the epithelial endometriotic cell line EEC145T<sup>+</sup> was carried out. The primer used here consists of an *Xho*I cleavage site and a poly(dT) sequence of 18 nucleotides in length. An adapter including an *Eco*RI cleavage site was ligated to the cDNA fragments produced. The two restriction sites permit directed insertion of the cDNA fragments into the ZAP Express<sup>TM</sup> vector. Inserts can be excised from the phage in the form of a kanamycin-resistant pBK CMV phagemid.

20

25

#### Example 6 Phage bank screening

30

35

The DDRT-PCR fragment 1 (394 bp) was used as a probe in order to screen 10<sup>6</sup> pfu (plaque forming units) of the cDNA phage bank EEC14 according to the manufacturer's protocol (Stratagene). Labeling of the probe with digoxigenin (Boehringer Mannheim) was carried out with the aid of PCR. The plaques formed after infection of the bacterial strain *XL 1blue MRF'* were transferred onto a nylon membrane and hybridized thereon with the

- 19 -

abovementioned probe. Detection of the hybridized, digoxigenin-labeled probe was carried out according to the chemiluminescence protocol by Boehringer Mannheim.

5 Positive plaques were selected and subjected to rescreening. The positive plaques from the rescreening were employed for the excision. Excising the vector portion from the phage by means of ExAssist helper phages resulted in kanamycin-resistant pBK CMV 10 phagemids which could be isolated and sequenced after amplification in the bacterial strain XLOLR<sup>TM</sup>. The isolated phagemid clone Q2A contained the longest insert of 2.3 kb in size whose sequence was determined 15 and is shown SEQ ID NO. 1. The DDRT-PCR fragment 1 sequence is found as nucleotides 1235 to 1628 in relation to SEQ ID NO. 1.

**Example 7 Southern blot analysis**

20 10 µg of genomic DNA from female and male subjects were cleaved with various restriction endonucleases. The fragments were fractionated in an agarose gel and transferred onto a nylon membrane. Hybridization with the digoxigenin-labeled DDRT-PCR fragment 1 was carried 25 out on this membrane.

Hybridization was detectable by chemiluminescence according to the Boehringer protocol. Using various restriction endonucleases only one band in each case 30 was detected in both the female and male DNA samples. This result suggests that the gene on which fragment 1 is based is a single, non-sex-specific gene. Since then, two genomic clones PAC J1472 and PAC N1977 have been isolated using DDRT-PCR fragment 1.

35

**Example 8 Fluorescence in situ hybridization (FISH)**

The genomic clones obtained in Example 7 were localized on chromosome 1 (1p36) by means of fluorescence in situ

- 20 -

hybridization (Lichter et al. (1990), Science 247:64-69).

**Example 9 Production of specific antibodies**

5

Nucleotides 584 to 909 of the abovementioned cDNA sequence were cloned by suitable restriction cleavage sites into the expression vector pMAL cRI. To express the sequence the construct was transformed into E.coli 10 DH5  $\alpha$  cells. The translated protein fragment was cut out of an SDS polyacrylamide gel and employed for immunizing rabbits.

**Example 10 RACE (rapid amplification of cDNA ends)**

15

Since the length of the cDNA clone Q2A (see Example 6) differs from the size of the detected mRNA (about 4 kb), RACE experiments were carried out to obtain further sequence information. With the aid of this 20 method it is possible to obtain cDNA sequences from an mRNA template between a defined internal sequence and unknown sequences at the 5' or 3' end. The 3' end of clone Q2A could be confirmed by 3'RACE experiments starting from the 5th exon.

25

For the 5'RACE, first strand synthesis of the cDNA was carried out using a gene-specific primer which hybridizes in the 1st exon, and then a homopolymeric nucleotide tail was attached with the aid of the enzyme 30 terminal transferase. This attached sequence permitted amplification of the sequence region located between the gene-specific primer and the homopolymeric nucleotide tail. This made it possible to obtain the following additional sequence which is located 5' from 35 the Q2A sequence and belongs to the first exon:

- 21 -

cc cgg ccg ccc cga gtg gag cgg atc cac ggg cag atg cag atg cct 47  
 Arg Pro Pro Arg Val Glu Arg Ile His Gly Gln Met Gln Met Pro  
 1 5 10 15

cga gcc aga cgg gcc cac agg ccc cgg gac cag gcg gcc gcc ctc gtg ... 95  
 Arg Ala Arg Arg Ala His Arg Pro Arg Asp Gln Ala Ala Ala Leu Val ...  
 20 25 30

The underlined sequence represents the first nucleotides of the Q2A sequence, the sequence in front of it corresponds to the novel sequence obtained by 5'RACE. The open reading frame fits into the one already derived for fragment and contains two putative start codons (underlined).

10 The nucleotide sequence which has the sequence previously obtained and is depicted in SEQ ID NO. 1 and the additional 84 nt at the 5' end is depicted in SEQ ID NO. 3.

**Claims**

1. Nucleic acid,  
characterized in  
5 that it comprises  
(a) the nucleotide sequences depicted in SEQ ID NO. 1, 3 or/and 5, a combination or a protein-encoding segment thereof,  
(b) a nucleotide sequence corresponding to the sequence in (a) within the scope of the degeneracy of the genetic code or  
10 (c) a nucleotide sequence hybridizing with the sequences in (a) and/or (b) under stringent conditions,  
15 with the proviso that the nucleic acid is different from the sequences stated with accession numbers Z98886, Ac003017, Aa453993, AL023586 and Aa452856 in the EMBL EST database.
- 20 2. Nucleic acid according to Claim 1,  
characterized in  
that it comprises a protein-encoding segment of  
the nucleotide sequences depicted in SEQ ID NO. 1,  
3 or/and 5.
- 25 3. Nucleic acid according to Claim 1  
characterized in  
that it has a homology of more than 80% to the  
nucleotide sequences depicted in SEQ ID NO. 1, 3  
30 or/and 5.
- 35 4. Nucleic acid according to any one of Claims  
1 to 3,  
characterized in  
that it codes for a polypeptide associated with  
invasive processes or for a segment thereof.
5. Modified nucleic acid or nucleic acid analog which

comprises a nucleotide sequence according to any one of Claims 1 to 4.

6. Polypeptide,  
5 characterized in  
that it is encoded by a nucleic acid according to any one of Claims 1 to 4, where the proviso of Claim 1 is to be disregarded.
- 10 7. Polypeptide according to Claim 6,  
characterized in  
that it has  
(a) the amino acid sequence depicted in SEQ ID  
NO. 2 or 4 or  
15 (b) a homology of more than 70% to the amino acid  
sequence according to (a).
8. Modified polypeptide comprising an amino acid  
sequence according to Claim 6 or 7.  
20
9. Peptide,  
characterized in  
that it represents a segment of at least 10 amino  
acids of the amino acid sequence depicted in SEQ  
25 ID NO. 2 or 4.
10. Vector,  
characterized in  
that it has at least one copy of a nucleic acid  
30 according to any one of Claims 1 to 4.
11. Vector according to Claim 10,  
characterized in  
that it facilitates expression of the nucleic acid  
35 in a suitable host cell.
12. Cell,  
characterized in

that it has been transformed with a nucleic acid according to any one of Claims 1 to 4 or with a vector according to Claim 10 or 11.

5 13. Antibody against a polypeptide according to any one of Claims 6 to 8 or against a peptide according to Claim 9.

10 14. Antibody according to Claim 13 characterized in that it is directed against the complete polypeptide or against a fragment thereof selected from a segment of amino acids 1 to 330 from SEQ ID NO. 4.

15 15. Composition for pharmaceutical application, characterized in that it comprises as active component:  
(a) a nucleic acid according to any one of Claims 20 1 to 5, where the proviso of Claim 1 is to be disregarded,  
(b) a vector according to Claim 10 or 11,  
(c) a cell according to Claim 12,  
(d) a polypeptide according to any one of Claims 25 6 to 8,  
(e) a peptide according to Claim 9 and/or  
(f) an antibody according to Claim 13 or 14.

30 16. Composition according to Claim 15, characterized in that it additionally contains pharmaceutically conventional carriers, excipients and/or additives.

35 17. Use of a polypeptide according to any one of Claims 6 to 8 or of a fragment of this polypeptide for the preparation of an immunogen for producing antibodies.

18. Use of a composition according to Claim 15 or 16 for the preparation of an agent for the diagnosis of diseases related to invasive processes.

5

19. Use of a composition according to Claim 15 or 16 for the preparation of an agent for the diagnosis of a predisposition for diseases related to invasive processes.

10

20. Use of a composition according to Claim 15 or 16 for the preparation of an agent for the therapy or prevention of diseases related to invasive processes.

15

21. Use of a composition according to Claim 15 or 16 for the preparation of an agent for the diagnosis, therapy or prevention of endometriosis.

20

22. Use of a composition according to Claim 15 or 16 for the preparation of an agent for the diagnosis, therapy or prevention of neoplastic diseases.

25

23. Use of a composition according to Claim 15 or 16 for the preparation of an agent for gene therapy.

24. Use of a composition according to Claim 15 or 16 for the preparation of an antisense inhibitor.

30

25. Use of a composition according to Claim 15 or 16 for the preparation of an agent in the implantation of embryos.

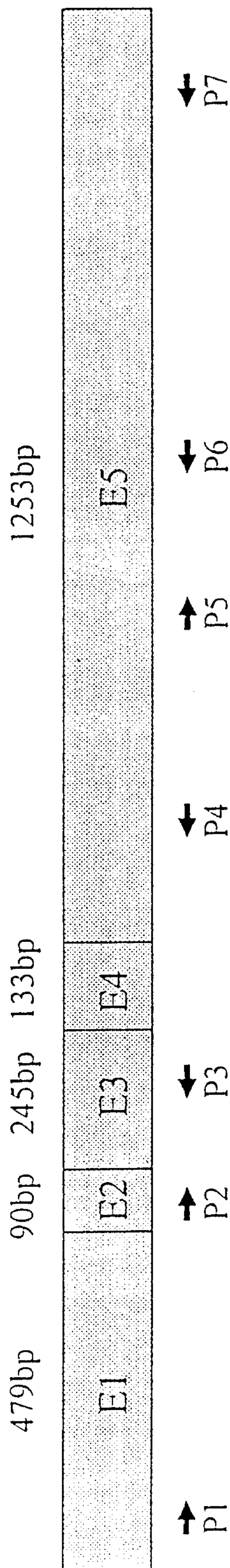
35

26. Use of a composition according to Claim 15 or 16 for the identification of inhibitors of a polypeptide according to any one of Claims 6 to 8 and/or of inhibitors of molecules which are capable of binding to the polypeptide.

27. Use of a polypeptide according to any one of  
Claims 6 to 8 or of a fragment of this polypeptide  
for the detection of antibodies against an  
endometriosis-associated protein or fragments  
thereof in a sample.  
5
28. Use of an antibody according to Claim 13 or 14 or  
of a fragment of this antibody for the detection  
of endometriosis-associated proteins or fragments  
thereof.  
10

Figure 1

DDRT-PCR-  
fragment-1 (394bp)



**E1-E5: Exon 1 up to exon 5**

The respective sizes of the exons are indicated above the drawing

**P1-P7: PCR-primer**

Sequences and exact designations of the primers are indicated in table 1.

The location of the sequence for the DDRT-PCR-fragment-1 has been illustrated by a black bar