(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 106 690 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
01.06.2005 Bulletin 2005/22

(21) Application number: **00310408.0**

(22) Date of filing: **23.11.2000**

(51) Int Cl.[7]: **C12N 15/12**, C07K 14/47, C12Q 1/68, C12N 5/10, C07K 16/18, G01N 33/53, G01N 33/557, C07K 19/00, A61K 38/17

(54) **Polynucleotide encoding autoantigens associated with endometriosis**

Endometriose-assoziierte-Autoantigene kodierendes Polynukleotid

Polynucléotide codant pour des auto-antigènes associés à l'endométriose

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

(30) Priority: **23.11.1999 US 447399**

(43) Date of publication of application:
13.06.2001 Bulletin 2001/24

(73) Proprietor: **DIAGNOSTIC PRODUCTS CORPORATION**
Los Angeles California 90045 (US)

(72) Inventors:
- **El Shami, A. Said**
  Camarillo, California 93032 (US)
- **Menon, Surendra Nath**
  Culver City, California 90232 (US)
- **French, Cynthia K.**
  Irvine, California 92612 (US)

(74) Representative:
Campbell, Patrick John Henry et al
J.A. Kemp & Co.,
14 South Square,
Gray's Inn
London WC1R 5JJ (GB)

(56) References cited:
WO-A-94/28021 **US-A- 5 880 261**

- **Emest24 Database Entry Hs888330 Accession number W19888; 5 May 1996 HILLIER ET AL.:"The WashU-Merck EST Project" XP002126944**
- **EMBL Database Entry AI832097 Accession number AI832097; 13 July 1999 ROBERT STRAUSBERG XP002141979**

EP 1 106 690 B1

Description

FIELD OF THE INVENTION

[0001] This invention is directed to the field of molecular biology in general, and, more specifically, to a polypeptide associated with endometriosis, an isolated polynucleotide encoding the polypeptide, and methods of using these molecules.

BACKGROUND OF THE INVENTION

[0002] Endometriosis is a painful disorder that is characterized by the ectopic implantation of functioning endometrial tissue into the abdominal wall and the outer surface of various organs including, most commonly, the lower bowel, ovaries and fallopian tubes. P. Vigano et al. (1991) *Fertility and Sterility* 56:894. Currently, endometriosis-specific genes have not been identified and the events relating to the development of endometriosis are poorly understood. However, several reports suggest that retrograde menstruation linked with abnormal immune function may play a role in establishing ectopic endometrium lesions. T. Ishimaru and H. Masuzaki (1991) *Am. J. Obstet. Gynecol*. 165:210--214. Many attempts to isolate antigens from ectopic endometrium lesions have failed, due to the necrotic nature of the lesions. WO 94/28021 describes isolation of an endometrial antigen from female patients with endometriosis.

[0003] Endometriosis also is recognized as having an autoimmune component. IgG and IgA auto-antibodies that react with multiple endometrial antigens have been documented in patients with endometriosis. However, attempts to develop IgG or IgA-based assays for the diagnosis of endometriosis has fallen short of fruition. S. Fernandez-Shaw et al., (1996) *Hum. Reprod.* 11:180-1184. R.A. Wild et al. (1991) *Obstetrics and Gynecology* 77:927. Studies have shown that circulating IgG antibodies that bind multiple endometrial proteins can be detected in women with endometriosis to varying degrees. Thirty-five percent to 74% of patients have sera reactive with endometrial proteins. O. Odukoya et al. (1996) *Acta Obstet. Gynecol. Scand.* 75:927-931; J.G. Kim et al. (1995) *Am. J. Reprod. Immunol.* 34:80-87; O.A. Odukoya et al. (1995) *Hum. Reprod.* 10:1214-1219. It has also been shown that endometrial antibody titers in patients that respond well to danazol are significantly lower (7/18 (39%) treated patients had elevated titers) than those patients with untreated endometriosis or patients that responded poorly to treatment (17/23 (74%) untreated patients had elevated titers). A. El-Roeiy et al. (1988) *Fertility and Sterility* 50:864-871; H.J. Chihal et al. (1986) *Fertility and Sterility* 46:408-411. In addition, it has been recently reported that women with endometriosis have elevated levels of IL-4, a Th2 mediating cytokine, and that treatment with danazol reduces the levels of IL-4 in women that respond well to treatment. C.-C. Hsu et al. (1997) *Fertility and Sterility* 67:1059-1064.

SUMMARY OF THE INVENTION

[0004] This invention provides an isolated cDNA molecule and an alternately spliced variant encoding autoantigens associated with endometriosis. The autoantigen is called Repro-EN-1.0. The alternately spliced variant is called IB1. Subjects diagnosed with endometriosis have been found to have antibodies that specifically bind to Repro-EN-1.0 polypeptide and/or a IB1 polypeptide. These antibodies represent a highly sensitive and specific diagnostic marker for endometriosis. Recombinant Repro-EN-1.0 protein and recombinant IB1 protein are useful to detect such antibodies in immunoassays.

[0005] In one aspect this invention provides a recombinant polynucleotide comprising a nucleotide sequence encoding polypeptide Repro-EN-1.0 of SEQ ID NO:2. or IB1 of SEQ ID NO:4 wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis. In one embodiment the nucleotide sequence is selected from the Repro-EN-1.0 sequence of SEQ ID NO:1 or IB1-sequence of SEQ ID NO: 3. In another embodiment the nucleotide sequence is identical to nucleotides 176 to 2755 of SEQ ID NO:1 or nucleotides 176 to 2986 of SEQ ID NO:3. In another embodiment the polynucleotide further comprises an expression control sequence operatively linked to the nucleofide sequence.

[0006] In another aspect this invention provides a polynucleotide primer pair which amplifies a nucleotide sequence encoding polypeptide Repro-EN-1.0 or IB1 wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis. The pair comprises: 1) a 3' primer of at least 7 nucleotides that specifically hybridizes to a 3' end of the nucleotide sequence or downstream from the sequence, and 2) a 5' primer of at least 7 nucleotides that specifically hybridizes to the 3' end of the complement of the nucleotide sequence or downstream from the complement of the sequence.

[0007] In another aspect this invention provides a recombinant cell comprising a recombinant polynucleotide comprising an expression control sequence operatively linked to a nucleotide sequence encoding polypeptide Repro-EN-1.0 of SEQ ID NO:2 or IB1 of SEQ ID NO:4, wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis.

[0008] In another aspect this invention provides a purified, recombinant Repro-EN-1.0 polypeptide whose amino acid sequence is identical to that of SEQ ID NO:2, or an allelic variant of SEQ ID NO:2, or an IB1 polypeptide whose amino acid sequence is identical to that of SEQ ID NO: 4, or an allelic variant of SEQ ID No: 4.

[0009] In another aspect this invention provides a purified polypeptide comprising Repro-EN-1.0 of SEQ ID NO: 2 or IB1 of SEQ ID NO: 4, wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis.

[0010] In another aspect this invention provides a composition consisting essentially of an antibody that specifically binds to Repro-EN-1.0 polypeptide of SEQ ID NO: 2, or IB1 polypeptide of SEQ ID NO: 4.

[0011] In another aspect this invention provides a method of diagnosing endometriosis in a subject. The method comprises the steps of: (a) detecting a test amount of an antibody that specifically binds to Repro-EN-1.0 polypeptide-whose amino acid sequence is identical to SEQ ID NO: 2 or an allelic variant of SEQ ID NO: 2 or IB1 polypeptide whose amino acid sequence is identical to SEQ ID NO: 4 or an allelic variant of SEQ ID NO: 4 in a sample from the subject; and (b) comparing the test amount with a normal range of the antibody in a control sample from a subject who does not suffer from endometriosis. A test amount above the normal range provides a positive indication in the diagnosis of endometriosis. The sample can be a blood product, e.g., serum, peritoneal fluid, menstrual fluid, vaginal secretion or urine. In one embodiment the antibody is an IgE, IgG or IG_4 immunoglobulin. In another embodiment the step of detecting comprises capturing the antibody from the sample with immobilized Repro-EN-1.0 or IB1, and detecting captured antibody. The step of detecting captured antibody can comprise contacting the captured antibody with a detectable antibody that specifically binds immunoglobulins and detecting binding between the captured antibody and the detectable antibody. In another embodiment the step of detecting can comprise capturing the antibody from the sample with an immobilized anti-immunoglobulin antibody and detecting captured antibody. The step of detecting captured antibody can comprise contacting the captured antibody with Repro-EN-1.0 or IB1, and detecting binding between the captured antibody and the Repro-EN-1.0 or IB1.

[0012] In another aspect this invention provides a method for use in following the progress of endometriosis in a subject. The method comprises the steps of: (a) detecting first and second amounts of an antibody that specifically bind Repro-EN-1.0 polypeptides whose amino acid sequence is identical to that of SEQ ID NO: 2 or an allelic variant of SEQ ID NO: 2 or IB1 polypeptide whose amino acid sequence is identical to that of SEQ ID NO: 4 or an allelic variant of SEQ ID NO: 4 in samples from the subject at a first and a second time, respectively; and (b) comparing the first and second amounts. An increase between the first and second amounts indicates progression of the endometriosis and a decrease between the first an second amounts indicates remission of the endometriosis.

[0013] In another aspect this invention provides a screening method for determining whether a compound increases or decreases the expression of Repro-EN-1 of SEQ ID NO: 2 or IB1 of SEQ ID NO: 4 in a cell, comprising contacting the cell with the compound and determining whether the-production of Repro-EN1.0 mRNA or polypeptide, or IB1 mRNA or polypeptide, is increased or decreased.

[0014] In another aspect this invention provides a method of detecting polymorphic forms of Repro-EN-1.0 of SEQ ID NO: 1 or 2 or IB1 of SEQ ID NO: 3 or 4 comprising the steps of: a) determining the identity of a nucleotide or amino acid at a selected position within the sequence of a test Repro-EN-1.0 gene orpolypeptide, or IB1 gene or polypeptide; b) determining the identity of the nucleotide or amino acid at the corresponding position of native Repro-EN-1.0 (gene or polypeptide of SEQ ID NO: 1 or 2 or IB1 gene or polypeptide of SEQ ID NO: 3 or 4) ; and c) comparing the identity from the test gene or polypeptide with the identity of the native gene or polypeptide, whereby a difference in identity indicates that the test gene or polypeptide a polymorphic form of Repro-EN-1.0 or IB1.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015]

Fig. 1 is a northern blot analysis of Repro-EN-1.0 expression in various tissues.
Fig. 2 is a northern blot analysis of Repro-EN-1.0 expression comparing various normal v. cancerous tissues.
Fig. 3 is a northern blot analysis of Repro-EN-1.0 expression in various tissue culture cells.

DETAILED DESCRIPTION OF THE INVENTION

I. DEFINITIONS

[0016] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY (2d ed. 1994); THE CAMBRIDGE DICTIONARY OF SCIENCE AND TECHNOLOGY

(Walker ed., 1988); THE GLOSSARY OF GENETICS, 5TH ED., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY (1991). As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[0017] . "Polynucleotide" refers to a polymer composed of nucleotide units. Polynucleotides include naturally occurring nucleic acids, such as deoxyribonucleic acid ("DNA") and ribonucleic acid ("RNA") as well as nucleic acid analogs. Nucleic acid analogs include those which include non-naturally occurring bases, nucleotides that engage in linkages with other nucleotides other than the naturally occurring phosphodiester bond or which include bases attached through linkages other than phosphodiester bonds. Thus, nucleotide analogs include, for example and without limitation, phosphorothioates, phosphorodithioates, phosphorotriesters, phosphoramidates, boranophosphates, methylphosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs), and the like. Such polynucleotides can be synthesized, for example, using an automated DNA synthesizer. The term "nucleic acid" typically refers to large polynucleotides. The term "oligonucleotide" typically refers to short polynucleotides, generally no greater than about 50 nucleotides. It will be understood that when a nucleotide sequence is represented by a DNA sequence (i.e., A, T, G, C), this also includes an RNA sequence (i.e., A, U, G, C) in which "U" replaces "T."

[0018] "cDNA" refers to a DNA that is complementary or identical to an mRNA, in either single stranded or double stranded form.

[0019] Conventional notation is used herein to describe polynucleotide sequences: the left-hand end of a single-stranded polynucleotide sequence is the 5'-end; the left-hand direction of a double-stranded polynucleotide sequence is referred to as the 5'-direction. The direction of 5' to 3' addition of nucleotides to nascent RNA transcripts is referred to as the transcription direction. The DNA strand having the same sequence as an mRNA is referred to as the "coding strand"; sequences on the DNA strand having the same sequence as an mRNA transcribed from that DNA and which are located 5' to the 5'-end of the RNA transcript are referred to as "upstream sequences"; sequences on the DNA strand having the same sequence as the RNA and which are 3' to the 3' end of the coding RNA transcript are referred to as "downstream sequences."

[0020] "Complementary" refers to the topological compatibility or matching together of interacting surfaces of two polynucleotides. Thus, the two molecules can be described as complementary, and furthermore, the con-tact surface characteristics are complementary to each other. A first polynucleotide is complementary to a second polynucleotide if the nucleotide sequence of the first polynucleotide is identical to the nucleotide sequence of the polynucleotide binding partner of the second polynucleotide. Thus, the polynucleotide whose sequence 5'-TATAC-3' is complementary to a polynucleotide whose sequence is 5'-GTATA-3'.

[0021] A nucleotide sequence is "substantially complementary" to a reference nucleotide sequence if the sequence complementary to the subject nucleotide sequence is substantially identical to the reference nucleotide sequence.

[0022] "Encoding" refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (i.e., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA produced by that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and non-coding strand, used as the template for transcription, of a gene or cDNA can be referred to as encoding the protein or other product of that gene or cDNA. Unless otherwise specified, a "nucleotide sequence encoding an amino acid sequence" includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns

[0023] "Recombinant polynucleotide" refers to a polynucleotide having sequences that are not naturally joined together. An amplified or assembled recombinant polynucleotide may be included in a suitable vector, and the vector can be used to transform a suitable host cell. A host cell that comprises the recombinant polynucleotide is referred to as a "recombinant host cell." The gene is then expressed in the recombinant host cell to produce, e.g., a "recombinant polypeptide." A recombinant polynucleotide may serve a non-coding function (e.g., promoter, origin of replication, ribosome-binding site, etc.) as well.

[0024] "Expression control sequence" refers to a nucleotide sequence in a polynucleotide that regulates the expression (transcription and/or translation) of a nucleotide sequence operatively linked thereto. "Operatively linked" refers to a functional relationship between two parts in which the activity of one part (e.g., the ability to regulate transcription) results in an action on the other part (e.g., transcription of the sequence). Expression control sequences can include, for example and without limitation, sequences of promoters (e.g., inducible or constitutive), enhancers, transcription terminators, a start codon (i.e., ATG), splicing signals for introns, and stop codons.

[0025] "Expression vector" refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operatively linked to a nucleotide sequence to be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or *in vitro* expression system. Expression vectors include all those known in the art, such as cosmids, plasmids (e.g., naked or contained in

liposomes) and viruses that incorporate the recombinant polynucleotide.

[0026] "Amplification" refers to any means by which a polynucleotide sequence is copied and thus expanded into a larger number of polynucleotide molecules, e.g., by reverse transcription, polymerase chain reaction, and ligase chain reaction.

[0027] "Primer" refers to a polynucleotide that is capable of specifically hybridizing to a designated polynucleotide template and providing a point of initiation for synthesis of a complementary polynucleotide. Such synthesis occurs when the polynucleotide primer is placed under conditions in which synthesis is induced, i.e., in the presence of nucleotides, a complementary polynucleotide template, and an agent for polymerization such as DNA polymerase. A primer is typically single-stranded, but may be double-stranded. Primers are typically deoxyribonucleic acids, but a wide variety of synthetic and naturally occurring primers are useful for many applications. A primer is complementary to the template to which it is designed to hybridize to serve as a site for the initiation of synthesis, but need not reflect the exact sequence of the template. In such a case, specific hybridization of the primer to the template depends on the stringency of the hybridization conditions. Primers can,be labeled with, e.g., chromogenic, radioactive, or fluorescent moieties and used as detectable moieties.

[0028] "Probe," when used in reference to a polynucleotide, refers to a polynucleotide that is capable of specifically hybridizing to a designated sequence of another polynucleotide. A probe specifically hybridizes to a target complementary polynucleotide, but need not reflect the exact complementary sequence of the template. In such a case, specific hybridization of the probe to the target depends on the stringency of the hybridization conditions. Probes can'be labeled with, e.g., chromogenic, radioactive, or fluorescent moieties and used as detectable moieties.

[0029] A first sequence is an "antisense sequence" with respect to a second sequence if a polynucleotide whose sequence is the first sequence specifically hybridizes with a polynucleotide whose sequence is the second sequence.

[0030] "Hybridizing specifically to" or "specific hybridization" or "selectively hybridize to", refers to the binding, duplexing, or hybridizing of a nucleic acid molecule preferentially to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA.

[0031] The term "stringent conditions" refers to conditions under which a probe will hybridize preferentially to its target subsequence, and to a lesser extent to, or not at all to, other sequences. "Stringent hybridization" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and northern hybridizations are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5° C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the Tm for a particular probe.

[0032] An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formalin with 1 mg of heparin at 42° C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.15 M NaCl at 72° C for about 15 minutes. An example of stringent wash conditions is a 0.2X SSC wash at 65° C for 15 minutes (see, Sambrook et al. for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1x SSC at 45° C for 15 minutes. An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6x SSC at 40° C for 15 minutes. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

[0033] "Polypeptide" refers to a polymer composed of amino acid residues, related naturally occurring structural variants, and synthetic non-naturally occurring analogs thereof linked via peptide bonds, related naturally occurring structural variants, and synthetic non-naturally occurring analogs thereof. Synthetic polypeptides can be synthesized, for example, using an automated polypeptide synthesizer. The term "protein" typically refers to large polypeptides. The term "peptide" typically refers to short polypeptides.

[0034] Conventional notation is used herein to portray polypeptide sequences: the left-hand end of a polypeptide sequence is the amino-terminus; the right-hand end of a polypeptide sequence is the carboxyl-terminus.

[0035] "Allelic variant" refers to any of two or more polymorphic forms of a gene occupying the same genetic locus. Allelic variations arise naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequences. "Allelic variants" also refer to cDNAs derived from mRNA transcripts of genetic allelic variants, as well as the proteins encoded by them.

[0036] The terms "identical" or percent "identity," in the context of two or more polynucleotide or polypeptide sequences, refer to two or more sequences or sub-sequences that are the same or have a specified percentage of

nucleotides or amino acid residues that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection.

[0037] "Antibody" refers to a polypeptide ligand substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof, which specifically binds and recognizes an epitope (e.g., an antigen). The recognized immunoglobulin genes include the kappa and lambda light chain constant region genes, the alpha, gamma, delta, epsilon and mu heavy chain constant region genes, and the myriad immunoglobulin variable region genes. Antibodies exist, e.g., as intact immunoglobulins or as a number of well characterized fragments produced by digestion with various peptidases. This includes, e.g., Fab' and F(ab)'2 fragments. The term "antibody," as used herein, also includes antibody fragments either produced by the modification of whole antibodies or those synthesized *de novo* using recombinant DNA methodologies. It also includes polyclonal antibodies, monoclonal antibodies, chimeric antibodies and humanized antibodies. "Fc" portion of an antibody refers to that portion of an immunoglobulin heavy chain that comprises one or more heavy chain constant region domains, CH1, CH2 and CH3, but does not include the heavy chain variable region.

[0038] A ligand or a receptor (e.g., an antibody) "specifically binds to" or "is specifically immunoreactive with" a compound analyte when the ligand or receptor functions in a binding reaction which is determinative of the presence of the analyte in a sample of heterogeneous compounds. Thus, under designated assay (e.g., immunoassay) conditions, the ligand or receptor binds preferentially to a particular analyte and does not bind in a significant amount to other compounds present in the sample. For example, a polynucleotide specifically binds under hybridization conditions to an analyte polynucledtide comprising a complementary sequence; an antibody specifically binds under immunoassay conditions to an antigen analyte bearing an epitope against which the antibody was raised; and an adsorbent specifically binds to an analyte under proper elution conditions.

[0039] "Immunoassay" refers to a method of detecting an analyte in a sample in which specificity for the analyte is conferred by the specific binding between an antibody and a ligand. This includes detecting an antibody analyte through specific binding between the antibody and a ligand. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual,* Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity.

[0040] "Substantially pure" or "isolated" means an object species is the predominant species present (i.e., on a molar basis, more abundant than any other individual macromolecular species in the composition), and a substantially purified fraction is a composition wherein the object species comprises at least about 50% (on a molar basis) of all macromolecular species present. Generally, a substantially pure composition means that about 80% to 90% or more of the macromolecular species present in the composition is the purified species of interest. The object species is purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods) if the composition consists essentially of a single macromolecular species. Solvent species, small molecules (<500 Daltons), stabilizers (e.g., BSA), and elemental ion species are not considered macromolecular species for purposes of this definition.

[0041] "Naturally-occurring" as applied to an object refers to the fact that the object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including viruses) that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally-occurring.

[0042] "Detecting" refers to determining the presence, absence, or amount of an analyte in a sample, and can include quantifying the amount of the analyte in a sample or per cell in a sample.

[0043] "Detectable moiety" or a "label" refers to a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include ^{32}P, ^{35}S, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin-streptavadin, dioxigenin, haptens and proteins for which antisera or monoclonal antibodies are available, or nucleic acid molecules with a sequence complementary to a target. The detectable moiety often generates a measurable signal, such as a radioactive, chromogenic, or fluorescent signal, that can be used to quantitate the amount of bound detectable moiety in a sample. The detectable moiety can be incorporated in or attached to a primer or probe either covalently, or through ionic, van der Waals or hydrogen bonds, e.g., incorporation of radioactive nucleotides, or biotinylated nucleotides that are recognized by streptavadin. The detectable moiety may be directly or indirectly detectable. Indirect detection can involve the binding of a second directly or indirectly detectable moiety to the detectable moiety. For example, the detectable moiety can be the ligand of a binding partner, such as biotin, which is a binding partner for streptavadin, or a nucleotide sequence, which is the binding partner for a complementary sequence, to which it can specifically hybridize. The binding partner may itself be directly detectable, for example, an antibody may be itself labeled with a fluorescent molecule. The binding partner also may be indirectly detectable, for example, a nucleic acid having a complementary nucleotide sequence can be a part of a branched DNA molecule that is in turn detectable through hybridization with other labeled nucleic acid molecules. (See, e.g., PD. Fahrlander and A. Klausner, *Bio/Technology* (1988) 6:1165.) Quantitation of the signal is achieved by, e.g., scintillation counting, densitometry, or flow cytometry.

[0044] A "subject" of diagnosis or treatment is a human or non-human mammal. Non-human mammals subject to

diagnosis or treatment include, for example, primates, ungulates, canines and felines.

[0045] "Diagnostic" means identifying the presence or nature of a pathologic condition. Diagnostic methods differ in their sensitivity and specificity. The "sensitivity" of a diagnostic assay is the percentage of diseased individuals who test positive (percent of true positives). The "specificity" of a diagnostic assay is 1 minus the false positive rate, where the false positive rate is defined as the proportion of those without the disease who test positive. While a particular diagnostic method may not provide a definitive diagnosis of a condition, it suffices if the method provides a positive indication that aids in diagnosis.

[0046] "Prognostic" means predicting the probable development (e.g., severity) of a pathologic condition.

[0047] "Test amount" refers to an amount of an analyte in a subject sample, which is then compared to a normal amount of the analyte in a sample (e.g., from a healthy individual) such that the relative comparison of the values provides a reference value for diagnosing a designated disease. Depending upon the method of detection, the test amount may be a determination of the amount of the analyte, but it is not necessarily an amount. The test amount may also be a relative value, such as a plus or a minus score, and also includes an amount indicating the presence or absence of the analyte in a sample.

[0048] "Normal amount" refers to an amount or a range of an analyte in a biological sample that indicates health or lack of pathology.

[0049] . "Diagnostic amount" refers to an amount of an analyte in a subject sample that is consistent with a particular diagnosis for a designated disease.

[0050] "Prognostic amount" refers to an amount or range of an analyte in a subject sample that is consistent with a particular prognosis for a designated disease.

[0051] "Plurality" means at least two.

[0052] An "epitope" is portion of a molecule that specifically binds to an antibody or a T cell receptor. A peptide epitope generally comprises a sequence of at least 6 amino acids from a polypeptide, although longer and shorter peptides can constitute epitopes.

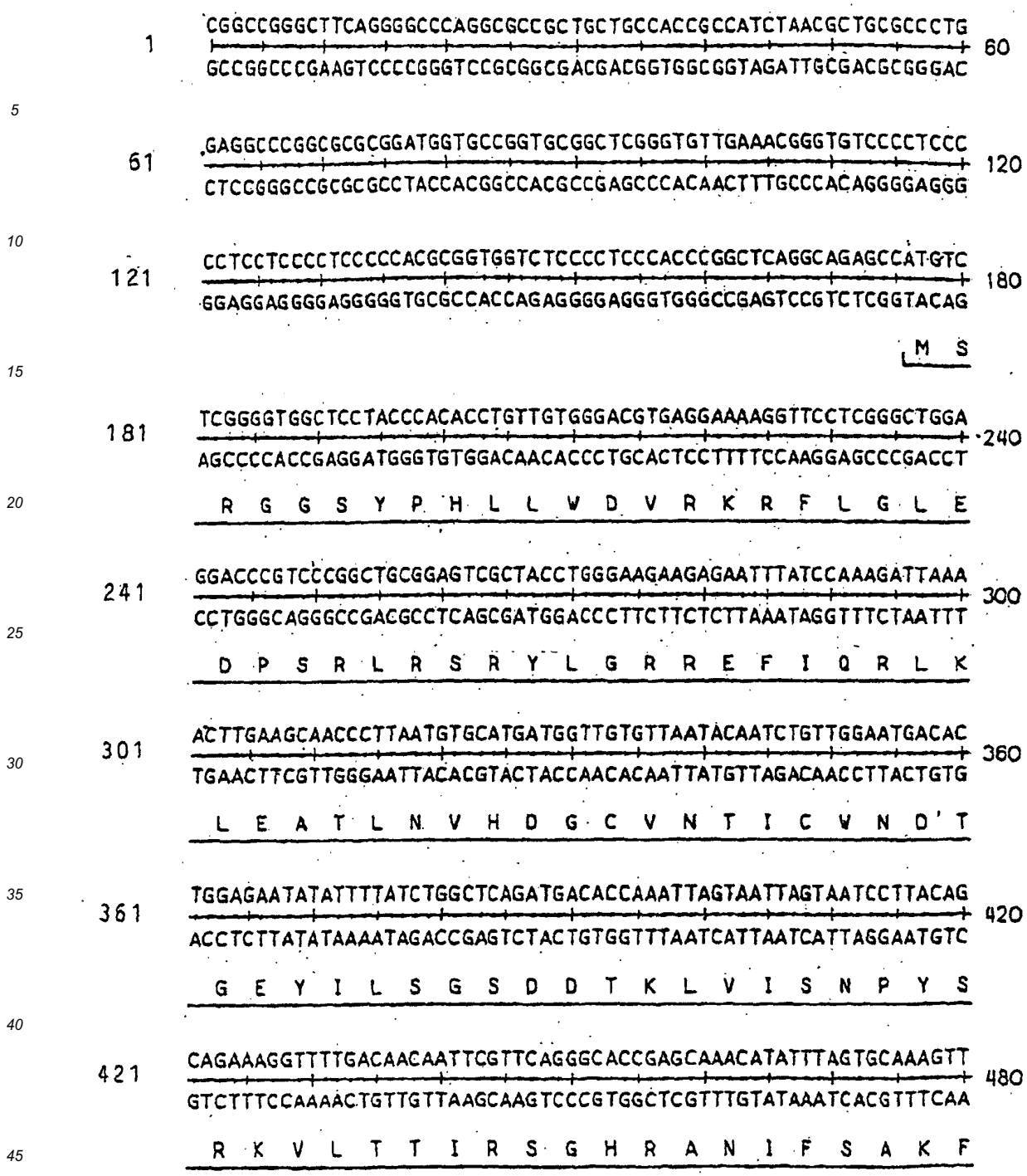
II. cDNA ENCODING REPRO-EN-1.0 and IB1

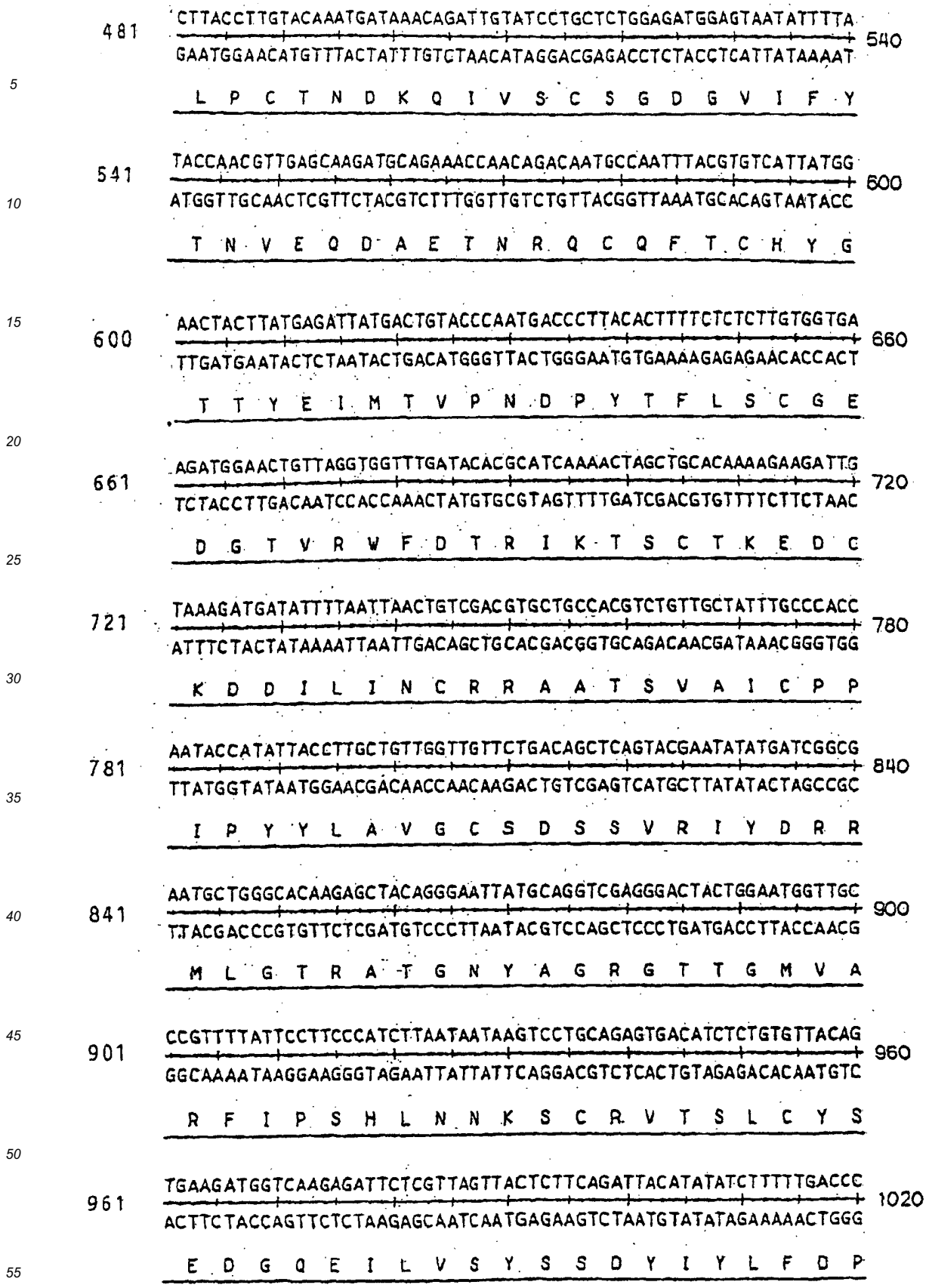
[0053] We have isolated a cDNA molecule encoding an autoantigen associated with endometriosis. The autoantigen is called Repro-EN-1.0. The presence of antibodies that specifically bind to an epitope of the Repro-EN-1.0 polypeptide is a highly sensitive and specific diagnostic marker for endometriosis.

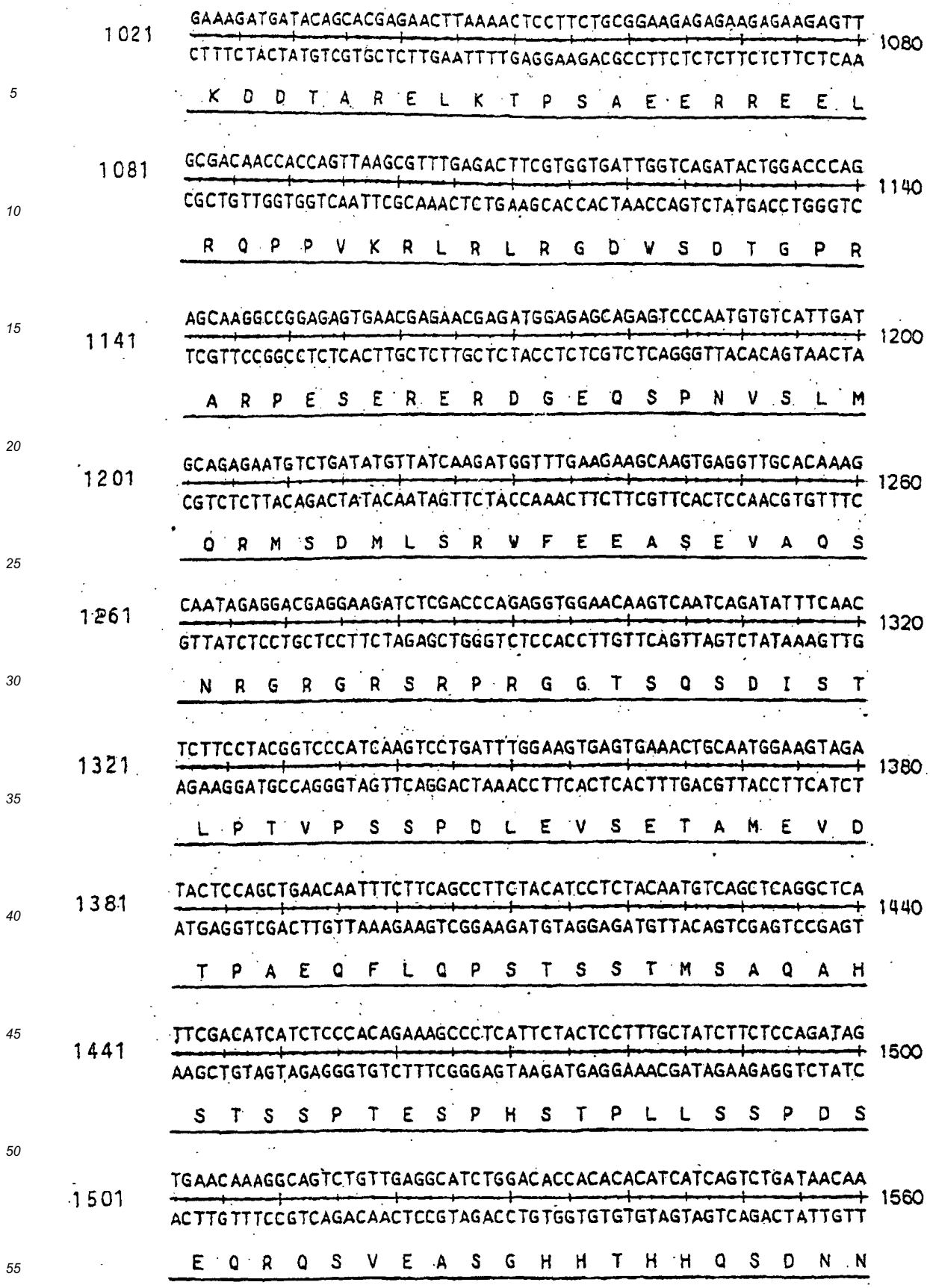
[0054] Polynucleotides encoding full-length Repro-EN-1.0 are useful in recombinant production of Repro-EN-1.0.

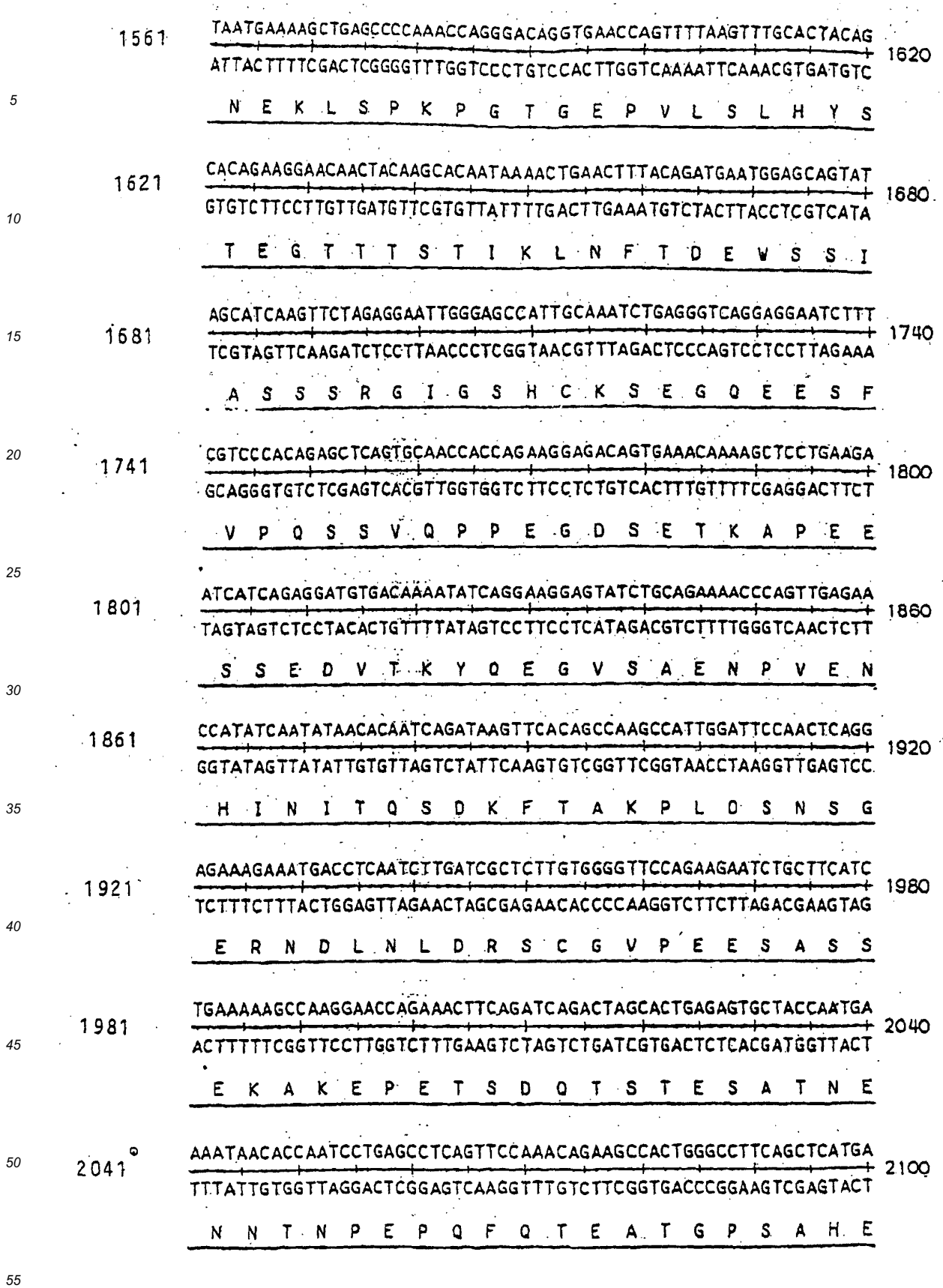
[0055] The Repro-EN-1.0 polypeptide is useful as a positive control in diagnostic assays to detect antibodies that specifically bind to Repro-EN-1.0 from patient serum samples. The polypeptide is also useful as an immunogen for eliciting production of antibodies against epitopes of the protein.

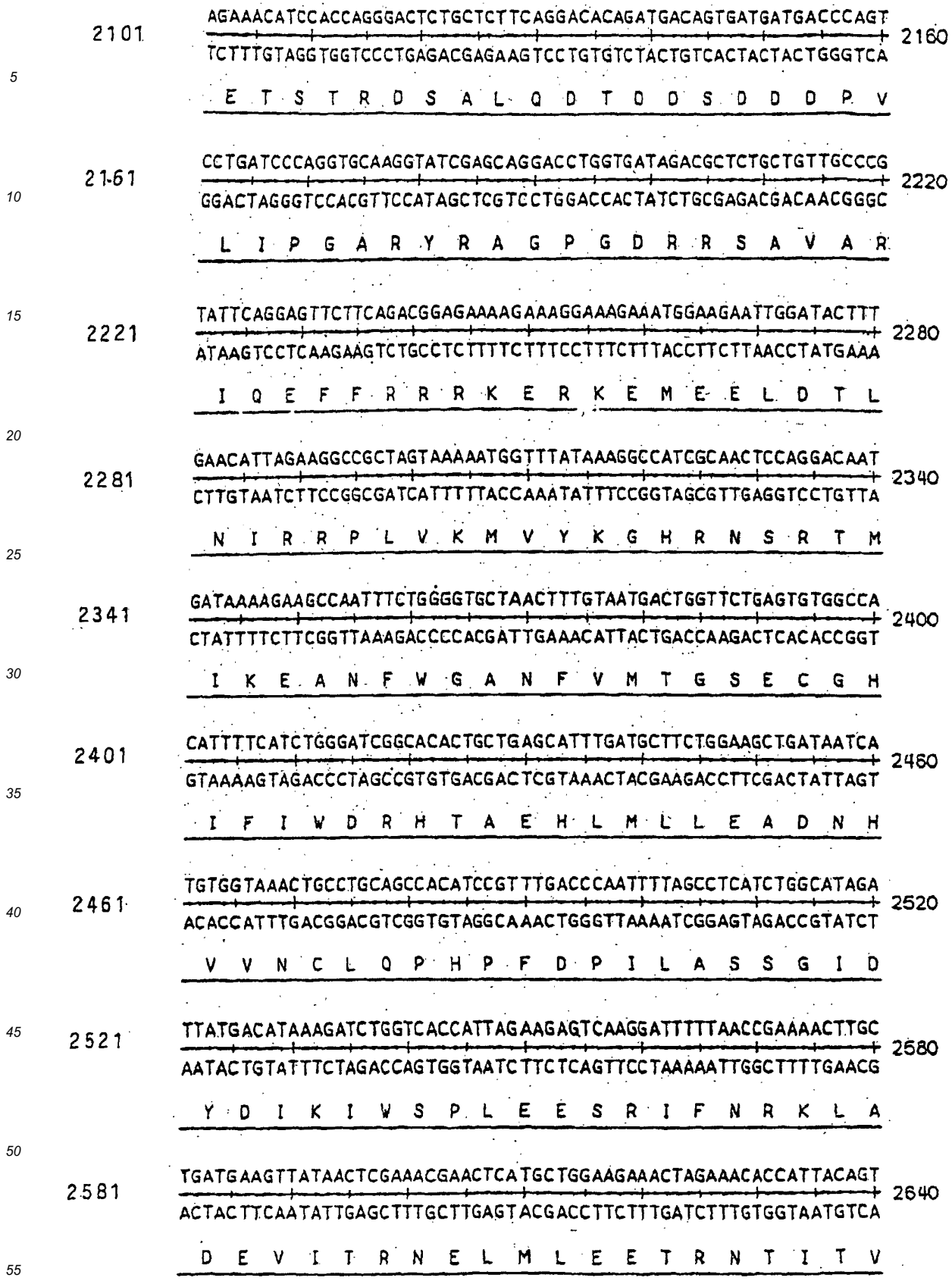
[0056] The nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of Repro-EN-1.0 follow:

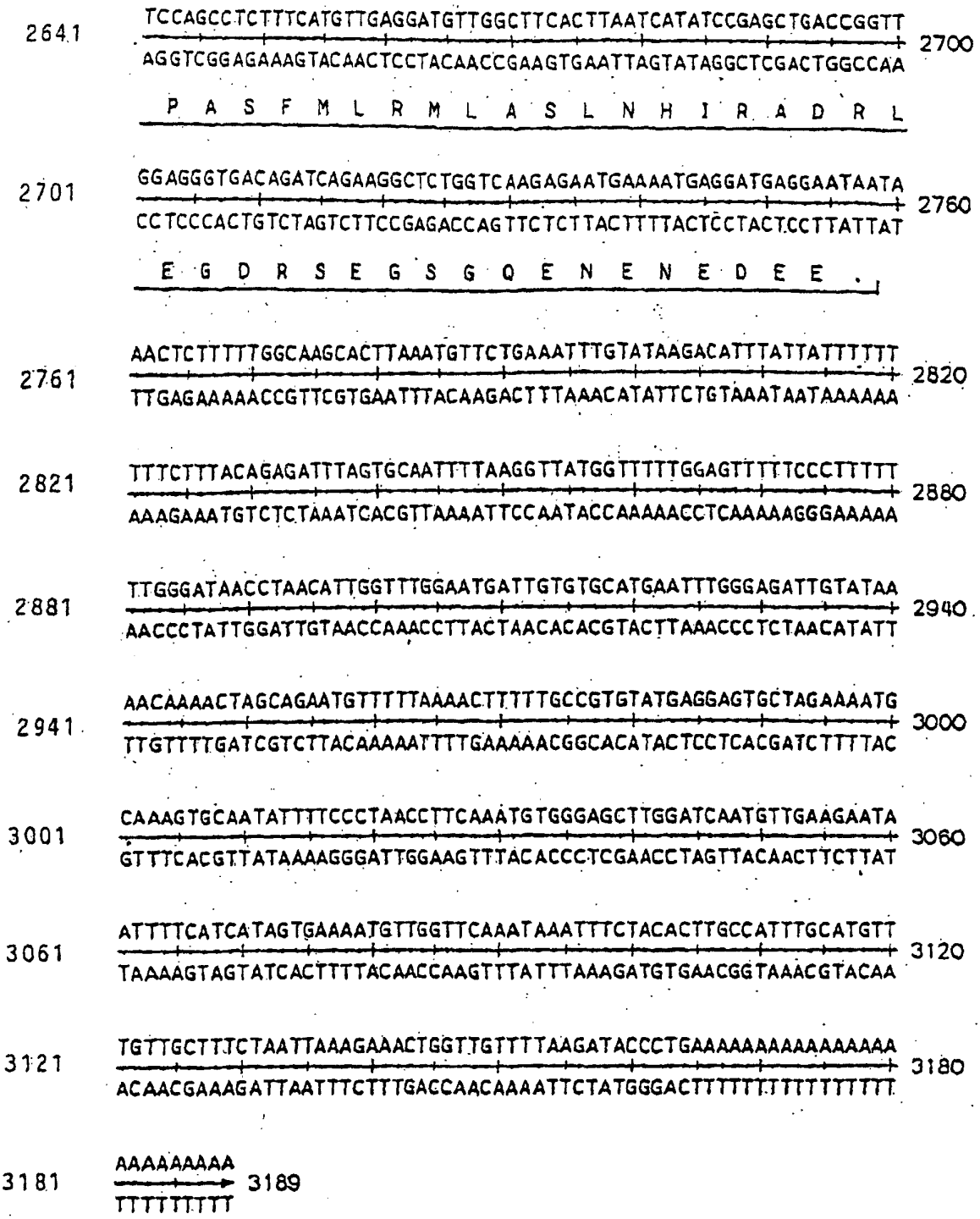

1
CGGCCGGGCTTCAGGGGCCCAGGCGCCGCTGCTGCCACCGCCATCTAACGCTGCGCCCTG
GCCGGCCCGAAGTCCCCGGGTCCGCGGCGACGACGGTGGCGGTAGATTGCGACGCGGGAC
60
61
GAGGCCCGGCGCGCGGATGGTGCCGGTGCGGCTCGGGTGTTGAAACGGGTGTCCCCTCCC
CTCCGGGCCGCGCGCCTACCACGGCCACGCCGAGCCCACAACTTTGCCCACAGGGGAGGG
120
121
CCTCCTCCCCTCCCCCACGCGGTGGTCTCCCCTCCCACCCGGCTCAGGCAGAGCCATGTC
GGAGGAGGGGAGGGGGTGCGCCACCAGAGGGGAGGGTGGGCCGAGTCCGTCTCGGTACAG
180
M S
181
TCGGGGTGGCTCCTACCCACACCTGTTGTGGGACGTGAGGAAAAGGTTCCTCGGGCTGGA
AGCCCCACCGAGGATGGGTGTGGACAACACCCTGCACTCCTTTTCCAAGGAGCCCGACCT
240
R G G S Y P H L L W D V R K R F L G L E
241
GGACCCGTCCCGGCTGCGGAGTCGCTACCTGGGAAGAAGAGAATTTATCCAAAGATTAAA
CCTGGGCAGGGCCGACGCCTCAGCGATGGACCCTTCTTCTCTTAAATAGGTTTCTAATTT
300
D P S R L R S R Y L G R R E F I Q R L K
301
ACTTGAAGCAACCCTTAATGTGCATGATGGTTGTGTTAATACAATCTGTTGGAATGACAC
TGAACTTCGTTGGGAATTACACGTACTACCAACACAATTATGTTAGACAACCTTACTGTG
360
L E A T L N V H D G C V N T I C W N D T
361
TGGAGAATATATTTTATCTGGCTCAGATGACACCAAATTAGTAATTAGTAATCCTTACAG
ACCTCTTATATAAAATAGACCGAGTCTACTGTGGTTTAATCATTAATCATTAGGAATGTC
420
G E Y I L S G S D D T K L V I S N P Y S
421
CAGAAAGGTTTTGACAACAATTCGTTCAGGGCACCGAGCAAACATATTTAGTGCAAAGTT
GTCTTTCCAAAACTGTTGTTAAGCAAGTCCCGTGGCTCGTTTGTATAAATCACGTTTCAA
480
R K V L T T I R S G H R A N I F S A K F


481
CTTACCTTGTACAAATGATAAACAGATTGTATCCTGCTCTGGAGATGGAGTAATATTTTA
GAATGGAACATGTTTACTATTTGTCTAACATAGGACGAGACCTCTACCTCATTATAAAAT
540
L P C T N D K Q I V S C S G D G V I F Y
541
TACCAACGTTGAGCAAGATGCAGAAACCAACAGACAATGCCAATTTACGTGTCATTATGG
ATGGTTGCAACTCGTTCTACGTCTTTGGTTGTCTGTTACGGTTAAATGCACAGTAATACC
600
T N V E Q D A E T N R Q C Q F T C H Y G
600
AACTACTTATGAGATTATGACTGTACCCAATGACCCTTACACTTTTCTCTCTTGTGGTGA
TTGATGAATACTCTAATACTGACATGGGTTACTGGGAATGTGAAAAGAGAGAACACCACT
660
T T Y E I M T V P N D P Y T F L S C G E
661
AGATGGAACTGTTAGGTGGTTTGATACACGCATCAAAACTAGCTGCACAAAAGAAGATTG
TCTACCTTGACAATCCACCAAACTATGTGCGTAGTTTTGATCGACGTGTTTTCTTCTAAC
720
D G T V R W F D T R I K T S C T K E D C
721
TAAAGATGATATTTTAATTAACTGTCGACGTGCTGCCACGTCTGTTGCTATTTGCCCACC
ATTTCTACTATAAAATTAATTGACAGCTGCACGACGGTGCAGACAACGATAAACGGGTGG
780
K D D I L I N C R R A A T S V A I C P P
781
AATACCATATTACCTTGCTGTTGGTTGTTCTGACAGCTCAGTACGAATATATGATCGGCG
TTATGGTATAATGGAACGACAACCAACAAGACTGTCGAGTCATGCTTATATACTAGCCGC
840
I P Y Y L A V G C S D S S V R I Y D R R
841
AATGCTGGGCACAAGAGCTACAGGGAATTATGCAGGTCGAGGGACTACTGGAATGGTTGC
TTACGACCCGTGTTCTCGATGTCCCTTAATACGTCCAGCTCCCTGATGACCTTACCAACG
900
M L G T R A T G N Y A G R G T T G M V A
901
CCGTTTTATTCCTTCCCATCTTAATAATAAGTCCTGCAGAGTGACATCTCTGTGTTACAG
GGCAAAATAAGGAAGGGTAGAATTATTATTCAGGACGTCTCACTGTAGAGACACAATGTC
960
R F I P S H L N N K S C R V T S L C Y S
961
TGAAGATGGTCAAGAGATTCTCGTTAGTTACTCTTCAGATTACATATATCTTTTTGACCC
ACTTCTACCAGTTCTCTAAGAGCAATCAATGAGAAGTCTAATGTATATAGAAAAACTGGG
1020
E D G Q E I L V S Y S S D Y I Y L F D P


1021
GAAAGATGATACAGCACGAGAACTTAAAACTCCTTCTGCGGAAGAGAGAAGAGAAGAGTT
CTTTCTACTATGTCGTGCTCTTGAATTTTGAGGAAGACGCCTTCTCTCTTCTCTTCTCAA
1080
K D D T A R E L K T P S A E E R R E E L
1081
GCGACAACCACCAGTTAAGCGTTTGAGACTTCGTGGTGATTGGTCAGATACTGGACCCAG
CGCTGTTGGTGGTCAATTCGCAAACTCTGAAGCACCACTAACCAGTCTATGACCTGGGTC
1140
R Q P P V K R L R L R G D W S D T G P R
1141
AGCAAGGCCGGAGAGTGAACGAGAACGAGATGGAGAGCAGAGTCCCAATGTGTCATTGAT
TCGTTCCGGCCTCTCACTTGCTCTTGCTCTACCTCTCGTCTCAGGGTTACACAGTAACTA
1200
A R P E S E R E R D G E Q S P N V S L M
1201
GCAGAGAATGTCTGATATGTTATCAAGATGGTTTGAAGAAGCAAGTGAGGTTGCACAAAG
CGTCTCTTACAGACTATACAATAGTTCTACCAAACTTCTTCGTTCACTCCAACGTGTTTC
1260
Q R M S D M L S R W F E E A S E V A Q S
1261
CAATAGAGGACGAGGAAGATCTCGACCCAGAGGTGGAACAAGTCAATCAGATATTTCAAC
GTTATCTCCTGCTCCTTCTAGAGCTGGGTCTCCACCTTGTTCAGTTAGTCTATAAAGTTG
1320
N R G R G R S R P R G G T S Q S D I S T
1321
TCTTCCTACGGTCCCATCAAGTCCTGATTTGGAAGTGAGTGAAACTGCAATGGAAGTAGA
AGAAGGATGCCAGGGTAGTTCAGGACTAAACCTTCACTCACTTTGACGTTACCTTCATCT
1380
L P T V P S S P D L E V S E T A M E V D
1381
TACTCCAGCTGAACAATTTCTTCAGCCTTCTACATCCTCTACAATGTCAGCTCAGGCTCA
ATGAGGTCGACTTGTTAAAGAAGTCGGAAGATGTAGGAGATGTTACAGTCGAGTCCGAGT
1440
T P A E Q F L Q P S T S S T M S A Q A H
1441
TTCGACATCATCTCCCACAGAAAGCCCTCATTCTACTCCTTTGCTATCTTCTCCAGATAG
AAGCTGTAGTAGAGGGTGTCTTTCGGGAGTAAGATGAGGAAACGATAGAAGAGGTCTATC
1500
S T S S P T E S P H S T P L L S S P D S
1501
TGAACAAAGGCAGTCTGTTGAGGCATCTGGACACCACACACATCATCAGTCTGATAACAA
ACTTGTTTCCGTCAGACAACTCCGTAGACCTGTGGTGTGTGTAGTAGTCAGACTATTGTT
1560
E Q R Q S V E A S G H H T H H Q S D N N


1561
TAATGAAAAGCTGAGCCCCAAACCAGGGACAGGTGAACCAGTTTTAAGTTTGCACTACAG
ATTACTTTTCGACTCGGGGTTTGGTCCCTGTCCACTTGGTCAAAATTCAAACGTGATGTC
1620
N E K L S P K P G T G E P V L S L H Y S
1621
CACAGAAGGAACAACTACAAGCACAATAAAACTGAACTTTACAGATGAATGGAGCAGTAT
GTGTCTTCCTTGTTGATGTTCGTGTTATTTTGACTTGAAATGTCTACTTACCTCGTCATA
1680
T E G T T T S T I K L N F T D E W S S I
1681
AGCATCAAGTTCTAGAGGAATTGGGAGCCATTGCAAATCTGAGGGTCAGGAGGAATCTTT
TCGTAGTTCAAGATCTCCTTAACCCTCGGTAACGTTTAGACTCCCAGTCCTCCTTAGAAA
1740
A S S S R G I G S H C K S E G Q E E S F
1741
CGTCCCACAGAGCTCAGTGCAACCACCAGAAGGAGACAGTGAAACAAAAGCTCCTGAAGA
GCAGGGTGTCTCGAGTCACGTTGGTGGTCTTCCTCTGTCACTTTGTTTTCGAGGACTTCT
1800
V P Q S S V Q P P E G D S E T K A P E E
1801
ATCATCAGAGGATGTGACAAAATATCAGGAAGGAGTATCTGCAGAAAACCCAGTTGAGAA
TAGTAGTCTCCTACACTGTTTTATAGTCCTTCCTCATAGACGTCTTTTGGGTCAACTCTT
1860
S S E D V T K Y Q E G V S A E N P V E N
1861
CCATATCAATATAACACAATCAGATAAGTTCACAGCCAAGCCATTGGATTCCAACTCAGG
GGTATAGTTATATTGTGTTAGTCTATTCAAGTGTCGGTTCGGTAACCTAAGGTTGAGTCC
1920
H I N I T Q S D K F T A K P L Q S N S G
1921
AGAAAGAAATGACCTCAATCTTGATCGCTCTTGTGGGGTTCCAGAAGAATCTGCTTCATC
TCTTTCTTTACTGGAGTTAGAACTAGCGAGAACACCCCAAGGTCTTCTTAGACGAAGTAG
1980
E R N D L N L D R S C G V P E E S A S S
1981
TGAAAAAGCCAAGGAACCAGAAACTTCAGATCAGACTAGCACTGAGAGTGCTACCAATGA
ACTTTTTCGGTTCCTTGGTCTTTGAAGTCTAGTCTGATCGTGACTCTCACGATGGTTACT
2040
E K A K E P E T S D Q T S T E S A T N E
2041
AAATAACACCAATCCTGAGCCTCAGTTCCAAACAGAAGCCACTGGGCCTTCAGCTCATGA
TTTATTGTGGTTAGGACTCGGAGTCAAGGTTTGTCTTCGGTGACCCGGAAGTCGAGTACT
2100
N N T N P E P Q F Q T E A T G P S A H E


2101
AGAAACATCCACCAGGGACTCTGCTCTTCAGGACACAGATGACAGTGATGATGACCCAGT
TCTTTGTAGGTGGTCCCTGAGACGAGAAGTCCTGTGTCTACTGTCACTACTACTGGGTCA
2160
E T S T R D S A L Q D T D D S D D D P V
2161
CCTGATCCCAGGTGCAAGGTATCGAGCAGGACCTGGTGATAGACGCTCTGCTGTTGCCCG
GGACTAGGGTCCACGTTCCATAGCTCGTCCTGGACCACTATCTGCGAGACGACAACGGGC
2220
L I P G A R Y R A G P G D R R S A V A R
2221
TATTCAGGAGTTCTTCAGACGGAGAAAAGAAAGGAAAGAAATGGAAGAATTGGATACTTT
ATAAGTCCTCAAGAAGTCTGCCTCTTTTCTTTCCTTTCTTTACCTTCTTAACCTATGAAA
2280
I Q E F F R R R K E R K E M E E L D T L
2281
GAACATTAGAAGGCCGCTAGTAAAAATGGTTTATAAAGGCCATCGCAACTCCAGGACAAT
CTTGTAATCTTCCGGCGATCATTTTTACCAAATATTTCCGGTAGCGTTGAGGTCCTGTTA
2340
N I R R P L V K M V Y K G H R N S R T M
2341
GATAAAAGAAGCCAATTTCTGGGGTGCTAACTTTGTAATGACTGGTTCTGAGTGTGGCCA
CTATTTTCTTCGGTTAAAGACCCCACGATTGAAACATTACTGACCAAGACTCACACCGGT
2400
I K E A N F W G A N F V M T G S E C G H
2401
CATTTTCATCTGGGATCGGCACACTGCTGAGCATTTGATGCTTCTGGAAGCTGATAATCA
GTAAAAGTAGACCCTAGCCGTGTGACGACTCGTAAACTACGAAGACCTTCGACTATTAGT
2460
I F I W D R H T A E H L M L L E A D N H
2461
TGTGGTAAACTGCCTGCAGCCACATCCGTTTGACCCAATTTTAGCCTCATCTGGCATAGA
ACACCATTTGACGGACGTCGGTGTAGGCAAACTGGGTTAAAATCGGAGTAGACCGTATCT
2520
V V N C L Q P H P F D P I L A S S G I D
2521
TTATGACATAAAGATCTGGTCACCATTAGAAGAGTCAAGGATTTTTAACCGAAAACTTGC
AATACTGTATTTCTAGACCAGTGGTAATCTTCTCAGTTCCTAAAAATTGGCTTTTGAACG
2580
Y D I K I W S P L E E S R I F N R K L A
2581
TGATGAAGTTATAACTCGAAACGAACTCATGCTGGAAGAAACTAGAAACACCATTACAGT
ACTACTTCAATATTGAGCTTTGCTTGAGTACGACCTTCTTTGATCTTTGTGGTAATGTCA
2640
D E V I T R N E L M L E E T R N T I T V



[0057] This 3164-base nucleotide sequence contains an open reading frame of 2580 nucleotides encoding Repro-EN-1.0 from nucleotide 176 to nucleotide 2755. The deduced amino acid sequence of Repro-EN-1.0 has 860 amino acids. Repro-EN-1.0 has a calculated molecular mass of 96.4 kD and a pI of 5.08.

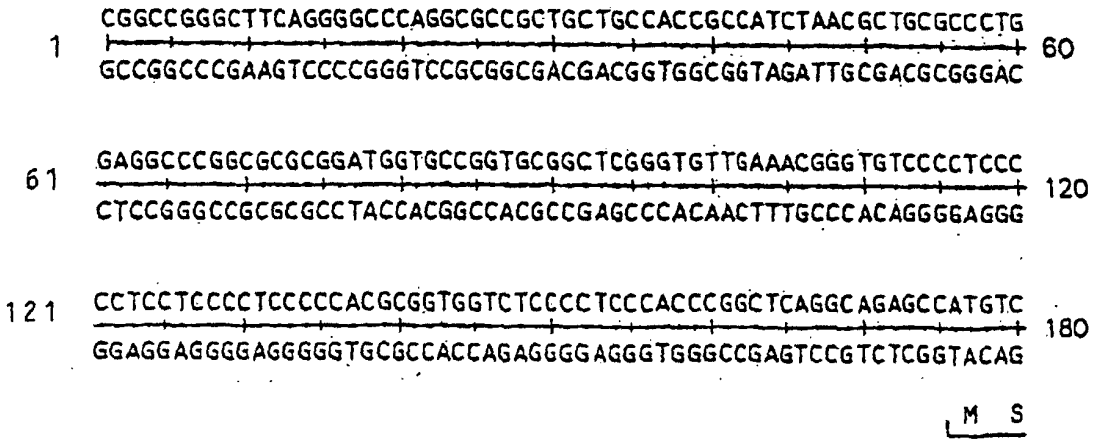
[0058] The Repro-EN-1.0 gene encodes a 3.4 kb mRNA. This mRNA is expressed primarily in skeletal muscle, heart and testis, and to a lesser extent in other tissues.
However, it is not detected in lung or peripheral blood mononuclear cells (PBMC).

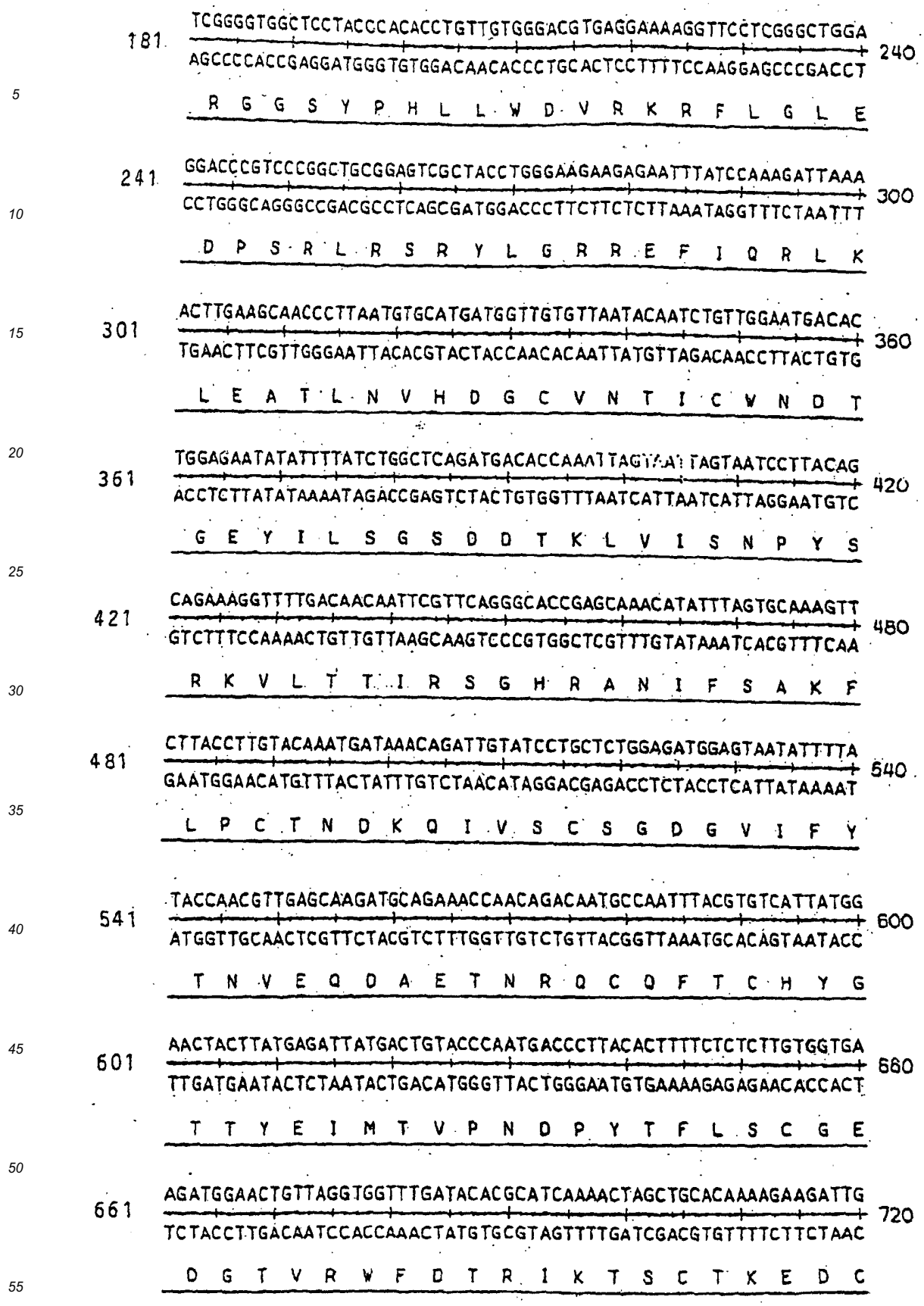
[0059] Analysis of the deduced amino acid sequence of Repro-EN-1.0 shows no significant sequence identity with any other protein.

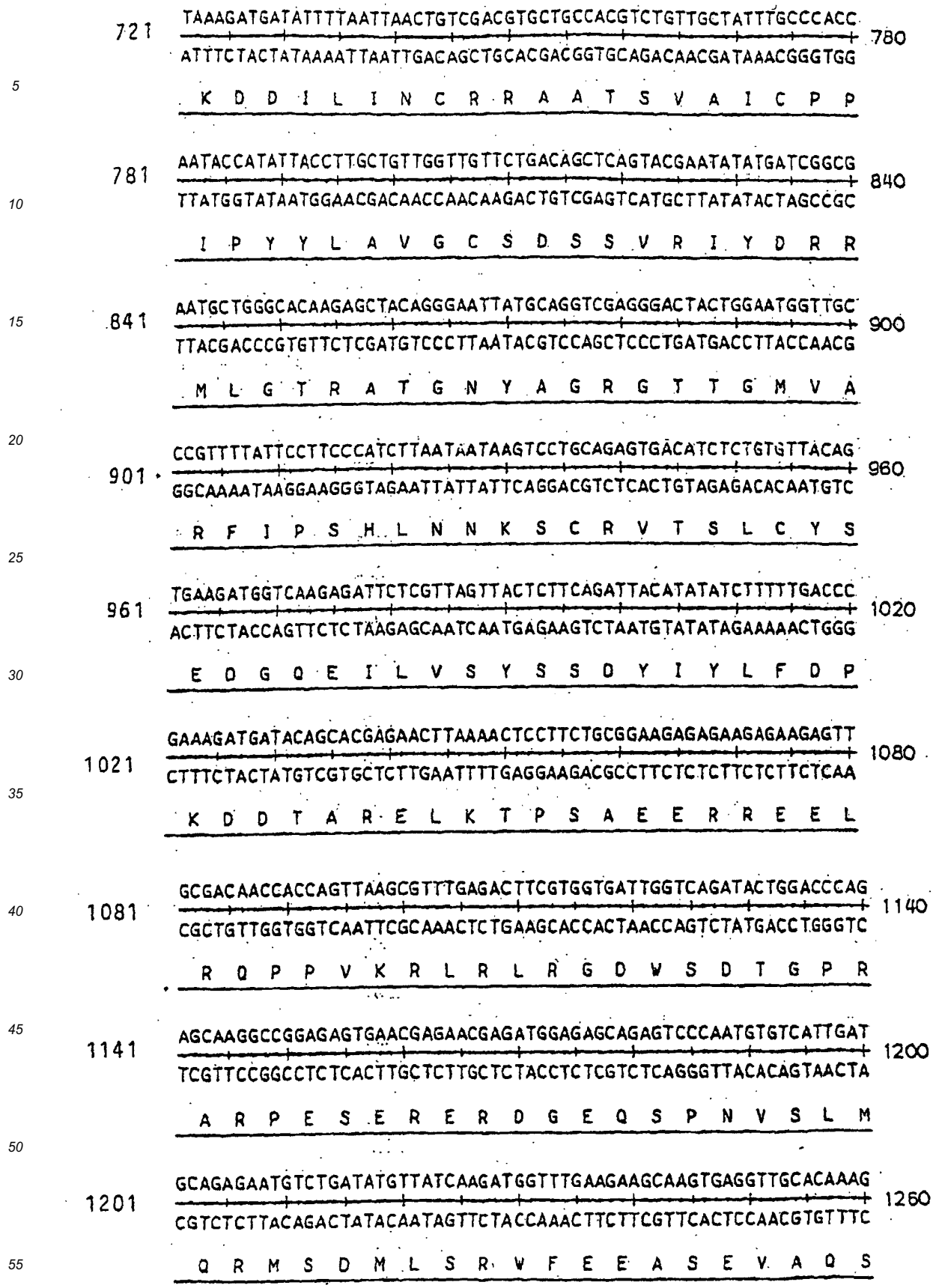
[0060] There is an alternately spliced variant that was isolated form a human heart cDNA library. This variant is called

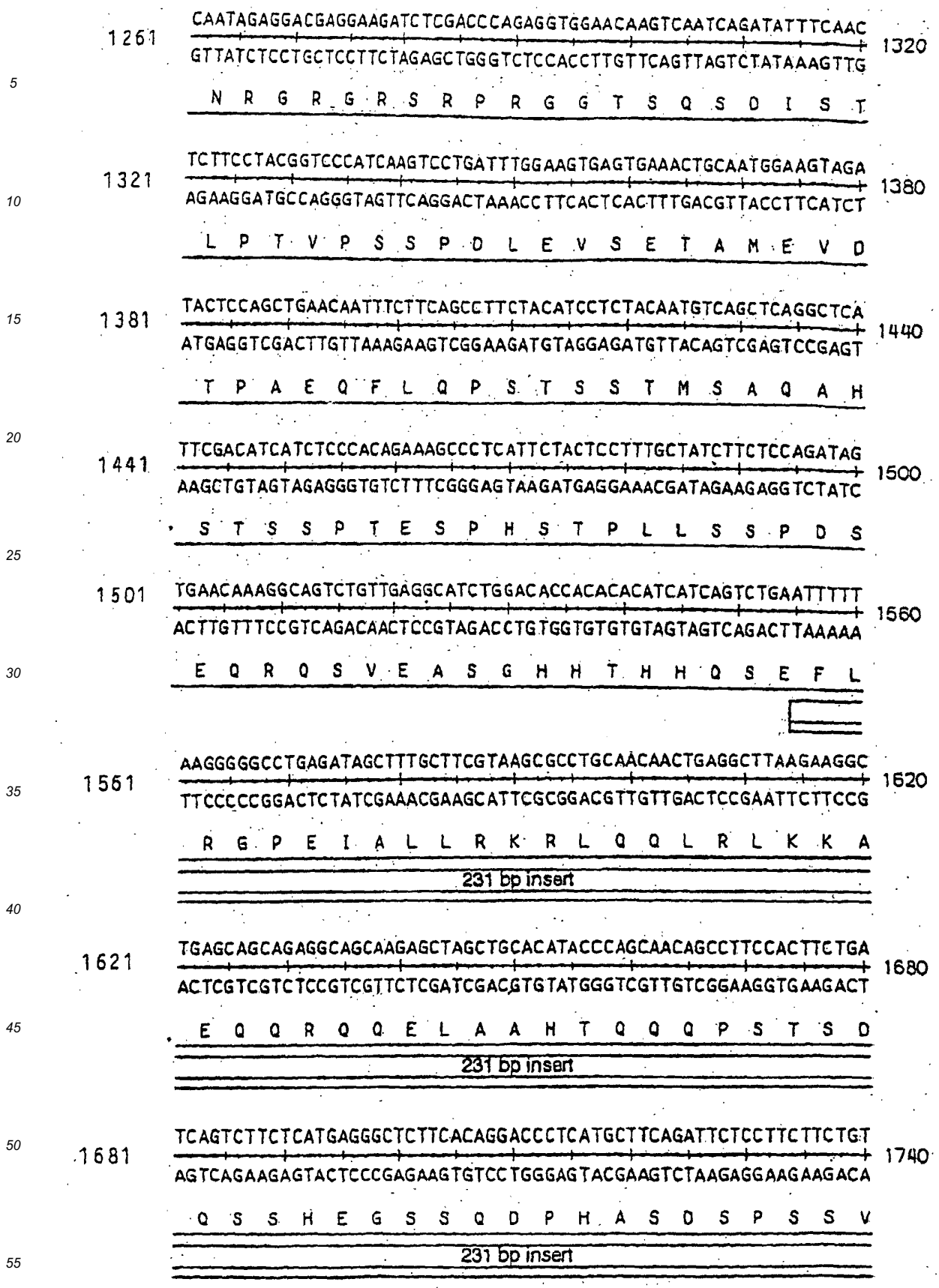
IB1 and is useful in the same ways as the nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of Repro-EN-1.0. The IB1 sequence was isolated from a heart cDNA library by screening with a nucleic acid probe obtained from the Repro-EN-1.0 sequence. The IB1 sequence is different from the Repro-EN-1.0 sequence in that it contains an additional 231 bp exon inserted into the cDNA sequence at position 1555. Therefore the IB1 sequence has similar properties, but is slightly larger.

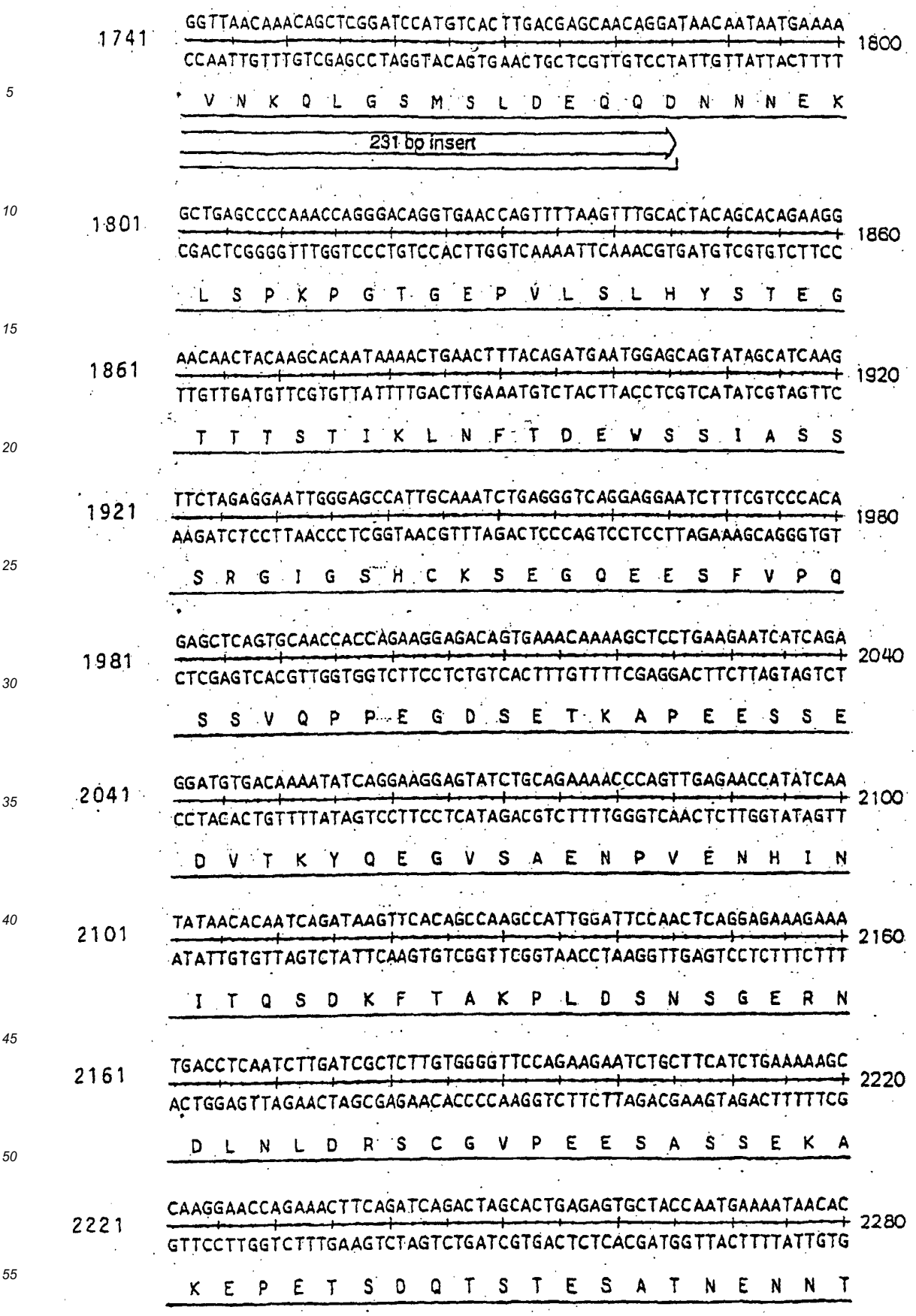
[0061] The nucleotide sequence (SEQ ID NO:3) and the deduced amino acid sequence (SEQ ID NO:4) of IB1 follow:

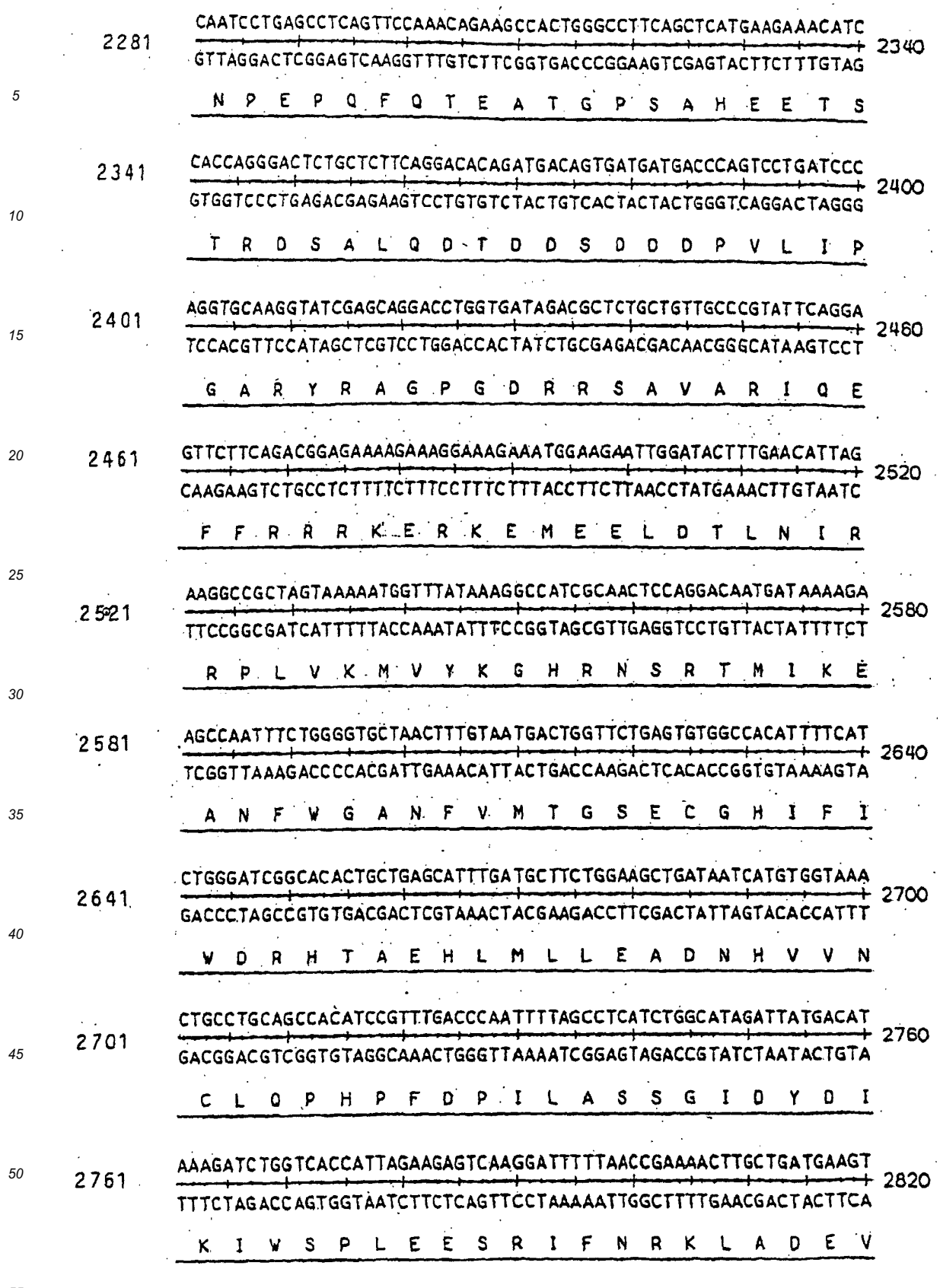


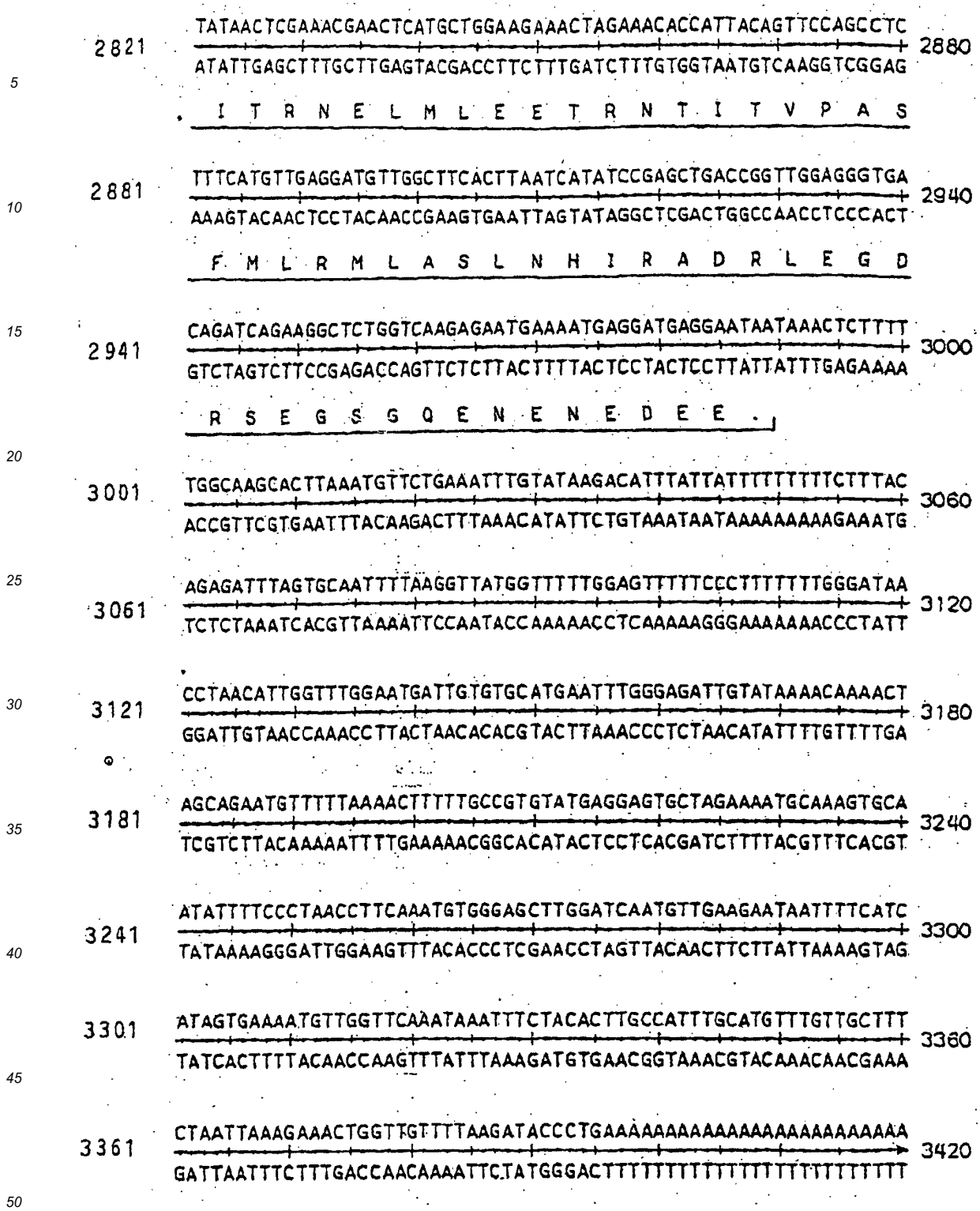

181
TCGGGGTGGCTCCTACCCACACCTGTTGTGGGACGTGAGGAAAAGGTTCCTCGGGCTGGA
AGCCCCACCGAGGATGGGTGTGGACAACACCCTGCACTCCTTTTCCAAGGAGCCCGACCT
240
R G G S Y P H L L W D V R K R F L G L E
241
GGACCCGTCCCGGCTGCGGAGTCGCTACCTGGGAAGAAGAGAATTTATCCAAAGATTAAA
CCTGGGCAGGGCCGACGCCTCAGCGATGGACCCTTCTTCTCTTAAATAGGTTTCTAATTT
300
D P S R L R S R Y L G R R E F I Q R L K
301
ACTTGAAGCAACCCTTAATGTGCATGATGGTTGTGTTAATACAATCTGTTGGAATGACAC
TGAACTTCGTTGGGAATTACACGTACTACCAACACAATTATGTTAGACAACCTTACTGTG
360
L E A T L N V H D G C V N T I C W N D T
361
TGGAGAATATATTTTATCTGGCTCAGATGACACCAAATTAGTAATAGTAATCCTTACAG
ACCTCTTATATAAAATAGACCGAGTCTACTGTGGTTTAATCATTAATCATTAGGAATGTC
420
G E Y I L S G S D D T K L V I S N P Y S
421
CAGAAAGGTTTTGACAACAATTCGTTCAGGGCACCGAGCAAACATATTTAGTGCAAAGTT
GTCTTTCCAAAACTGTTGTTAAGCAAGTCCCGTGGCTCGTTTGTATAAATCACGTTTCAA
480
R K V L T T I R S G H R A N I F S A K F
481
CTTACCTTGTACAAATGATAAACAGATTGTATCCTGCTCTGGAGATGGAGTAATATTTTA
GAATGGAACATGTTTACTATTTGTCTAACATAGGACGAGACCTCTACCTCATTATAAAAT
540
L P C T N D K Q I V S C S G D G V I F Y
541
TACCAACGTTGAGCAAGATGCAGAAACCAACAGACAATGCCAATTTACGTGTCATTATGG
ATGGTTGCAACTCGTTCTACGTCTTTGGTTGTCTGTTACGGTTAAATGCACAGTAATACC
600
T N V E Q D A E T N R Q C Q F T C H Y G
601
AACTACTTATGAGATTATGACTGTACCCAATGACCCTTACACTTTTCTCTCTTGTGGTGA
TTGATGAATACTCTAATACTGACATGGGTTACTGGGAATGTGAAAAGAGAGAACACCACT
660
T T Y E I M T V P N D P Y T F L S C G E
661
AGATGGAACTGTTAGGTGGTTTGATACACGCATCAAAACTAGCTGCACAAAAGAAGATTG
TCTACCTTGACAATCCACCAAACTATGTGCGTAGTTTTGATCGACGTGTTTTCTTCTAAC
720
D G T V R W F D T R I K T S C T K E D C


721
TAAAGATGATATTTTAATTAACTGTCGACGTGCTGCCACGTCTGTTGCTATTTGCCCACC
780
ATTTCTACTATAAAATTAATTGACAGCTGCACGACGGTGCAGACAACGATAAACGGGTGG
K D D I L I N C R R A A T S V A I C P P
781
AATACCATATTACCTTGCTGTTGGTTGTTCTGACAGCTCAGTACGAATATATGATCGGCG
840
TTATGGTATAATGGAACGACAACCAACAAGACTGTCGAGTCATGCTTATATACTAGCCGC
I P Y Y L A V G C S D S S V R I Y D R R
841
AATGCTGGGCACAAGAGCTACAGGGAATTATGCAGGTCGAGGGACTACTGGAATGGTTGC
900
TTACGACCCGTGTTCTCGATGTCCCTTAATACGTCCAGCTCCCTGATGACCTTACCAACG
M L G T R A T G N Y A G R G T T G M V A
901
CCGTTTTATTCCTTCCCATCTTAATAATAAGTCCTGCAGAGTGACATCTCTGTGTTACAG
960
GGCAAAATAAGGAAGGGTAGAATTATTATTCAGGACGTCTCACTGTAGAGACACAATGTC
R F I P S H L N N K S C R V T S L C Y S
961
TGAAGATGGTCAAGAGATTCTCGTTAGTTACTCTTCAGATTACATATATCTTTTTGACCC
1020
ACTTCTACCAGTTCTCTAAGAGCAATCAATGAGAAGTCTAATGTATATAGAAAAACTGGG
E D G Q E I L V S Y S S D Y I Y L F D P
1021
GAAAGATGATACAGCACGAGAACTTAAAACTCCTTCTGCGGAAGAGAGAAGAGAAGAGTT
1080
CTTTCTACTATGTCGTGCTCTTGAATTTTGAGGAAGACGCCTTCTCTCTTCTCTTCTCAA
K D D T A R E L K T P S A E E R R E E L
1081
GCGACAACCACCAGTTAAGCGTTTGAGACTTCGTGGTGATTGGTCAGATACTGGACCCAG
1140
CGCTGTTGGTGGTCAATTCGCAAACTCTGAAGCACCACTAACCAGTCTATGACCTGGGTC
R Q P P V K R L R L R G D W S D T G P R
1141
AGCAAGGCCGGAGAGTGAACGAGAACGAGATGGAGAGCAGAGTCCCAATGTGTCATTGAT
1200
TCGTTCCGGCCTCTCACTTGCTCTTGCTCTACCTCTCGTCTCAGGGTTACACAGTAACTA
A R P E S E R E R D G E Q S P N V S L M
1201
GCAGAGAATGTCTGATATGTTATCAAGATGGTTTGAAGAAGCAAGTGAGGTTGCACAAAG
1260
CGTCTCTTACAGACTATACAATAGTTCTACCAAACTTCTTCGTTCACTCCAACGTGTTTC
Q R M S D M L S R W F E E A S E V A Q S


1261
CAATAGAGGACGAGGAAGATCTCGACCCAGAGGTGGAACAAGTCAATCAGATATTTCAAC
GTTATCTCCTGCTCCTTCTAGAGCTGGGTCTCCACCTTGTTCAGTTAGTCTATAAAGTTG
1320
N R G R G R S R P R G G T S Q S D I S T
1321
TCTTCCTACGGTCCCATCAAGTCCTGATTTGGAAGTGAGTGAAACTGCAATGGAAGTAGA
AGAAGGATGCCAGGGTAGTTCAGGACTAAACCTTCACTCACTTTGACGTTACCTTCATCT
1380
L P T V P S S P D L E V S E T A M E V D
1381
TACTCCAGCTGAACAATTTCTTCAGCCTTCTACATCCTCTACAATGTCAGCTCAGGCTCA
ATGAGGTCGACTTGTTAAAGAAGTCGGAAGATGTAGGAGATGTTACAGTCGAGTCCGAGT
1440
T P A E Q F L Q P S T S S T M S A Q A H
1441
TTCGACATCATCTCCCACAGAAAGCCCTCATTCTACTCCTTTGCTATCTTCTCCAGATAG
AAGCTGTAGTAGAGGGTGTCTTTCGGGAGTAAGATGAGGAAACGATAGAAGAGGTCTATC
1500
S T S S P T E S P H S T P L L S S P D S
1501
TGAACAAAGGCAGTCTGTTGAGGCATCTGGACACCACACACATCATCAGTCTGAATTTTT
ACTTGTTTCCGTCAGACAACTCCGTAGACCTGTGGTGTGTGTAGTAGTCAGACTTAAAAA
1560
E Q R Q S V E A S G H H T H H Q S E F L
1561
AAGGGGGCCTGAGATAGCTTTGCTTCGTAAGCGCCTGCAACAACTGAGGCTTAAGAAGGC
TTCCCCCGGACTCTATCGAAACGAAGCATTCGCGGACGTTGTTGACTCCGAATTCTTCCG
1620
R G P E I A L L R K R L Q Q L R L K K A
231 bp insert
1621
TGAGCAGCAGAGGCAGCAAGAGCTAGCTGCACATACCCAGCAACAGCCTTCCACTTCTGA
ACTCGTCGTCTCCGTCGTTCTCGATCGACGTGTATGGGTCGTTGTCGGAAGGTGAAGACT
1680
E Q Q R Q Q E L A A H T Q Q Q P S T S D
231 bp insert
1681
TCAGTCTTCTCATGAGGGCTCTTCACAGGACCCTCATGCTTCAGATTCTCCTTCTTCTGT
AGTCAGAAGAGTACTCCCGAGAAGTGTCCTGGGAGTACGAAGTCTAAGAGGAAGAAGACA
1740
Q S S H E G S S Q D P H A S D S P S S V
231 bp insert


1741
GGTTAACAAACAGCTCGGATCCATGTCACTTGACGAGCAACAGGATAACAATAATGAAAA
1800
CCAATTGTTTGTCGAGCCTAGGTACAGTGAACTGCTCGTTGTCCTATTGTTATTACTTTT
V N K Q L G S M S L D E Q Q D N N N E K
231 bp insert
1801
GCTGAGCCCCAAACCAGGGACAGGTGAACCAGTTTTAAGTTTGCACTACAGCACAGAAGG
1860
CGACTCGGGGTTTGGTCCCTGTCCACTTGGTCAAAATTCAAACGTGATGTCGTGTCTTCC
L S P K P G T G E P V L S L H Y S T E G
1861
AACAACTACAAGCACAATAAAACTGAACTTTACAGATGAATGGAGCAGTATAGCATCAAG
1920
TTGTTGATGTTCGTGTTATTTTGACTTGAAATGTCTACTTACCTCGTCATATCGTAGTTC
T T T S T I K L N F T D E W S S I A S S
1921
TTCTAGAGGAATTGGGAGCCATTGCAAATCTGAGGGTCAGGAGGAATCTTTCGTCCCACA
1980
AAGATCTCCTTAACCCTCGGTAACGTTTAGACTCCCAGTCCTCCTTAGAAAGCAGGGTGT
S R G I G S H C K S E G Q E E S F V P Q
1981
GAGCTCAGTGCAACCACCAGAAGGAGACAGTGAAACAAAAGCTCCTGAAGAATCATCAGA
2040
CTCGAGTCACGTTGGTGGTCTTCCTCTGTCACTTTGTTTTCGAGGACTTCTTAGTAGTCT
S S V Q P P E G D S E T K A P E E S S E
2041
GGATGTGACAAAATATCAGGAAGGAGTATCTGCAGAAAACCCAGTTGAGAACCATATCAA
2100
CCTACACTGTTTTATAGTCCTTCCTCATAGACGTCTTTTGGGTCAACTCTTGGTATAGTT
D V T K Y Q E G V S A E N P V E N H I N
2101
TATAACACAATCAGATAAGTTCACAGCCAAGCCATTGGATTCCAACTCAGGAGAAAGAAA
2160
ATATTGTGTTAGTCTATTCAAGTGTCGGTTCGGTAACCTAAGGTTGAGTCCTCTTTCTTT
I T Q S D K F T A K P L D S N S G E R N
2161
TGACCTCAATCTTGATCGCTCTTGTGGGGTTCCAGAAGAATCTGCTTCATCTGAAAAAGC
2220
ACTGGAGTTAGAACTAGCGAGAACACCCCAAGGTCTTCTTAGACGAAGTAGACTTTTTCG
D L N L D R S C G V P E E S A S S E K A
2221
CAAGGAACCAGAAACTTCAGATCAGACTAGCACTGAGAGTGCTACCAATGAAAATAACAC
2280
GTTCCTTGGTCTTTGAAGTCTAGTCTGATCGTGACTCTCACGATGGTTACTTTTATTGTG
K E P E T S D Q T S T E S A T N E N N T


2281
CAATCCTGAGCCTCAGTTCCAAACAGAAGCCACTGGGCCTTCAGCTCATGAAGAAACATC
GTTAGGACTCGGAGTCAAGGTTTGTCTTCGGTGACCCGGAAGTCGAGTACTTCTTTGTAG
2340
N P E P Q F Q T E A T G P S A H E E T S
2341
CACCAGGGACTCTGCTCTTCAGGACACAGATGACAGTGATGATGACCCAGTCCTGATCCC
GTGGTCCCTGAGACGAGAAGTCCTGTGTCTACTGTCACTACTACTGGGTCAGGACTAGGG
2400
T R D S A L Q D T D D S D D D P V L I P
2401
AGGTGCAAGGTATCGAGCAGGACCTGGTGATAGACGCTCTGCTGTTGCCCGTATTCAGGA
TCCACGTTCCATAGCTCGTCCTGGACCACTATCTGCGAGACGACAACGGGCATAAGTCCT
2460
G A R Y R A G P G D R R S A V A R I Q E
2461
GTTCTTCAGACGGAGAAAAGAAAGGAAAGAAATGGAAGAATTGGATACTTTGAACATTAG
CAAGAAGTCTGCCTCTTTTCTTTCCTTTCTTTACCTTCTTAACCTATGAAACTTGTAATC
2520
F F R R R K E R K E M E E L D T L N I R
2521
AAGGCCGCTAGTAAAAATGGTTTATAAAGGCCATCGCAACTCCAGGACAATGATAAAAGA
TTCCGGCGATCATTTTTACCAAATATTTCCGGTAGCGTTGAGGTCCTGTTACTATTTTCT
2580
R P L V K M V Y K G H R N S R T M I K E
2581
AGCCAATTTCTGGGGTGCTAACTTTGTAATGACTGGTTCTGAGTGTGGCCACATTTTCAT
TCGGTTAAAGACCCCACGATTGAAACATTACTGACCAAGACTCACACCGGTGTAAAAGTA
2640
A N F W G A N F V M T G S E C G H I F I
2641
CTGGGATCGGCACACTGCTGAGCATTTGATGCTTCTGGAAGCTGATAATCATGTGGTAAA
GACCCTAGCCGTGTGACGACTCGTAAACTACGAAGACCTTCGACTATTAGTACACCATTT
2700
W D R H T A E H L M L L E A D N H V V N
2701
CTGCCTGCAGCCACATCCGTTTGACCCAATTTTAGCCTCATCTGGCATAGATTATGACAT
GACGGACGTCGGTGTAGGCAAACTGGGTTAAAATCGGAGTAGACCGTATCTAATACTGTA
2760
C L Q P H P F D P I L A S S G I D Y D I
2761
AAAGATCTGGTCACCATTAGAAGAGTCAAGGATTTTTAACCGAAAACTTGCTGATGAAGT
TTTCTAGACCAGTGGTAATCTTCTCAGTTCCTAAAAATTGGCTTTTGAACGACTACTTCA
2820
K I W S P L E E S R I F N R K L A D E V



IB1 Alternately Spliced Variant from Heart Library:

[0062]

IB1 cDNA is 3,395 base pairs
IB1 open reading frame is 2,811 nucleotides

IB1 open reading frame maps to nucleotides 176 to 2986
Deduced amino acid sequence of IB1 has 937 amino acids
Calculated moecular mass of IB1 protein is 104,969 Daltons (105kD)
Calculated pI of IB1 protein is 5.17

[0063] Analysis of the amino acid sequence identified several amino acid motifs that will be apparent to those skilled in the art including a myb 1 DNA binding domain, a WD 40 site, an RGD cell-attachment sequence, an N-myristolylation site and several phosphorylation and glycosylation sites. Analysis also shows that the protein is largely hydrophilic. This implies that most of the amino acid sequence is exposed to the immune system and can be recognized as epitopes.

[0064] Analysis of expressed sequence tags (ESTs) from a public database (Genbank) identified many overlapping ESTs that, together, covered most of the Repro-EN-1.0 cDNA sequence.

III. REPRO-EN-1.0 and IB1 NUCLEIC ACIDS

[0065] This invention provides recombinant polynucleotides comprising a nucleotide sequence encoding Repro-EN-1.0 and IB1 proteins as described herein. The polynucleotides are useful for expressing the mRNA or polypeptides they encode and in the preparation of probes or primers, among other things.

[0066] In one embodiment, the recombinant polynucleotide molecule comprises a nucleotide sequence encoding polypeptide Repro-EN-1.0 of SEQ ID NO:2 or IB1 of SEQ ID NO:4.

[0067] The nucleotide sequence can be identical to a sequence from Repro-EN-1.0 cDNA or its complement or IB1 cDNA or its complement, or can include degenerate codons. In one embodiment of a nucleotide sequence encoding full-length Repro-EN-1.0 or IB1, the sequence is identical to the coding sequence of Repro-EN-1.0 of SEQ ID NO:1 or IB1 of SEQ ID NO:3.

[0068] In another embodiment, the polynucleotide encodes a fusion protein between Repro-EN-1.0 or IB1 polypeptide or Repro-EN-1.0 or IB1 analog amino acid sequences and a second amino acid sequence. The second amino acid sequence can be, for example, a detectable label such as a fluorescent protein, enzyme marker of protein from a two-hybrid system.

[0069] The polynucleotides of the present invention are cloned or amplified by *in vitro* methods, such as the polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (3SR) and the Qβ replicase amplification system (QB). For example, a polynucleotide encoding the protein can be isolated by polymerase chain reaction of cDNA from a human endometrial carcinoma cell line using primers based on the DNA sequence of Repro-EN-1.0 of SEQ ID NO:1. One pair of primers useful for amplifying Repro-EN-1.0 DNA, including allelic variants, is:

Upstream sense: 5'-caggacacagatgacagtgat-3' (SEQ ID NO:5)

Downstream antisense: 5'-agagccttctgatctgtcac-3' (SEQ ID NO:6).

[0070] A wide variety of cloning and *in vitro* amplification methodologies are well-known to persons of skill. PCR methods are described in, for example, U.S. Pat. No. 4,683,195; Mullis et al. (1987) *Cold Spring Harbor Symp. Quant. Biol*. 51:263; and Erlich, ed., *PCR Technology,* (Stockton Press, NY, 1989). One useful format is real time PCR. See, e.g., Luch et al. (1997) *J. Molec. Endocrinol*. 18:77-85 and Arold et al., (1997) *Proc. Nat'l Acad. Sci.,* USA 94:2438-43. Polynucleotides also can be isolated by screening genomic or cDNA libraries with probes selected from the sequences of SEQ ID NO:1 under stringent hybridization conditions.

[0071] Mutant versions of the proteins can be made by site-specific mutagenesis of other polynucleotides encoding the proteins, or by random mutagenesis caused by increasing the error rate of PCR of the original polynucleotide with 0.1 mM MnCl2 and unbalanced nucleotide concentrations.

[0072] This invention also provides expression vectors, e.g., recombinant polynucleotide molecules comprising expression control sequences operatively linked to a nucleotide sequence encoding the target polypeptide. Expression vectors can be adapted for function in prokaryotes or eukaryotes by inclusion of appropriate promoters, replication sequences, markers, etc. for transcription and translation of mRNA. The construction of expression vectors and the expression of genes in transfected cells involves the use of molecular cloning techniques also well known in the art. Sambrook et al., *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, (1989) and *Current Protocols in Molecular Biology,* F.M. Ausubel et al., eds., (Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc.) Useful promoters for such purposes include a metallothionein promoter, a constitutive adenovirus major late promoter, a dexamethasone-inducible MMTV promoter, a

SV40 promoter, a MRP polIII promoter, a constitutive MPSV promoter, a tetracycline-inducible CMV promoter (such as the human immediate-early CMV promoter), and a constitutive CMV promoter.

[0073] . Methods for transfecting genes into mammalian cells and obtaining their expression for *in vitro* use or for gene therapy, are well known to the art. See, e.g., *Methods in Enzymology,* vol. 185, Academic Press, Inc., San Diego, CA (D.V. Goeddel, ed.) (1990) or M. Krieger, *Gene Transfer and Expression -- A Laboratory Manual,* Stockton Press, New York, NY, (1990).

[0074] Expression vectors useful in this invention depend on their intended use. Such expression vectors must, of course, contain expression and replication signals compatible with the host cell. Expression vectors useful for expressing the protein of this invention include viral vectors such as alpha viruses, retroviruses, adenoviruses and adeno-associated viruses, plasmid vectors, cosmids, liposomes and the like. Viral and plasmid vectors are preferred for transfecting mammalian cells. The expression vector pcDNA1 (Invitrogen, San Diego, CA), in which the expression control sequence comprises the CMV promoter, provides good rates of transfection and expression. Adeno-associated viral vectors are useful in the gene therapy methods of this invention.

[0075] The construct can also contain a tag to simplify isolation of the protein. For example, a polyhistidine tag of, e.g., six histidine residues, can be incorporated at the amino terminal end of the protein. The polyhistidine tag allows convenient isolation of the protein in a single step by nickel-chelate chromatography.

[0076] In another embodiment, endogenous genes are transcribed by operatively linking them to expression control sequences supplied endogenously that recombine with genomic DNA. In one method, one provides the cell with a recombinant polynucleotide containing a targeting sequence, which permits homologous recombination into the genome upstream of the transcriptional start site of target gene; the expression control sequences; an exon of the target gene; and an unpaired splice-donor site which pairs with a splice acceptor in the target gene. Such methods are discussed in Treco et al., WO 94/12650; Treco et al., WO 95/31560 and Treco et al., WO 96/29411.

[0077] . The invention also provides recombinant cells comprising an expression vector for expression of the nucleotide sequences encoding a polypeptide of this invention. Host cells can be selected for high levels of expression in order to purify the protein. Mammalian cells are preferred for this purpose, but prokaryotic cells, such as *E. coli,* also are useful. The cell can be, e.g., a recombinant cell in culture or a cell *in vivo*.

IV. POLYNUCLEOTIDE PROBES AND PRIMERS

[0078] This invention provides polynucleotide probes and primers that specifically hybridize to a sub-sequence of Repro-EN-1.0 cDNA or its complement or IB1 cDNA or its complement, under stringent hybridization conditions. The probes and primers of this invention are polynucleotides of at least 7 nucleotides, at least 10 nucleotides, at least 15 nucleotides, at least 20 nucleotides or at least 25 nucleotides. In one embodiment, the sequence of the polynucleotide is a contiguous sequence from SEQ ID NO: 1 or its complement. Any suitable region of the Repro-EN-1.0 or IB1 gene may be chosen as a target for polynucleotide hybridization. Nucleotide substitutions, deletions, and additions may be incorporated into the polynucleotides as long as the characteristic ability to specifically hybridize to the target sequence or its complement is retained. Nucleotide sequence variation may result from sequence polymorphisms of various alleles, minor sequencing errors, and the like.

[0079] The probes and primers of the invention are useful as probes in hybridization assays, such as Southern and northern blots, for identifying polynucleotides having a nucleotide sequence encoding a Repro-EN-1.0 or IB1 polypeptide, and as primers for amplification procedures. The probes and primers of the invention are also useful in detecting the presence, absence or amount of Repro-EN-1.0 or IB1 in tissue biopsies and histological sections where the detection method is carried out *in situ,* typically after amplification of Repro-EN-1.0 or IB1 sequences using a primer set.

[0080] The probes and primers of this invention also are useful for identifying allelic forms of Repro-EN-1.0 and animal cognate genes or IB1 and animal cognate genes. Probes and primers can be used to screen human or animal genomic DNA or cDNA libraries under, e.g., stringent conditions. DNA molecules that specifically hybridize to the probe are then further examined to determine whether they are Repro-EN-1.0 allelic variants or animal cognates or IB1 allelic variants or animal cognates.

[0081] The probes also are useful in oligonucleotide arrays. Such arrays are used in hybridization assays to check the identity of bases in a target polynucleotide. In essence, when a target hybridizes perfectly to a probe on the array, the target contains the nucleotide sequence of the probe. When the target hybridizes less well, or does not hybridize at all, then the target and probe differ in sequence by one or more nucleotide. By proper selection of probes, one can check bases on a target molecule. See, e.g., Chee et al., WO 95/11995. The use the Repro-EN-1.0 or IB1 sequence in genomics is described further below.

[0082] In one embodiment, the polynucleotide is directly or indirectly detectable through a detectable moiety. A detectable moiety bound to either an oligonucleotide primer or a probe is subsequently used to detect hybridization of an oligonucleotide primer to the RNA component. Detection of labeled material bound to a Repro-EN-1.0 or IB1 polynucleotide in a sample provides a means of determining a diagnostic or prognostic value.

[0083] Although primers and probes can differ in sequence and length, the primary differentiating factor is one of function: primers serve as an initiation point for DNA synthesis of a target polynucleotide, as in RT and PCR reactions, while probes are typically used for hybridization to and detection of a target polynucleotide. Typical lengths of primers or probes can range from 7-50 nucleotides, preferably from 10-40 nucleotides, and most preferably from 15-35 nucleotides. A primer or probe can also be labeled with a detectable moiety for detection of hybridization of the primer or probe to the target polynucleotide.

[0084] In general, those of skill in the art recognize that the polynucleotides used in the invention include both DNA and RNA molecules and naturally occurring modifications thereof, as well as synthetic, non-naturally occurring analogs of the same, and heteropolymers, of deoxyribonucleotides, ribonucleotides, and/or analogs of either. The particular composition of a polynucleotide or polynucleotide analog will depend upon the purpose for which the material will be used and the environment in which the material will be placed. Modified or synthetic, non-naturally occurring nucleotides have been designed to serve a variety of purposes and to remain stable in a variety of environments, such as those in which nucleases are present.

[0085] Oligonucleotides preferably are synthesized, e.g., on an Applied BioSystems or other commercially available oligonucleotide synthesizer according to specifications provided by the manufacturer. Oligonucleotides may be prepared using any suitable method, such as the phosphotriester and phosphodiester methods, or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidates are used as starting materials and may be synthesized as described by Beaucage et al., *Tetrahedron Letters* 22: 1859 (1981), and U.S. Patent No. 4,458,066.

[0086] Polynucleotides, e.g., probes, also can be recombinantly produced through the use of plasmids or other vectors.

[0087] In one aspect this invention provides a probe that specifically hybridizes to the 5' untranslated region of Repro-EN-1.0 or IB1, the coding region of Repro-EN-1.0 or IB1, or a region of Repro-EN-1.0 or IB1 encoding an epitope of the Repro-EN-1.0 or IB1 polypeptide.

[0088] In another aspect, this invention provides a primer pair which amplifies a nucleotide sequence encoding a polypeptide epitope of Repro-EN-1.0 or IB1 recognized by an antibody from an individual diagnosed with endometriosis. A primer pair that amplifies a particular nucleotide (given in the 5' to 3' orientation) includes a 5' primer and a 3' primer. The 3' primer hybridizes to the 3' end of the nucleotide sequence or downstream from it. The 5' primer hybridizes to the 3' end of the complement of the nucleotide sequence or downstream from it. In this way, the primers can amplify a polynucleotide that comprises the nucleotide sequence. One nucleotide sequence encoding a polypeptide epitope of Repro-EN-1.0 has been identified within the about 2.2 kb from 3' end of the coding sequences of Repro-EN-1.0 (SEQ ID NO:1).

V. REPRO-EN-1.0 AND IB1 POLYPEPIDES

[0089] This invention also provides purified, recombinant Repro-EN-1.0 and IB1 polypeptide. The invention provides recombinant Repro-EN-1.0 polypeptides which include the polypeptide whose amino acid sequence is presented in SEQ ID NO:2, as well as allelic variants of it.

[0090] Repro-EN-1.0 polypeptide refers to native Repro-EN-1.0, the polypeptide whose amino acid sequence is the amino acid sequence of SEQ ID NO:2, and to allelic variants of it. Polynucleotide molecules that encode allelic variants of Repro-EN-1.0 are isolatable from endometrial cancer cell cDNA or genomic DNA and typically hybridize under stringent conditions to the nucleotide sequence encoding Repro-EN-1.0 (SEQ ID NO:1). They can be obtained by amplification using, e.g., PCR primers taken from the sequence of Repro-EN-1.0 described herein.

[0091] Repro-EN-1.0 polypeptides are useful as immunogens to elicit the production of anti-Repro-EN-1.0 antibodies, as affinity capture molecules to isolate such antibodies from a mixture, and as controls in diagnostic methods aimed at detecting Repro-EN-1.0 in a sample.

[0092] The cDNA encoding Repro-EN-1.0 of SEQ ID NO:1 was discovered by screening an expression library of cDNA from an endometrial carcinoma cell line with serum pooled from subjects diagnosed with endometriosis. Therefore, polypeptides comprising an epitope of Repro-EN-1.0 can be identified by screening with such serum. Preferably, the test serum is a serum pooled from several subjects positively diagnosed with endometriosis. At least one epitope of Repro-EN-1.0 exists in a fragment of 567 amino acids from the carboxy-terminus of the molecule (amino acids 293-860 of SEQ ID NO:2). At least one epitope of IB1 exists in a fragment of 644 amino acids from the carboxy terminus of the molecule (amino acids 293-937 of SEQ ID NO: 4) to detect the presence of antibodies against Repro-EN-1.0 or IB1 in patient serum samples. This test is useful in diagnosis because the presence of such antibodies is a diagnostic marker.

[0093] Repro-EN-1.0 or IB1 are most easily produced recombinantly, as described herein. Recombinant Repro-EN-1.0 or IB1 can be purified by affinity purification In one method, recombinant Repro-EN-1.0 or IB1 comprise a polyhistidine tag. The protein is purified on a nickel-chelate affinity matrix. In another method, Repro-EN-1.0 or IB1 is purified using an affinity matrix carrying anti-Repro-EN-1.0 or IB1 antibodies.

VI. ANTIBODIES AND HYBRIDOMAS

[0094] In one aspect this invention provides a composition comprising an antibody that specifically binds Repro-EN-1.0 or IB1 polypeptides. Antibodies preferably have affinity of at least 10^6 M^{-1}, 10^7 M^{-1}, 10^8 M^{-1}, or 10^9 M^{-1}. This invention contemplates both polyclonal and monoclonal antibody compositions.

[0095] Immunotoxins against Repro-EN-1.0- or IB1-expressing cells are also described herein. Immunotoxins are antibodies and the like as described herein coupled to a compound, e.g., a toxin, that is toxic to a target cell. Toxins can include, for example, radioactive isotopes, ricin, cisplatin, antisense molecules, *Diphtheria* toxin, *Pseudomonas* exotoxin A or *Bacillus anthraces* protective antigen. Immunotoxins bind to cancer cells that express Repro-EN-1.0 or IB1 and kill them. They are useful in the therapeutic methods of this invention. The antibodies of the invention have many uses. For example, such antibodies are useful for detecting Repro-EN-1.0 or IB1 polypeptides in immunoassays. The antibodies also can be used to screen expression libraries for particular expression products such as mammalian Repro-EN-1.0 or IB1. These are useful for detecting or diagnosing various pathological conditions related to the presence of the respective antigens. Usually the antibodies in such a procedure are labeled with a moiety allowing easy detection of presence of antigen by antibody binding. Antibodies raised against Repro-EN-1.0 or IB1 can also be used to raise anti-idiotypic antibodies.

A. Production of Antibodies

[0096] A number of immunogens are used to produce antibodies that specifically bind Repro-EN-1.0 or IB1 polypeptides. Full-length Repro-EN-1.0 or IB1 is a suitable immunogen. Typically, the immunogen of interest is a peptide of at least about 3 amino acids, more typically the peptide is 5 amino acids in length, preferably, the fragment is 10 amino acids in length and more preferably the fragment is 15 amino acids in length or greater, Thepeptides can be coupled to a carrier protein (e.g., as a fusion protein), or are recombinantly expressed in an immunization vector. Antigenic determinants on peptides to which antibodies bind are typically 3 to 10 amino acids in length. Naturally occurring polypeptides are also used either in pure or impure form.

[0097] Recombinant polypeptides are expressed in eukaryotic or prokaryotic cells and purified using standard techniques. The polypeptide, or a synthetic version thereof, is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies can be generated for subsequent use in immunoassays to measure the presence and quantity of the polypeptide.

[0098] Methods for producing polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified polypeptide, a polypeptide coupled to an appropriate carrier *(e.g.,* GST, keyhole limpet hemanocyanin, etc.), or a polypeptide incorporated into an immunization vector such as a recombinant vaccinia virus (see, U. S. Patent No. 4,722,848) is mixed with an adjuvant and animals are immunized with the mixture. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the polypeptide of interest. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the polypeptide is performed where desired. See, e.g., Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY.

[0099] Antibodies, including binding fragments and single chain recombinant versions thereof, against predetermined fragments of Repro-EN-1.0 or IB1 proteins are raised by immunizing animals, e.g., with conjugates of the fragments with carrier proteins as described above.

[0100] Monoclonal antibodies are prepared from cells secreting the desired antibody. These antibodies are screened for binding to normal or modified polypeptides, or screened for agonistic or antagonistic activity, e.g., activity mediated through Repro-EN-1.0 or IB1. In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description of techniques for preparing such monoclonal antibodies are found in, e.g., Stites et al. (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane, *Supra;* Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY; and Kohler and Milstein (1975) *Nature* 256: 495-497.

[0101] Other suitable techniques involve selection of libraries of recombinant antibodies in phage or similar vectors. See, Huse et al. (1989) *Science* 246: 1275-1281; and Ward, et al. (1989) *Nature* 341: 544-546.

[0102] Also, recombinant immunoglobulins may be produced. See, Cabilly, U.S. Patent No. 4,816,567; and Queen et al. (1989) *Proc. Nat'l Acad. Sci.* USA 86: 10029-10033.

[0103] Frequently, the polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Thus, an antibody used for detecting an analyte can be directly labeled with a detectable moiety, or may be indirectly labeled by, for example, binding to the antibody a secondary antibody that is, itself directly or indirectly labeled.

[0104] The antibodies of this invention are also used for affinity chromatography in isolating Repro-EN-1.0 or IB1 proteins. Columns are prepared, e.g., with the antibodies linked to a solid support, e.g., particles, such as agarose, Sephadex, or the like, where a cell lysate is passed through the column, washed, and treated with increasing concentrations of a mild denaturant, whereby purified Repro-EN-1.0 or IB1 polypeptides are released

[0105] An alternative approach is the generation of humanized immunoglobulins by linking the CDR regions of non-human antibodies to human constant regions by recombinant DNA techniques. See Queen et al., United States patent 5,585,089.

[0106] .A further approach for isolating DNA sequences which encode a human monoclonal antibody or a binding fragment thereof is by screening a DNA library from human B cells according to the general protocol outlined by Huse et al., *Science* 246:1275-1281 (1989) and then cloning and amplifying the sequences which encode the antibody (or binding fragment) of the desired specificity. The protocol described by Huse is rendered more efficient in combination with phage display technology. See, e.g., Dower et al., WO 91/17271 and McCafferty et al., WO 92/01047. Phage display technology can also be used to mutagenize CDR regions of antibodies previously shown to have affinity for Repro-EN-1.0 or IB1 protein receptors or their ligands. Antibodies having improved binding affinity are selected.

[0107] Fragments of antibodies against Repro-EN-1.0 or IB1 protein or protein analogs are also described herein. Typically, these fragments exhibit specific binding to the Repro-EN-1.0 protein receptor similar to that of a complete immunoglobulin. Antibody fragments include separate heavy chains, light chains Fab, Fab' $F(ab')_2$ and Fv. Fragments are produced by recombinant DNA techniques, or by enzymic or chemical separation of intact immunoglobulins

VII. METHODS FOR DETECTING REPRO-EN-1.0 AND IB1 POLYPEPTIDES

[0108] Repro-EN-1.0 or IB1 polypeptides can be identified by any methods known in the art. In one embodiment, the methods involve detecting the polypeptide with a ligand that specifically recognizes the polypeptide (e.g., an immunoassay). The antibodies of the invention are particularly useful for specific detection of Repro-EN-1.0 or IB1 polypeptides. . A variety of antibody-based detection methods are known in the art. These include, for example, radioimmunoassay, sandwich immunoassays (including ELISA), immunofluorescent assays, western blot, affinity chromatography (affinity ligand bound to a solid phase), and *in situ* detection with labeled antibodies. Another method for detecting Repro-EN-1.0 or IB1 polypeptides involves identifying the polypeptide according to its mass through, for example, gel electrophoresis, mass spectrometry or HPLC. Subject samples can be taken from any number of appropriate sources, such as saliva, peritoneal fluid, blood or a blood product (e.g., serum), urine, tissue biopsy (e.g., lymph node tissue), etc.

a. Immunoassays

[0109] The methods for detection of Repro-EN-1.0 or IB1 polypeptides may employ one or more anti-Repro-EN-1.0 or IB1 antibody reagents (i.e., immunoassays). A number of well established immunological binding assay formats suitable for the practice of the invention are known (see, e.g., U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). See, e.g., METHODS IN CELL BIOLOGY VOLUME 37: ANTIBODIES IN CELL BIOLOGY, Asai, ed. Academic Press, Inc. New York (1993); BASIC AND CLINICAL IMMUNOLOGY 7th Edition, Stites & Terr, eds. (1991); Harlow and Lane, *supra* [e.g., Chapter 14], and Ausubel et al., *supra,* [e.g., Chapter 11]. Typically, immunological binding assays (or immunoassays) utilize a "capture agent" to specifically bind to and, often, immobilize the analyte to a solid phase. In one embodiment embodiment, the capture agent is a moiety that specifically binds to a Repro-EN-1.0 or IB1 polypeptide or subsequence, such as an anti-Repro-EN-1.0 or anti-IB1 antibody.

[0110] Usually the Repro-EN-1.0 or IB1 polypeptide being assayed is detected directly or indirectly using a detectable label. The particular label or detectable group used in the assay is usually not a critical aspect of the invention, so long as it does not significantly interfere with the specific binding of the antibody or antibodies used in the assay. The label may be covalently attached to the capture agent (e.g., an anti-Repro-EN-1.0 or or anti-IB1 antibody), or may be attached to a third moiety, such as another antibody, that specifically binds to the Repro-EN-1.0 polypeptide.

[0111] The methods employ competitive and noncompetitive immunoassays for detecting Repro-EN-1.0 or IB1 polypeptides. Non-competitive immunoassays are assays in which the amount of captured analyte (in this case Repro-EN-1.0 or IB1) is directly measured. One such assay is a two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the Repro-EN-1.0 or IB1 protein. See, e.g., Maddox et al., 1983, *J. Exp. Med.* 158:1211 for background information. In one preferred "sandwich" assay, the capture agent (e.g., an anti-Repro-EN-1.0 or anti-IB1 antibody) is bound directly to a solid substrate where it is immobilized. These immobilized antibodies then capture any Repro-EN-1.0 or IB1 protein present in the test sample. The Repro-EN-1.0 or IB1 polypeptide thus immobilized can then be labeled, i.e., by binding to a second anti-Repro-EN-1.0 or IB1 antibody bearing a label. Alternatively, the second anti-Repro-EN-1.0 or IB1 antibody may lack a label, but be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second antibody

alternatively can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

[0112] In competitive assays, the amount of Repro-EN-1.0 or IB1 protein present in the sample is measured indirectly by measuring the amount of an added (exogenous) Repro-EN-1.0 or IB1 displaced (or competed away) from a capture agent (e.g., anti-Repro-EN-1.0 or anti-IB1 antibody) by the Repro-EN-1.0 or IB1 protein present in the sample.

[0113] A hapten inhibition assay is another example of a competitive assay. In this assay Repro-EN-1.0 or IB1 protein is immobilized on a solid substrate. A known amount of anti-Repro-EN-1.0 or anti-IB1 antibody is added to the sample, and the sample is then contacted with the immobilized Repro-EN-1.0 or IB1 protein. In this case, the amount of anti-Repro-EN-1.0 or anti-IB1 antibody bound to the immobilized Repro-EN-1.0 or IB1 protein is inversely proportional to the amount of Repro-EN-1.0 protein present in the sample. The amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. In this aspect, detection may be direct, where the antibody is labeled, or indirect where the label is bound to a molecule that specifically binds to the antibody as described above.

b. Other Antibody-based Assay Formats

[0114] The methods detect and quantify the presence of Repro-EN-1.0 or IB1 polypeptide in the sample by using an immunoblot (Western blot) format. Another immunoassay is the so-called "lateral flow chromatography." In a non-competitive version of lateral flow chromatography, a sample moves across a substrate by, e.g., capillary action, and encounters a mobile labeled antibody that binds the analyte forming a conjugate. The conjugate then moves across the substrate and encounters an immobilized second antibody that binds the analyte. Thus, immobilized analyte is detected by detecting the labeled antibody. In a competitive version of lateral flow chromatography a labeled version of the analyte moves across the carrier and competes with unlabeled analyte for binding with the immobilized antibody. The greater the amount of the analyte in the sample, the less the binding by labeled analyte and, therefore, the weaker the signal. See, e.g., May et aL, U.S. patent 5,622,871 and Rosenstein, U.S. patent 5,591,645.

c. Solid Phases: Substrates, Solid Supports, Membranes, Filters

[0115] As noted *supra,* depending upon the assay, various components, including the antigen, target antibody, or anti-Repro-EN-1.0 or anti-IB1 antibody, may be bound to a solid surface or support (i.e., a substrate, membrane, or filter paper). Many methods for immobilizing biomolecules to a variety of solid surfaces are known in the art. For instance, the solid surface may be a membrane *(e.g.,* nitrocellulose), a microtiter dish *(e.g.,* PVC, polypropylene, or polystyrene), a test tube (glass or plastic), a dipstick *(e.g.* glass, PVC, polypropylene, polystyrene, latex, and the like), a microcentrifuge tube, or a glass or plastic bead. The desired component may be covalently bound or noncovalently attached through nonspecific bonding.

[0116] A wide variety of organic and inorganic polymers, both natural and synthetic may be employed as the material for the solid surface. Illustrative polymers include polyethylene, polypropylene, poly(4-methylbutene), polystyrene, polymethacrylate, poly(ethylene terephthalate), rayon, nylon, poly(vinyl butyrate), polyvinyldene difluoride (PVDF), silicones, polyformaldehyde, cellulose, cellulose acetate, nitrocellulose, and the like. Other materials which may be employed, include paper, glasses, ceramics, metals, metalloids, semiconductive materials, cements or the like. In addition, substances that form gels, such as proteins (*e.g.*, gelatins), lipopolysaccharides, silicates, agarose and polyacrylamides can be used. Polymers which form several aqueous phases, such as dextrans, polyalkylene glycols or surfactants, such as phospholipids, long chain (12-24 carbon atoms) alkyl ammonium salts and the like are also suitable. Where the splid surface is porous, various pore sizes may be employed depending upon the nature of the system.

d. Mass Spectrometry

[0117] The mass of a molecule firquently can be used as an identifier of the molecule. Therefore, methods of mass spectrometry can be used to identify a protein analyte. Mass spectrometers can measure mass by determining the time required for an ionized analyte to travel down a flight tube and to be detected by an ion detector.

[0118] One method of mass spectrometry for proteins is matrix-assisted laser desorption/ionization mass spectrometry ("MALDI"). In MALDI the analyte is mixed with an energy absorbing matrix material that absorbs energy of the wavelength of a laser and placed on the surface of a probe. Upon striking the matrix with the laser, the analyte is desorbed from the probe surface, ionized, and detected by the ion detector. See, for example, Hillenkamp et al., United States patent 5,118,937

[0119] . Other methods of mass spectrometry for proteins are described in Hutchens and Yip, United States patent 5,719,060. In one such method referred to as Surfaces Enhanced for Affinity Capture ("SEAC") a solid phase affinity reagent that binds the analyte specifically or non-specifically, such as an antibody or a metal ion, is used to separate

the analyte from other materials in a sample. Then the captured analyte is desorbed from the solid phase by, e.g., laser energy, ionized, and detected by the detector.

e. Assay Combinations

[0120] The diagnostic and prognostic assays described herein can be carried out in various combinations and can also be carried out in conjunction with other diagnostic or prognostic tests. For example, when the present methods are used to diagnose endometriosis, the presence of a Repro-EN-1.0 or IB1 polypeptide can be used to determine the stage of the disease. Tests that may provide additional information include microscopic analysis of biopsy samples, detection of antigens (e.g., cell-surface markers) associated with endometriosis (e.g., using histocytochemistry, FACS, or the like).

VIII. DIAGNOSTIC, MONITORING AND PROGNOSTIC METHODS

A. Methods Of Diagnosing Endometriosis

[0121] We have detected circulating antibodies against Repro-EN-1.0 or IB1 in the blood of women diagnosed with endometriosis. This supports the idea that endometriosis has an autoimmune component. Further, Repro-EN-1.0 or IB1 and auto-antibodies against Repro-EN-1.0 or IB1 represent two targets in the diagnosis of endometriosis.

[0122] Repro-EN-1.0 or IB1 that is shed into the peritoneal fluid of women with endometriosis is useful in methods of diagnosing endometriosis. These methods include detecting Repro-EN-1.0 or IB1 in a biological sample of a subject. Suitable samples include, without limitation, saliva, blood or a blood product (e.g., serum), urine, menstrual fluid, vaginal secretion and, in particular, peritoneal fluid. Repro-EN-1.0 or IB1 can be detected by any of the methods described herein. Any detection of Repro-EN-1.0 or IB1 above a normal range is a positive sign in the diagnosis of endometriosis.

[0123] In another aspect, this invention provides methods for diagnosing endometriosis in a subject by detecting in a sample from the subject a diagnostic amount of an antibody that specifically binds to Repro-EN-1.0 or IB1 polypeptide. Suitable patient samples include, without limitation, saliva, blood or a blood product (e.g., serum), peritoneal fluid, urine, menstrual fluid, vaginal secretion. The antibodies can be detected by any of the methods for detecting proteins described herein. However, sandwich type assays are particularly useful. In one version, all antibodies are captured onto a solid phase, for example using protein A, and antibodies specific for Repro-EN-1.0 or IB1 are detected using a directly or indirectly labeled Repro-EN-1.0 or IB1 or polypeptide fragment of it having an epitope of Repro-EN-1.0 or IB1. In another version of the assay, Repro-EN-1.0 or IB1 or an antigenic fragment of it can be used as the capture molecule and captured antibodies can be detected.

[0124] While the detection of antibodies, in general, against Repro-EN-1.0 or IB1 is a positive sign of endometriosis, IgE an IgG_4 class antibodies are particularly specific and sensitive for the diagnosis of endometriosis. Therefore, in one embodiment, the diagnostic method involves specifically detecting IgE or IgG_4 antibodies that specifically recognize Repro-EN-1.0 or IB1. Anti-human IgE antibodies and anti-human IgG_4 antibodies can be easily bought or made.

IX. METHODS FOR SCREENING FOR COMPOUNDS THAT REGULATE EXPRESSION OF REPRO-EN-1.0 OR IB1

[0125] Compounds that regulate the expression of Repro-EN-1.0 and IB1 are candidates as therapeutic agents. This invention provides methods for determining whether a compound regulates (e.g., activates or inhibits) expression of Repro-EN-1.0 or IB1.

[0126] Methods for determining whether a compound regulates Repro-EN-1.0 or IB1 expression involve administering to a cell or a test animal having an expressible Repro-EN-1.0 or IB1 gene with the compound, and determining whether expression Repro-EN-1.0 or IB1 is altered. In one embodiment, the methods involve administering the compound to a culture comprising the cell or to a test animal that has cells expressing Repro-EN-1.0 or IB1, measuring the amount of the Repro-EN-1.0 or IB1 polynucleotide or polypeptide in a sample from the culture or the animal, and determining whether the measured amount is different than the amount in a sample from the culture or from the animal under control conditions (e.g., to which no compound has been administered). Statistically significant ($p<0.05$) differences between the amount measured from the test sample and from the control sample are recorded and indicate that the compound alters the amount of Repro-EN-1.0 or IB1 produced by the cell.

[0127] The compound to be tested can be selected from a number of sources. For example, combinatorial libraries of molecules are available for screening experiments. Using such libraries, thousands of molecules can be screened for regulatory activity. In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve

as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

[0128] Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.*, U.S. Patent 5,010,175, Furka (1991) *Int. J. Pept. Prot. Res.,* 37: 487-493, Houghton *et al*. (1991) *Nature*, 354: 84-88). Peptide synthesis is by no means the only approach envisioned and intended for use with the present invention. Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (PCT.Publication No WO 91/19735, 26 Dec. 1991), encoded peptides (PCT Publication WO 93/20242, 14 Oct. 1993), random bio-oligomers (PCT Publication WO 92/00091, 9 Jan. 1992), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al*., (1993) *Proc. Nat. Acad. Sci.* USA 90: 6909-6913), vinylogous polypeptides (Hagihara *et al*. (1992) *J. Amer. Chem. Soc.* 114: 6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann *et al*., (1992) *J. Amer. Chem. Soc.* 114: 9217-9218), analogous organic syntheses of small compound libraries (Chen *et al*. (1994) *J. Amer. Chem. Soc*. 116: 2661), oligocarbamates (Cho, *et al*., (1993) *Science* 261:1303), and/or peptidyl phosphonates (Campbell *et al*., (1994) *J. Org. Chem*. 59: 658). *See, generally*, Gordon *et al*., (1994) *J. Med. Chem*. 37:1385, nucleic acid libraries, peptide nucleic acid libraries (*see, e.g.*, U. S. Patent 5,539,083) antibody libraries (*see, e.g*., Vaughn *et al*. (1996) *Nature Biotechnology*, 14(3): 309-314), and PCT/US96/10287), carbohydrate libraries (*see*, *e.g*., Liang *et al*. (1996) *Science*, 274: 1520-1522, and U.S. Patent 5,593,853), and small organic molecule libraries (*see, e.g*., benzodiazepines, Baum (1993) C&EN, Jan 18, page 33, isoprenoids U.S. Patent 5,569,588, thiazolidinones and metathiazanones U.S. Patent 5,549,974, pyrrolidines U.S. Patents 5,525,735 and 5,519,134, morpholino compounds U.S. Patent 5,506,337, benzodiazepines 5,288,514, and the like).

[0129] Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

X GENOMICS

[0130] The identification of cognate or polymorphic forms of the Repro-EN-1.0 or IB1 gene and the tracking of those polymorphisms in individuals and families is important in genetic screening. Accordingly, this invention provides methods useful in detecting polymorphic forms of the Repro-EN-1.0 or IB1 gene. The methods involve comparing the identity of a nucleotide or amino acid at a selected position within the sequence of a test Repro-EN-1.0 or IB1 gene with the nucleotide or amino acid at the corresponding position from the sequence of native Repro-EN-1.0 (SEQ ID NO:1) or IB1 (SEQ ID NO:3). The comparison can be carried out by any methods known in the art, including direct sequence comparison by nucleotide sequencing, sequence comparison or determination by hybridization or ident fication of RFLPs.

[0131] In one embodiment, the method involves nucleotide or amino acid sequencing of the entire test polynucleotide orpolypeptide, or a subsequence from it, and comparing that sequence with the sequence of native Repro-EN-1.0 or IB1. In another embodiment, the method involves identifying restriction fragments produced upon restriction enzyme digestion of the test polynucleotide and comparing those fragments with fragments produced by restriction enzyme digestion of native Repro-EN-1.0 or IB1 gene. Restriction fragments from the native gene can be identified by analysis of the sequence to identify restriction sites. Another embodiment involves the use of oligonucleotide arrays. (See, e. g., Fodor et al., United States patent 5,445,934.) The method involves providing an oligonucleotide array comprising a set of oligonucleotide probes that define sequences selected from the native Repro-EN-1.0 or IB1 sequence, generating hybridization data by performing a hybridization reaction between the target polynucleotide molecules and the probes in the set and detecting hybridization between the target molecules and each of the probes in the set and processing the hybridization data to determine nucleotide positions at which the identity of the target molecule differs from that of native Repro-EN-1.0 or IB1. The comparison can be done manually, but is more conveniently done by a programmable, digital computer.

EXAMPLES

I. Construction of a human endometrial carcinoma cell line (RL95-2) cDNA expression library

A. Material and Methods

[0132] Poly A+ RNA isolated from RL95-2, was used as a template for first strand cDNA synthesis. The poly A+ RNA was analyzed by denaturing gel electrophoresis and ranged in size from 0.2 to 10 Kb (10). An oligo dT primer containing an internal, protected *Xhol* site was annealed in the presence of a nucleotide mixture containing 5-methyl dCTP, and extended with MMTV reverse transcriptase.

[0133] Second strand synthesis of the RL95-2 cDNA/RNA hybrid was completed by the addition of RNase H, DNA polymerase 1, and dNTPs to the first strand synthesis reaction. *Pfu* DNA polymerase was used to blunt-end the double stranded RL95-2 cDNA followed by the ligation of *EcoRI* adapters. The cDNA was kinased and digested with *XhoI* and *EcoRI* before size fractionation on Sephacryl S-500 columns, The size fractionated cDNA was recovered and the quantified on ethidium bromide containing plates against a set of serially-diluted DNA standards. The cDNA contained in the first two column fractionations was directionally ligated, in the sense orientation, to *XhoI*/*EcoRI*-digested uniZAP phage vector arms. Initially, approximately 25 ng (per fraction) of the cDNA was ligated and packaged into bacteriophage particles. Subsequently, the approximately 100 ng remaining cDNA in fractions 1 and 2 was packaged into bacteriophage particles using several reactions of the lambda phage packaging extract (Stratagene).

[0134] After packaging, the primary human RL95-2 library was titered by infection of the XL1 Blue host strain. The ratio of recombinant:nonrecombinant phage was determined by plating infected XL1 Blue in the presence of IPTG and XGal. The number of blue (non-recombinant) or white (recombinant) plaques were quantified using a Manostat colony counter. Ninety-eight and one-half (98.5) percent of the phage in the human RL95-2 cDNA library were recombinant and the primary phage library base consisted of 2.2×10^6 independent clones. The human RL95-2 cDNA library was amplified once and re-titered as above. The titer of the amplified library was 3×10^9 pfu/ml.

[0135] The average size of the cDNA inserts was determined by PCR amplification. Twenty well-isolated phage plaques were cored and the cDNA inserts were amplified using T7/T3 specific oligonucleotides which hybridize to sites flanking the cDNA insertion site contained within the lambda phage vector. The PCR products were analyzed by agarose gel electrophoresis and visualized by ethidium bromide staining. Ninety--five percent of the cored plaques contained inserts varying in size from 0.5 to 2.5 Kb with an average size of 1.5 Kb.

II. Library Screening

[0136] To identify endometrial autoantigens that could be used to develop an endometriosis immunoassay, pooled sera from patients (n=17) with laparoscopy confirmed endometriosis was used as a probe to screen the RL95-2 cDNA expression library using the following methods.

A. Preadsorption of serum. Antibodies that react with expression library host strains were immunoadsorbed from patient serum by using BNN97 and Y1090 *E. coli* lysate-conjugated sepharose beads (5' → 3') following the manufacturer's protocols. Briefly, 2 ml of host stain lysate-conjugated sepharose beads were washed twice with sterile Tris buffered saline (TBS). The beads were resuspended in 4 mls of serum diluted 1/2 in TBS. Following a 16 hour incubation at 4°C, the sepharose beads were collected by centrifugation at 1,000 x g for 2 min. The supernatant was removed and the beads were washed with 4 mls of sterile TBS. After centrifugation at 1,000 x g for 2 min, the supernatants were collected, pooled and used to screen the RL95-2-specific cDNA expression library.

B. Screening the RL95-2 cDNA library. Approximately 10^6 infectious phage particles were incubated with XL-1 blue host cells and plated at density of 50,000 phage per 150 mm dish using standard protocols (Stratagene). After incubating for 5 hours at 42°C, the phage plaques were overlaid with nitrocellulose membranes (Protran; Schleicher & Schuell) that had been soaked in a 10 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) solution. Following a 4 hour incubation at 37°C, the membranes were removed and washed three times for 15 min in TBS with 0.05% $Tween_{20}$ (TTBS). The membranes were incubated with blocking solution (1% Bovine serum albumin [fraction V] in TBS) for 1 hour at room temperature prior to a 1 hour incubation with preadsorbed patient serum diluted 1/10 (final dilution of 1/40) in blocking solution. The membranes were washed three time for 15 min with TTBS prior to the addition of alkaline phospbatase-conjugated goat anti-human Ig(G,A,M) (Pierce) diluted 1/25,000 in blocking solution. After a 1 hour incubation at room temperature the membranes were washed three times with TTBS as described above and once with TBS. The membranes were incubated with enzyme substrate (Western Blue; Promega) for approximately 30 min and the enzymatic reaction was terminated by briefly incubating the membranes with stop solution (Tris-HCl pH 2.9; 1 mM EDTA). Several immunoreactive phage plaques were selected and transferred to 500 µl of SM buffer (100 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris-HCl pH 7.5, 0.01% gelatin) containing 20 µl of chloroform. The selected phage were eluted from the agar and plated at a density of approximately 1,000 phage per 100 mm dish and screen as described above. To insure that the selected phage plaque, named Repro-EN-1.0, represented a single clone the screening process was repeated a third time as described above.

C. Excision of Repro-EN-1.0 phagemid. Plasmid containing the cDNA insert for Repro-EN-1.0 was excised from the phage clone using the manufacturer's protocols (Stratagene). The size of the Repro-EN-1.0 insert was determined by releasing the cDNA fragment from the rescued pBluescript/Repro-EN-1.0 plasmid with the restriction enzymes *EcoRI* and *XhoI.* The released insert was size fractionated by agarose-gel electrophoresis and the ap-

parent length of the insert was determined by comparing its migration position with a DNA standard (1 kb ladder; Gibco BRL). The insert migrated at approximately 2.0 Kb.

D. Identification of the IB1 clone. An alternately spliced variant of Repro-EN-1.0. A commercial human heart cDNA library (Clontech) was screened with a radiolabeled probe mapping within the amino terminus of the Repro-EN-1.0 coding sequence (nt 203 to 897). One of the two clones isolated contained a cDNA insert of 3.4 Kb which possessed an extra 231 base pair insert within the Repro-EN-1.0 coding sequence.

III. Characterization

A. Sequence analysis

[0137] The nucleotide sequence of Repro-EN-1.0 was determined by using a modified protocol of the dideoxy chain termination method of Sanger et al. and USB Sequenase 2.0 (Barker, D.F. 1993, *Biotechniques*). The amino acid sequence was predicted using the Intelligenetic TRANSLATE program and sequence homologies were determined with BLAST data base search algorithms. The deduced amino acid sequence (in the expected frame for a fusion protein) was novel.

B. Tissue expression analysis

[0138] The Repro-EN-1.0 expression distribution was determined by Northern blot analysis using poly A+ RNA collected from human spleen, thymus, prostate, testis, ovary, small intestine, colon, and peripheral blood leukocytes (MTN human blot 11; Clontech) and human heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas (MTN human blot 1- Clontech) and the manufacturer' suggested protocols (Clontech). The immobilized poly A+ RNA samples were incubated with a random prime labeled probe that represented the Repro-EN-1.0 insert. The probe was generated by using and *EcoRI*/*XhoI* released fragment of pRepro-EN-1.0 as template for a random prime labeled reaction as described in the manufacturer's manual (Megaprime kit; Amersham). The Northern blot membranes were prehybridized with 6 mls of ExpressHyb solution (Clontech) followed by a 1 hour incubation at 68°C with approximately 0.5 ng of radiolabeled probe (approx. 2X105 cpm) in 5 mls of ExpressHyb solution. Unbound probe was removed by washing the membranes three times with 2 X SSC, 0.05% SDS at room temperature for 10 min for each wash followed by twice with 0.1 X SSC, 0.1% SDS at 50° C for 15 min for each wash. The apparent length of RNA species, a 3.4 Kb mRNA, and tissue distribution was determined by autoradiography.
Expression was detected primarily in skeletal muscle, heart and testis; and to a lesser extent in other tissues, but was not detected in lung or peripheral blood mononuclear cells (PBMC). Expression of Repro-EN-1.0 was up-regulated in breast and uterine carcinomas relative to their normal counterparts, was highly expressed in both normal fallopian tube and fallopian tube carcinoma, and was expressed at low levels in both normal ovary and ovarian carcinoma (Fig. 2). Expression of Repro-EN-1.0 in RL95-2 (endometrium carcinoma) cells is lower than in LNCaP (human prostate adenocarcinoma), PC-12 (rat cell line) and BT12 (human breast carcinoma cell line) cells and undetectable in a mouse hybridoma cell line (3E10; negative control) (Fig.3). In addition, expression in normal endometrium is undetectable.

C. Homologue analysis

[0139] To determine the level of nucleotide conservation of Repro-EN-1.0 in different species, a Southern blot analysis using *EcoRI* digested genomic DNA collected from human, monkey, rat, mouse, dog, cow, rabbit, chicken and yeast was performed as described in the manufactures manual (ZooBlot; Clontech). Briefly, the immobilized *EcoRI* digested genomic DNA samples were prehybridized with 6 mls of ExpressHyb (Clontech) and incubated for 1 hour at 68°C with 1.0 ng (approx. $4X10^5$ cpm in 5 mls) of a random prime labeled probe that represented the Repro-EN-1.0 insert. Unbound probe was removed by washing the membranes once with 2 X SSC, 0.05% SDS at room temperature for 30 min followed by one wash with 0.1 X SSC, 0.1% SDS at 50C for 30 min. Identification of homologues in different species was determine by autoradiography. The sequence is highly conserved between human and non-human primates (Monkey).

IV. Antibodies

[0140] Antibodies to peptides of the clone and/or to recombinant protein were generated in rabbits. This antisera was used to develop a Repro-EN-1.0 ELISA.

V. Recombinant Protein

[0141] The predicted ORF for Repro-EN-1 was subcloned into an expression vector, and the recombinant protein was expressed and purified by Ni--Chelate chromatography using standard methodologies.

VI. ELISA

[0142] The purified recombinant Repro-EN-1 protein was used as a target antigen in an ELISA (EndX™ ELISA) designed to detect antigen-specific autoantibodies in patient serum.

SEQUENCE LISTING

[0143]

<110> Diagnostic Products Corporation

<120> Polynucleotide Encoding Autoantigens Associates With Endometriosis

<130> N80980

<160> 20

<170> PatentIn Ver. 2.1

<210> 1
<211> 3189
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (176)..(2755)

<400> 1

```
cggccgggct tcaggggccc aggcgccgct gctgccaccg ccatctaacg ctgcgccctg 60

gaggcccggc gcgcggatgg tgccggtgcg gctcgggtgt tgaaacgggt gtcccctccc 120

cctcctcccc tcccccacgc ggtggtctcc cctcccaccc ggctcaggca gagcc atg  178
                                                            Met
                                                              1

tct cgg ggt ggc tcc tac cca cac ctg ttg tgg gac gtg agg aaa agg   226
Ser Arg Gly Gly Ser Tyr Pro His Leu Leu Trp Asp Val Arg Lys Arg
            5                  10                  15

ttc ctc ggg ctg gag gac ccg tcc cgg ctg cgg agt cgc tac ctg gga   274
Phe Leu Gly Leu Glu Asp Pro Ser Arg Leu Arg Ser Arg Tyr Leu Gly
```

```
                20                  25                  30

aga aga gaa ttt atc caa aga tta aaa ctt gaa gca acc ctt aat gtg   322
Arg Arg Glu Phe Ile Gln Arg Leu Lys Leu Glu Ala Thr Leu Asn Val
     35                  40                  45

cat gat ggt tgt gtt aat aca atc tgt tgg aat gac act gga gaa tat   370
His Asp Gly Cys Val Asn Thr Ile Cys Trp Asn Asp Thr Gly Glu Tyr
 50                  55                  60                  65

att tta tct ggc tca gat gac acc aaa tta gta att agt aat cct tac   418
Ile Leu Ser Gly Ser Asp Asp Thr Lys Leu Val Ile Ser Asn Pro Tyr
                 70                  75                  80

agc aga aag gtt ttg aca aca att cgt tca ggg cac cga gca aac ata   466
Ser Arg Lys Val Leu Thr Thr Ile Arg Ser Gly His Arg Ala Asn Ile
             85                  90                  95

ttt agt gca aag ttc tta cct tgt aca aat gat aaa cag att gta tcc   514
Phe Ser Ala Lys Phe Leu Pro Cys Thr Asn Asp Lys Gln Ile Val Ser
        100                 105                 110

tgc tct gga gat gga gta ata ttt tat acc aac gtt gag caa gat gca   562
Cys Ser Gly Asp Gly Val Ile Phe Tyr Thr Asn Val Glu Gln Asp Ala
    115                 120                 125

gaa acc aac aga caa tgc caa ttt acg tgt cat tat gga act act tat   610
Glu Thr Asn Arg Gln Cys Gln Phe Thr Cys His Tyr Gly Thr Thr Tyr
130                 135                 140                 145

gag att atg act gta ccc aat gac cct tac act ttt ctc tct tgt ggt   658
Glu Ile Met Thr Val Pro Asn Asp Pro Tyr Thr Phe Leu Ser Cys Gly
                150                 155                 160

gaa gat gga act gtt agg tgg ttt gat aca cgc atc aaa act agc tgc   706
Glu Asp Gly Thr Val Arg Trp Phe Asp Thr Arg Ile Lys Thr Ser Cys
            165                 170                 175

aca aaa gaa gat tgt aaa gat gat att tta att aac tgt cga cgt gct   754
Thr Lys Glu Asp Cys Lys Asp Asp Ile Leu Ile Asn Cys Arg Arg Ala
        180                 185                 190

gcc acg tct gtt gct att tgc cca cca ata cca tat tac ctt gct gtt   802
Ala Thr Ser Val Ala Ile Cys Pro Pro Ile Pro Tyr Tyr Leu Ala Val
    195                 200                 205

ggt tgt tct gac agc tca gta cga ata tat gat cgg cga atg ctg ggc   850
Gly Cys Ser Asp Ser Ser Val Arg Ile Tyr Asp Arg Arg Met Leu Gly
```

```
210                 215                 220                 225

aca aga gct aca ggg aat tat gca ggt cga ggg act act gga atg gtt   898
Thr Arg Ala Thr Gly Asn Tyr Ala Gly Arg Gly Thr Thr Gly Met Val
                230                 235                 240

gcc cgt ttt att cct tcc cat ctt aat aat aag tcc tgc aga gtg aca   946
Ala Arg Phe Ile Pro Ser His Leu Asn Asn Lys Ser Cys Arg Val Thr
            245                 250                 255

tct ctg tgt tac agt gaa gat ggt caa gag att ctc gtt agt tac tct   994
Ser Leu Cys Tyr Ser Glu Asp Gly Gln Glu Ile Leu Val Ser Tyr Ser
        260                 265                 270

tca gat tac ata tat ctt ttt gac ccg aaa gat gat aca gca cga gaa   1042
Ser Asp Tyr Ile Tyr Leu Phe Asp Pro Lys Asp Asp Thr Ala Arg Glu
    275                 280                 285

ctt aaa act cct tct gcg gaa gag aga aga gaa gag ttg cga caa cca   1090
Leu Lys Thr Pro Ser Ala Glu Glu Arg Arg Glu Glu Leu Arg Gln Pro
290                 295                 300                 305

cca gtt aag cgt ttg aga ctt cgt ggt gat tgg tca gat act gga ccc   1138
Pro Val Lys Arg Leu Arg Leu Arg Gly Asp Trp Ser Asp Thr Gly Pro
                310                 315                 320

aga gca agg ccg gag agt gaa cga gaa cga gat gga gag cag agt ccc   1186
Arg Ala Arg Pro Glu Ser Glu Arg Glu Arg Asp Gly Glu Gln Ser Pro
            325                 330                 335

aat gtg tca ttg atg cag aga atg tct gat atg tta tca aga tgg ttt   1234
Asn Val Ser Leu Met Gln Arg Met Ser Asp Met Leu Ser Arg Trp Phe
        340                 345                 350

gaa gaa gca agt gag gtt gca caa agc aat aga gga cga gga aga tct   1282
Glu Glu Ala Ser Glu Val Ala Gln Ser Asn Arg Gly Arg Gly Arg Ser
    355                 360                 365

cga ccc aga ggt gga aca agt caa tca gat att tca act ctt cct acg   1330
Arg Pro Arg Gly Gly Thr Ser Gln Ser Asp Ile Ser Thr Leu Pro Thr
370                 375                 380                 385

gtc cca tca agt cct gat ttg gaa gtg agt gaa act gca atg gaa gta   1378
Val Pro Ser Ser Pro Asp Leu Glu Val Ser Glu Thr Ala Met Glu Val
                390                 395                 400

gat act cca gct gaa caa ttt ctt cag cct tct aca tcc tct aca atg   1426
Asp Thr Pro Ala Glu Gln Phe Leu Gln Pro Ser Thr Ser Ser Thr Met
```

```
                405                 410                 415

tca gct cag gct cat tcg aca tca tct ccc aca gaa agc cct cat tct   1474
Ser Ala Gln Ala His Ser Thr Ser Ser Pro Thr Glu Ser Pro His Ser
        420                 425                 430

act cct ttg cta tct tct cca gat agt gaa caa agg cag tct gtt gag   1522
Thr Pro Leu Leu Ser Ser Pro Asp Ser Glu Gln Arg Gln Ser Val Glu
    435                 440                 445

gca tct gga cac cac aca cat cat cag tct gat aac aat aat gaa aag   1570
Ala Ser Gly His His Thr His His Gln Ser Asp Asn Asn Asn Glu Lys
450                 455                 460                 465

ctg agc ccc aaa cca ggg aca ggt gaa cca gtt tta agt ttg cac tac   1618
Leu Ser Pro Lys Pro Gly Thr Gly Glu Pro Val Leu Ser Leu His Tyr
                470                 475                 480

agc aca gaa gga aca act aca agc aca ata aaa ctg aac ttt aca gat   1666
Ser Thr Glu Gly Thr Thr Thr Ser Thr Ile Lys Leu Asn Phe Thr Asp
            485                 490                 495

gaa tgg agc agt ata gca tca agt tct aga gga att ggg agc cat tgc   1714
Glu Trp Ser Ser Ile Ala Ser Ser Ser Arg Gly Ile Gly Ser His Cys
        500                 505                 510

aaa tct gag ggt cag gag gaa tct ttc gtc cca cag agc tca gtg caa   1762
Lys Ser Glu Gly Gln Glu Glu Ser Phe Val Pro Gln Ser Ser Val Gln
    515                 520                 525

cca cca gaa gga gac agt gaa aca aaa gct cct gaa gaa tca tca gag   1810
Pro Pro Glu Gly Asp Ser Glu Thr Lys Ala Pro Glu Glu Ser Ser Glu
530                 535                 540                 545

gat gtg aca aaa tat cag gaa gga gta tct gca gaa aac cca gtt gag   1858
Asp Val Thr Lys Tyr Gln Glu Gly Val Ser Ala Glu Asn Pro Val Glu
                550                 555                 560

aac cat atc aat ata aca caa tca gat aag ttc aca gcc aag cca ttg   1906
Asn His Ile Asn Ile Thr Gln Ser Asp Lys Phe Thr Ala Lys Pro Leu
            565                 570                 575

gat tcc aac tca gga gaa aga aat gac ctc aat ctt gat cgc tct tgt   1954
Asp Ser Asn Ser Gly Glu Arg Asn Asp Leu Asn Leu Asp Arg Ser Cys
        580                 585                 590

ggg gtt cca gaa gaa tct gct tca tct gaa aaa gcc aag gaa cca gaa   2002
Gly Val Pro Glu Glu Ser Ala Ser Ser Glu Lys Ala Lys Glu Pro Glu
```

595 600 605

act tca gat cag act agc act gag agt gct acc aat gaa aat aac acc 2050
Thr Ser Asp Gln Thr Ser Thr Glu Ser Ala Thr Asn Glu Asn Asn Thr
610 615 620 625

aat cct gag cct cag ttc caa aca gaa gcc act ggg cct tca gct cat 2098
Asn Pro Glu Pro Gln Phe Gln Thr Glu Ala Thr Gly Pro Ser Ala His
630 635 640

gaa gaa aca tcc acc agg gac tct gct ctt cag gac aca gat gac agt 2146
Glu Glu Thr Ser Thr Arg Asp Ser Ala Leu Gln Asp Thr Asp Asp Ser
645 650 655

gat gat gac cca gtc ctg atc cca ggt gca agg tat cga gca gga cct 2194
Asp Asp Asp Pro Val Leu Ile Pro Gly Ala Arg Tyr Arg Ala Gly Pro
660 665 670

ggt gat aga cgc tct gct gtt gcc cgt att cag gag ttc ttc aga cgg 2242
Gly Asp Arg Arg Ser Ala Val Ala Arg Ile Gln Glu Phe Phe Arg Arg
675 680 685

aga aaa gaa agg aaa gaa atg gaa gaa ttg gat act ttg aac att aga 2290
Arg Lys Glu Arg Lys Glu Met Glu Glu Leu Asp Thr Leu Asn Ile Arg
690 695 700 705

agg ccg cta gta aaa atg gtt tat aaa ggc cat cgc aac tcc agg aca 2338
Arg Pro Leu Val Lys Met Val Tyr Lys Gly His Arg Asn Ser Arg Thr
710 715 720

atg ata aaa gaa gcc aat ttc tgg ggt gct aac ttt gta atg act ggt 2386
Met Ile Lys Glu Ala Asn Phe Trp Gly Ala Asn Phe Val Met Thr Gly
725 730 735

tct gag tgt ggc cac att ttc atc tgg gat cgg cac act gct gag cat 2434
Ser Glu Cys Gly His Ile Phe Ile Trp Asp Arg His Thr Ala Glu His
740 745 750

ttg atg ctt ctg gaa gct gat aat cat gtg gta aac tgc ctg cag cca 2482
Leu Met Leu Leu Glu Ala Asp Asn His Val Val Asn Cys Leu Gln Pro
755 760 765

cat ccg ttt gac cca att tta gcc tca tct ggc ata gat tat gac ata 2530
His Pro Phe Asp Pro Ile Leu Ala Ser Ser Gly Ile Asp Tyr Asp Ile
770 775 780 785

aag atc tgg tca cca tta gaa gag tca agg att ttt aac cga aaa ctt 2578
Lys Ile Trp Ser Pro Leu Glu Glu Ser Arg Ile Phe Asn Arg Lys Leu

790 795 800

gct gat gaa gtt ata act cga aac gaa ctc atg ctg gaa gaa act aga 2626
Ala Asp Glu Val Ile Thr Arg Asn Glu Leu Met Leu Glu Glu Thr Arg
805 810 815

aac acc att aca gtt cca gcc tct ttc atg ttg agg atg ttg gct tca 2674
Asn Thr Ile Thr Val Pro Ala Ser Phe Met Leu Arg Met Leu Ala Ser
820 825 830

ctt aat cat atc cga gct gac cgg ttg gag ggt gac aga tca gaa ggc 2722
Leu Asn His Ile Arg Ala Asp Arg Leu Glu Gly Asp Arg Ser Glu Gly
835 840 845

tct ggt caa gag aat gaa aat gag gat gag gaa taataaactc tttttggcaa 2775
Ser Gly Gln Glu Asn Glu Asn Glu Asp Glu Glu
850 855 860

gcacttaaat gttctgaaat ttgtataaga catttattat ttttttttct ttacagagat 2835

ttagtgcaat tttaaggtta tggtttttgg agtttttccc ttttttgggg ataacctaac 2895

attggtttgg aatgattgtg tgcatgaatt tgggagattg tataaaacaa aactagcaga 2955

atgtttttaa aacttttttgc cgtgtatgag gagtgctaga aaatgcaaag tgcaatattt 3015

tccctaacct tcaaatgtgg gagcttggat caatgttgaa gaataatttt catcatagtg 3075

aaaatgttgg ttcaaataaa tttctacact tgccatttgc atgtttgttg ctttctaatt 3135

aaagaaactg gttgttttaa gataccctga aaaaaaaaaa aaaaaaaaaa aaaa 3189

<210> 2
<211> 860
<212> PRT
<213> Homo sapiens

<400> 2
Met Ser Arg Gly Gly Ser Tyr Pro His Leu Leu Trp Asp Val Arg Lys
1 5 10 15

Arg Phe Leu Gly Leu Glu Asp Pro Ser Arg Leu Arg Ser Arg Tyr Leu
20 25 30

Gly Arg Arg Glu Phe Ile Gln Arg Leu Lys Leu Glu Ala Thr Leu Asn
35 40 45

```
Val His Asp Gly Cys Val Asn Thr Ile Cys Trp Asn Asp Thr Gly Glu
 50                  55                  60

Tyr Ile Leu Ser Gly Ser Asp Asp Thr Lys Leu Val Ile Ser Asn Pro
65                  70                  75                  80

Tyr Ser Arg Lys Val Leu Thr Thr Ile Arg Ser Gly His Arg Ala Asn
                85                  90                  95

Ile Phe Ser Ala Lys Phe Leu Pro Cys Thr Asn Asp Lys Gln Ile Val
            100                 105                 110

Ser Cys Ser Gly Asp Gly Val Ile Phe Tyr Thr Asn Val Glu Gln Asp
        115                 120                 125

Ala Glu Thr Asn Arg Gln Cys Gln Phe Th: Cys His Tyr Gly Thr Thr
    130                 135                 140

Tyr Glu Ile Met Thr Val Pro Asn Asp Pro Tyr Thr Phe Leu Ser Cys
145                 150                 155                 160

Gly Glu Asp Gly Thr Val Arg Trp Phe Asp Thr Arg Ile Lys Thr Ser
                165                 170                 175

Cys Thr Lys Glu Asp Cys Lys Asp Asp Ile Leu Ile Asn Cys Arg Arg
            180                 185                 190

Ala Ala Thr Ser Val Ala Ile Cys Pro Pro Ile Pro Tyr Tyr Leu Ala
        195                 200                 205

Val Gly Cys Ser Asp Ser Ser Val Arg Ile Tyr Asp Arg Arg Met Leu
    210                 215                 220

Gly Thr Arg Ala Thr Gly Asn Tyr Ala Gly Arg Gly Thr Thr Gly Met
225                 230                 235                 240

Val Ala Arg Phe Ile Pro Ser His Leu Asn Asn Lys Ser Cys Arg Val
                245                 250                 255

Thr Ser Leu Cys Tyr Ser Glu Asp Gly Gln Glu Ile Leu Val Ser Tyr
            260                 265                 270

Ser Ser Asp Tyr Ile Tyr Leu Phe Asp Pro Lys Asp Asp Thr Ala Arg
        275                 280                 285

Glu Leu Lys Thr Pro Ser Ala Glu Glu Arg Arg Glu Glu Leu Arg Gln
    290                 295                 300
```

```
Pro Pro Val Lys Arg Leu Arg Leu Arg Gly Asp Trp Ser Asp Thr Gly
305                 310                 315                 320

Pro Arg Ala Arg Pro Glu Ser Glu Arg Glu Arg Asp Gly Glu Gln Ser
                325                 330                 335

Pro Asn Val Ser Leu Met Gln Arg Met Ser Asp Met Leu Ser Arg Trp
            340                 345                 350

Phe Glu Glu Ala Ser Glu Val Ala Gln Ser Asn Arg Gly Arg Gly Arg
        355                 360                 365

Ser Arg Pro Arg Gly Gly Thr Ser Gln Ser Asp Ile Ser Thr Leu Pro
    370                 375                 380

Thr Val Pro Ser Ser Pro Asp Leu Glu Val Ser Glu Thr Ala Met Glu
385                 390                 395                 400

Val Asp Thr Pro Ala Glu Gln Phe Leu Gln Pro Ser Thr Ser Ser Thr
                405                 410                 415

Met Ser Ala Gln Ala His Ser Thr Ser Ser Pro Thr Glu Ser Pro His
            420                 425                 430

Ser Thr Pro Leu Leu Ser Ser Pro Asp Ser Glu Gln Arg Gln Ser Val
        435                 440                 445

Glu Ala Ser Gly His His Thr His His Gln Ser Asp Asn Asn Asn Glu
    450                 455                 460

Lys Leu Ser Pro Lys Pro Gly Thr Gly Glu Pro Val Leu Ser Leu His
465                 470                 475                 480

Tyr Ser Thr Glu Gly Thr Thr Thr Ser Thr Ile Lys Leu Asn Phe Thr
                485                 490                 495

Asp Glu Trp Ser Ser Ile Ala Ser Ser Ser Arg Gly Ile Gly Ser His
            500                 505                 510

Cys Lys Ser Glu Gly Gln Glu Glu Ser Phe Val Pro Gln Ser Ser Val
        515                 520                 525

Gln Pro Pro Glu Gly Asp Ser Glu Thr Lys Ala Pro Glu Glu Ser Ser
    530                 535                 540

Glu Asp Val Thr Lys Tyr Gln Glu Gly Val Ser Ala Glu Asn Pro Val
545                 550                 555                 560
```

Glu Asn His Ile Asn Ile Thr Gln Ser Asp Lys Phe Thr Ala Lys Pro
565 570 575

Leu Asp Ser Asn Ser Gly Glu Arg Asn Asp Leu Asn Leu Asp Arg Ser
580 585 590

Cys Gly Val Pro Glu Glu Ser Ala Ser Ser Glu Lys Ala Lys Glu Pro
595 600 605

Glu Thr Ser Asp Gln Thr Ser Thr Glu Ser Ala Thr Asn Glu Asn Asn
610 615 620

Thr Asn Pro Glu Pro Gln Phe Gln Thr Glu Ala Thr Gly Pro Ser Ala
625 630 635 640

His Glu Glu Thr Ser Thr Arg Asp Ser Ala Leu Gln Asp Thr Asp Asp
645 650 655

Ser Asp Asp Asp Pro Val Leu Ile Pro Gly Ala Arg Tyr Arg Ala Gly
660 665 670

Pro Gly Asp Arg Arg Ser Ala Val Ala Arg Ile Gln Glu Phe Phe Arg
675 680 685

Arg Arg Lys Glu Arg Lys Glu Met Glu Glu Leu Asp Thr Leu Asn Ile
690 695 700

Arg Arg Pro Leu Val Lys Met Val Tyr Lys Gly His Arg Asn Ser Arg
705 710 715 720

Thr Met Ile Lys Glu Ala Asn Phe Trp Gly Ala Asn Phe Val Met Thr
725 730 735

Gly Ser Glu Cys Gly His Ile Phe Ile Trp Asp Arg His Thr Ala Glu
740 745 750

His Leu Met Leu Leu Glu Ala Asp Asn His Val Val Asn Cys Leu Gln
755 760 765

Pro His Pro Phe Asp Pro Ile Leu Ala Ser Ser Gly Ile Asp Tyr Asp
770 775 780

Ile Lys Ile Trp Ser Pro Leu Glu Glu Ser Arg Ile Phe Asn Arg Lys
785 790 795 800

Leu Ala Asp Glu Val Ile Thr Arg Asn Glu Leu Met Leu Glu Glu Thr
805 810 815

```
Arg Asn Thr Ile Thr Val Pro Ala Ser Phe Met Leu Arg Met Leu Ala
            820                 825                 830

Ser Leu Asn His Ile Arg Ala Asp Arg Leu Glu Gly Asp Arg Ser Glu
        835                 840                 845

Gly Ser Gly Gln Glu Asn Glu Asn Glu Asp Glu Glu
    850                 855                 860
```

<210> 3
<211> 3420
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (176)..(2986)

<400> 3

```
cggccgggct tcaggggccc aggcgccgct gctgccaccg ccatctaacg ctgcgccctg 60

gaggcccggc gcgcggatgg tgccggtgcg gctcgggtgt tgaaacgggt gtcccctccc 120

cctcctcccc tccccacgc ggtggtctcc cctcccaccc ggctcaggca gagcc atg  178
                                                          Met
                                                           1

tct cgg ggt ggc tcc tac cca cac ctg ttg tgg gac gtg agg aaa agg   226
Ser Arg Gly Gly Ser Tyr Pro His Leu Leu Trp Asp Val Arg Lys Arg
              5                  10                  15

ttc ctc ggg ctg gag gac ccg tcc cgg ctg cgg agt cgc tac ctg gga   274
Phe Leu Gly Leu Glu Asp Pro Ser Arg Leu Arg Ser Arg Tyr Leu Gly
          20                  25                  30

aga aga gaa ttt atc caa aga tta aaa ctt gaa gca acc ctt aat gtg   322
Arg Arg Glu Phe Ile Gln Arg Leu Lys Leu Glu Ala Thr Leu Asn Val
      35                  40                  45

cat gat ggt tgt gtt aat aca atc tgt tgg aat gac act gga gaa tat   370
His Asp Gly Cys Val Asn Thr Ile Cys Trp Asn Asp Thr Gly Glu Tyr
  50                  55                  60                  65

att tta tct ggc tca gat gac acc aaa tta gta att agt aat cct tac   418
Ile Leu Ser Gly Ser Asp Asp Thr Lys Leu Val Ile Ser Asn Pro Tyr
                  70                  75                  80
```

```
agc aga aag gtt ttg aca aca att cgt tca ggg cac cga gca aac ata   466
Ser Arg Lys Val Leu Thr Thr Ile Arg Ser Gly His Arg Ala Asn Ile
            85                  90                  95

ttt agt gca aag ttc tta cct tgt aca aat gat aaa cag att gta tcc   514
Phe Ser Ala Lys Phe Leu Pro Cys Thr Asn Asp Lys Gln Ile Val Ser
        100                 105                 110

tgc tct gga gat gga gta ata ttt tat acc aac gtt gag caa gat gca   562
Cys Ser Gly Asp Gly Val Ile Phe Tyr Thr Asn Val Glu Gln Asp Ala
    115                 120                 125

gaa acc aac aga caa tgc caa ttt acg tgt cat tat gga act act tat   610
Glu Thr Asn Arg Gln Cys Gln Phe Thr Cys His Tyr Gly Thr Thr Tyr
130                 135                 140                 145

gag att atg act gta ccc aat gac cct tac act ttt ctc tct tgt ggt   658
Glu Ile Met Thr Val Pro Asn Asp Pro Tyr Thr Phe Leu Ser Cys Gly
                150                 155                 160

gaa gat gga act gtt agg tgg ttt gat aca cgc atc aaa act agc tgc   706
Glu Asp Gly Thr Val Arg Trp Phe Asp Thr Arg Ile Lys Thr Ser Cys
            165                 170                 175

aca aaa gaa gat tgt aaa gat gat att tta att aac tgt cga cgt gct   754
Thr Lys Glu Asp Cys Lys Asp Asp Ile Leu Ile Asn Cys Arg Arg Ala
        180                 185                 190

gcc acg tct gtt gct att tgc cca cca ata cca tat tac ctt gct gtt   802
Ala Thr Ser Val Ala Ile Cys Pro Pro Ile Pro Tyr Tyr Leu Ala Val
    195                 200                 205

ggt tgt tct gac agc tca gta cga ata tat gat cgg cga atg ctg ggc   850
Gly Cys Ser Asp Ser Ser Val Arg Ile Tyr Asp Arg Arg Met Leu Gly
210                 215                 220                 225

aca aga gct aca ggg aat tat gca ggt cga ggg act act gga atg gtt   898
Thr Arg Ala Thr Gly Asn Tyr Ala Gly Arg Gly Thr Thr Gly Met Val
                230                 235                 240

gcc cgt ttt att cct tcc cat ctt aat aat aag tcc tgc aga gtg aca   946
Ala Arg Phe Ile Pro Ser His Leu Asn Asn Lys Ser Cys Arg Val Thr
            245                 250                 255

tct ctg tgt tac agt gaa gat ggt caa gag att ctc gtt agt tac tct   994
Ser Leu Cys Tyr Ser Glu Asp Gly Gln Glu Ile Leu Val Ser Tyr Ser
        260                 265                 270
```

```
tca gat tac ata tat ctt ttt gac ccg aaa gat gat aca gca cga gaa   1042
Ser Asp Tyr Ile Tyr Leu Phe Asp Pro Lys Asp Asp Thr Ala Arg Glu
    275                 280                 285

ctt aaa act cct tct gcg gaa gag aga aga gaa gag ttg cga caa cca   1090
Leu Lys Thr Pro Ser Ala Glu Glu Arg Arg Glu Glu Leu Arg Gln Pro
290                 295                 300                 305

cca gtt aag cgt ttg aga ctt cgt ggt gat tgg tca gat act gga ccc   1138
Pro Val Lys Arg Leu Arg Leu Arg Gly Asp Trp Ser Asp Thr Gly Pro
                310                 315                 320

aga gca agg ccg gag agt gaa cga gaa cga gat gga gag cag agt ccc   1186
Arg Ala Arg Pro Glu Ser Glu Arg Glu Arg Asp Gly Glu Gln Ser Pro
            325                 330                 335

aat gtg tca ttg atg cag aga atg tct gat atg tta tca aga tgg ttt   1234
Asn Val Ser Leu Met Gln Arg Met Ser Asp Met Leu Ser Arg Trp Phe
        340                 345                 350

gaa gaa gca agt gag gtt gca caa agc aat aga gga cga gga aga tct   1282
Glu Glu Ala Ser Glu Val Ala Gln Ser Asn Arg Gly Arg Gly Arg Ser
    355                 360                 365

cga ccc aga ggt gga aca agt caa tca gat att tca act ctt cct acg   1330
Arg Pro Arg Gly Gly Thr Ser Gln Ser Asp Ile Ser Thr Leu Pro Thr
370                 375                 380                 385

gtc cca tca agt cct gat ttg gaa gtg agt gaa act gca atg gaa gta   1378
Val Pro Ser Ser Pro Asp Leu Glu Val Ser Glu Thr Ala Met Glu Val
                390                 395                 400

gat act cca gct gaa caa ttt ctt cag cct tct aca tcc tct aca atg   1426
Asp Thr Pro Ala Glu Gln Phe Leu Gln Pro Ser Thr Ser Ser Thr Met
            405                 410                 415

tca gct cag gct cat tcg aca tca tct ccc aca gaa agc cct cat tct   1474
Ser Ala Gln Ala His Ser Thr Ser Ser Pro Thr Glu Ser Pro His Ser
        420                 425                 430

act cct ttg cta tct tct cca gat agt gaa caa agg cag tct gtt gag   1522
Thr Pro Leu Leu Ser Ser Pro Asp Ser Glu Gln Arg Gln Ser Val Glu
    435                 440                 445

gca tct gga cac cac aca cat cat cag tct gaa ttt tta agg ggg cct   1570
Ala Ser Gly His His Thr His His Gln Ser Glu Phe Leu Arg Gly Pro
450                 455                 460                 465
```

```
gag ata gct ttg ctt cgt aag cgc ctg caa caa ctg agg ctt aag aag   1618
Glu Ile Ala Leu Leu Arg Lys Arg Leu Gln Gln Leu Arg Leu Lys Lys
                470                 475                 480

gct gag cag cag agg cag caa gag cta gct gca cat acc cag caa cag   1666
Ala Glu Gln Gln Arg Gln Gln Glu Leu Ala Ala His Thr Gln Gln Gln
            485                 490                 495

cct tcc act tct gat cag tct tct cat gag ggc tct tca cag gac cct   1714
Pro Ser Thr Ser Asp Gln Ser Ser His Glu Gly Ser Ser Gln Asp Pro
        500                 505                 510

cat gct tca gat tct cct tct tct gtg gtt aac aaa cag ctc gga tcc   1762
His Ala Ser Asp Ser Pro Ser Ser Val Val Asn Lys Gln Leu Gly Ser
    515                 520                 525

atg tca ctt gac gag caa cag gat aac aat aat gaa aag ctg agc ccc   1810
Met Ser Leu Asp Glu Gln Gln Asp Asn Asn Asn Glu Lys Leu Ser Pro
530                 535                 540                 545

aaa cca ggg aca ggt gaa cca gtt tta agt ttg cac tac agc aca gaa   1858
Lys Pro Gly Thr Gly Glu Pro Val Leu Ser Leu His Tyr Ser Thr Glu
                550                 555                 560

gga aca act aca agc aca ata aaa ctg aac ttt aca gat gaa tgg agc   1906
Gly Thr Thr Thr Ser Thr Ile Lys Leu Asn Phe Thr Asp Glu Trp Ser
            565                 570                 575

agt ata gca tca agt tct aga gga att ggg agc cat tgc aaa tct gag   1954
Ser Ile Ala Ser Ser Ser Arg Gly Ile Gly Ser His Cys Lys Ser Glu
        580                 585                 590

ggt cag gag gaa tct ttc gtc cca cag agc tca gtg caa cca cca gaa   2002
Gly Gln Glu Glu Ser Phe Val Pro Gln Ser Ser Val Gln Pro Pro Glu
    595                 600                 605

gga gac agt gaa aca aaa gct cct gaa gaa tca tca gag gat gtg aca   2050
Gly Asp Ser Glu Thr Lys Ala Pro Glu Glu Ser Ser Glu Asp Val Thr
610                 615                 620                 625

aaa tat cag gaa gga gta tct gca gaa aac cca gtt gag aac cat atc   2098
Lys Tyr Gln Glu Gly Val Ser Ala Glu Asn Pro Val Glu Asn His Ile
                630                 635                 640

aat ata aca caa tca gat aag ttc aca gcc aag cca ttg gat tcc aac   2146
Asn Ile Thr Gln Ser Asp Lys Phe Thr Ala Lys Pro Leu Asp Ser Asn
            645                 650                 655
```

```
tca gga gaa aga aat gac ctc aat ctt gat cgc tct tgt ggg gtt cca   2194
Ser Gly Glu Arg Asn Asp Leu Asn Leu Asp Arg Ser Cys Gly Val Pro
        660                 665                 670

gaa gaa tct gct tca tct gaa aaa gcc aag gaa cca gaa act tca gat   2242
Glu Glu Ser Ala Ser Ser Glu Lys Ala Lys Glu Pro Glu Thr Ser Asp
    675                 680                 685

cag act agc act gag agt gct acc aat gaa aat aac acc aat cct gag   2290
Gln Thr Ser Thr Glu Ser Ala Thr Asn Glu Asn Asn Thr Asn Pro Glu
690                 695                 700                 705

cct cag ttc caa aca gaa gcc act ggg cct tca gct cat gaa gaa aca   2338
Pro Gln Phe Gln Thr Glu Ala Thr Gly Pro Ser Ala His Glu Glu Thr
                710                 715                 720

tcc acc agg gac tct gct ctt cag gac aca gat gac agt gat gat gac   2386
Ser Thr Arg Asp Ser Ala Leu Gln Asp Thr Asp Asp Ser Asp Asp Asp
            725                 730                 735

cca gtc ctg atc cca ggt gca agg tat cga gca gga cct ggt gat aga   2434
Pro Val Leu Ile Pro Gly Ala Arg Tyr Arg Ala Gly Pro Gly Asp Arg
        740                 745                 750

cgc tct gct gtt gcc cgt att cag gag ttc ttc aga cgg aga aaa gaa   2482
Arg Ser Ala Val Ala Arg Ile Gln Glu Phe Phe Arg Arg Arg Lys Glu
    755                 760                 765

agg aaa gaa atg gaa gaa ttg gat act ttg aac att aga agg ccg cta   2530
Arg Lys Glu Met Glu Glu Leu Asp Thr Leu Asn Ile Arg Arg Pro Leu
770                 775                 780                 785

gta aaa atg gtt tat aaa ggc cat cgc aac tcc agg aca atg ata aaa   2578
Val Lys Met Val Tyr Lys Gly His Arg Asn Ser Arg Thr Met Ile Lys
                790                 795                 800

gaa gcc aat ttc tgg ggt gct aac ttt gta atg act ggt tct gag tgt   2626
Glu Ala Asn Phe Trp Gly Ala Asn Phe Val Met Thr Gly Ser Glu Cys
            805                 810                 815

ggc cac att ttc atc tgg gat cgg cac act gct gag cat ttg atg ctt   2674
Gly His Ile Phe Ile Trp Asp Arg His Thr Ala Glu His Leu Met Leu
        820                 825                 830

ctg gaa gct gat aat cat gtg gta aac tgc ctg cag cca cat ccg ttt   2722
Leu Glu Ala Asp Asn His Val Val Asn Cys Leu Gln Pro His Pro Phe
    835                 840                 845
```

```
gac cca att tta gcc tca tct ggc ata gat tat gac ata aag atc tgg   2770
Asp Pro Ile Leu Ala Ser Ser Gly Ile Asp Tyr Asp Ile Lys Ile Trp
850                 855                 860                 865

tca cca tta gaa gag tca agg att ttt aac cga aaa ctt gct gat gaa   2818
Ser Pro Leu Glu Glu Ser Arg Ile Phe Asn Arg Lys Leu Ala Asp Glu
                870                 875                 880

gtt ata act cga aac gaa ctc atg ctg gaa gaa act aga aac acc att   2866
Val Ile Thr Arg Asn Glu Leu Met Leu Glu Glu Thr Arg Asn Thr Ile
            885                 890                 895

aca gtt cca gcc tct ttc atg ttg agg atg ttg gct tca ctt aat cat   2914
Thr Val Pro Ala Ser Phe Met Leu Arg Met Leu Ala Ser Leu Asn His
        900                 905                 910

atc cga gct gac cgg ttg gag ggt gac aga tca gaa ggc tct ggt caa   2962
Ile Arg Ala Asp Arg Leu Glu Gly Asp Arg Ser Glu Gly Ser Gly Gln
    915                 920                 925

gag aat gaa aat gag gat gag gaa taataaactc tttttggcaa gcacttaaat  3016
Glu Asn Glu Asn Glu Asp Glu Glu
930                 935

gttctgaaat ttgtataaga catttattat ttttttttct ttacagagat ttagtgcaat 3076

tttaaggtta tggtttttgg agtttttccc ttttttgggg ataacctaac attggtttgg 3136

aatgattgtg tgcatgaatt tgggagattg tataaaacaa aactagcaga atgtttttaa 3196

aacttttttgc cgtgtatgag gagtgctaga aaatgcaaag tgcaatattt tccctaacct 3256

tcaaatgtgg gagcttggat caatgttgaa gaataatttt catcatagtg aaaatgttgg 3316

ttcaaataaa tttctacact tgccatttgc atgtttgttg ctttctaatt aaagaaactg 3376

gttgttttaa gataccctga aaaaaaaaaa aaaaaaaaaa aaaa                  3420
```

<210>
<211> 937
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Ser Arg Gly Gly Ser Tyr Pro His Leu Leu Trp Asp Val Arg Lys
 1               5                  10                  15
```

```
Arg Phe Leu Gly Leu Glu Asp Pro Ser Arg Leu Arg Ser Arg Tyr Leu
            20                  25                  30

Gly Arg Arg Glu Phe Ile Gln Arg Leu Lys Leu Glu Ala Thr Leu Asn
        35                  40                  45

Val His Asp Gly Cys Val Asn Thr Ile Cys Trp Asn Asp Thr Gly Glu
    50                  55                  60

Tyr Ile Leu Ser Gly Ser Asp Asp Thr Lys Leu Val Ile Ser Asn Pro
65                  70                  75                  80

Tyr Ser Arg Lys Val Leu Thr Thr Ile Arg Ser Gly His Arg Ala Asn
                85                  90                  95

Ile Phe Ser Ala Lys Phe Leu Pro Cys Thr Asn Asp Lys Gln Ile Val
            100                 105                 110

Ser Cys Ser Gly Asp Gly Val Ile Phe Tyr Thr Asn Val Glu Gln Asp
        115                 120                 125

Ala Glu Thr Asn Arg Gln Cys Gln Phe Thr Cys His Tyr Gly Thr Thr
    130                 135                 140

Tyr Glu Ile Met Thr Val Pro Asn Asp Pro Tyr Thr Phe Leu Ser Cys
145                 150                 155                 160

Gly Glu Asp Gly Thr Val Arg Trp Phe Asp Thr Arg Ile Lys Thr Ser
                165                 170                 175

Cys Thr Lys Glu Asp Cys Lys Asp Asp Ile Leu Ile Asn Cys Arg Arg
            180                 185                 190

Ala Ala Thr Ser Val Ala Ile Cys Pro Pro Ile Pro Tyr Tyr Leu Ala
        195                 200                 205

Val Gly Cys Ser Asp Ser Ser Val Arg Ile Tyr Asp Arg Arg Met Leu
    210                 215                 220

Gly Thr Arg Ala Thr Gly Asn Tyr Ala Gly Arg Gly Thr Thr Gly Met
225                 230                 235                 240

Val Ala Arg Phe Ile Pro Ser His Leu Asn Asn Lys Ser Cys Arg Val
                245                 250                 255

Thr Ser Leu Cys Tyr Ser Glu Asp Gly Gln Glu Ile Leu Val Ser Tyr
            260                 265                 270
```

```
Ser Ser Asp Tyr Ile Tyr Leu Phe Asp Pro Lys Asp Asp Thr Ala Arg
        275                 280                 285

Glu Leu Lys Thr Pro Ser Ala Glu Glu Arg Arg Glu Glu Leu Arg Gln
    290                 295                 300

Pro Pro Val Lys Arg Leu Arg Leu Arg Gly Asp Trp Ser Asp Thr Gly
305                 310                 315                 320

Pro Arg Ala Arg Pro Glu Ser Glu Arg Glu Arg Asp Gly Glu Gln Ser
                325                 330                 335

Pro Asn Val Ser Leu Met Gln Arg Met Ser Asp Met Leu Ser Arg Trp
            340                 345                 350

Phe Glu Glu Ala Ser Glu Val Ala Gln Ser Asn Arg Gly Arg Gly Arg
        355                 360                 365

Ser Arg Pro Arg Gly Gly Thr Ser Gln Ser Asp Ile Ser Thr Leu Pro
    370                 375                 380

Thr Val Pro Ser Ser Pro Asp Leu Glu Val Ser Glu Thr Ala Met Glu
385                 390                 395                 400

Val Asp Thr Pro Ala Glu Gln Phe Leu Gln Pro Ser Thr Ser Ser Thr
                405                 410                 415

Met Ser Ala Gln Ala His Ser Thr Ser Ser Pro Thr Glu Ser Pro His
            420                 425                 430

Ser Thr Pro Leu Leu Ser Ser Pro Asp Ser Glu Gln Arg Gln Ser Val
        435                 440                 445

Glu Ala Ser Gly His His Thr His His Gln Ser Glu Phe Leu Arg Gly
    450                 455                 460

Pro Glu Ile Ala Leu Leu Arg Lys Arg Leu Gln Gln Leu Arg Leu Lys
465                 470                 475                 480

Lys Ala Glu Gln Gln Arg Gln Gln Glu Leu Ala Ala His Thr Gln Gln
                485                 490                 495

Gln Pro Ser Thr Ser Asp Gln Ser Ser His Glu Gly Ser Ser Gln Asp
            500                 505                 510

Pro His Ala Ser Asp Ser Pro Ser Ser Val Val Asn Lys Gln Leu Gly
        515                 520                 525
```

```
Ser Met Ser Leu Asp Glu Gln Gln Asp Asn Asn Asn Glu Lys Leu Ser
 530                 535                 540

Pro Lys Pro Gly Thr Gly Glu Pro Val Leu Ser Leu His Tyr Ser Thr
545                 550                 555                 560

Glu Gly Thr Thr Thr Ser Thr Ile Lys Leu Asn Phe Thr Asp Glu Trp
                565                 570                 575

Ser Ser Ile Ala Ser Ser Ser Arg Gly Ile Gly Ser His Cys Lys Ser
            580                 585                 590

Glu Gly Gln Glu Glu Ser Phe Val Pro Gln Ser Ser Val Gln Pro Pro
        595                 600                 605

Glu Gly Asp Ser Glu Thr Lys Ala Pro Glu Glu Ser Ser Glu Asp Val
    610                 615                 620

Thr Lys Tyr Gln Glu Gly Val Ser Ala Glu Asn Pro Val Glu Asn His
625                 630                 635                 640

Ile Asn Ile Thr Gln Ser Asp Lys Phe Thr Ala Lys Pro Leu Asp Ser
                645                 650                 655

Asn Ser Gly Glu Arg Asn Asp Leu Asn Leu Asp Arg Ser Cys Gly Val
            660                 665                 670

Pro Glu Glu Ser Ala Ser Ser Glu Lys Ala Lys Glu Pro Glu Thr Ser
        675                 680                 685

Asp Gln Thr Ser Thr Glu Ser Ala Thr Asn Glu Asn Asn Thr Asn Pro
    690                 695                 700

Glu Pro Gln Phe Gln Thr Glu Ala Thr Gly Pro Ser Ala His Glu Glu
705                 710                 715                 720

Thr Ser Thr Arg Asp Ser Ala Leu Gln Asp Thr Asp Asp Ser Asp Asp
                725                 730                 735

Asp Pro Val Leu Ile Pro Gly Ala Arg Tyr Arg Ala Gly Pro Gly Asp
            740                 745                 750

Arg Arg Ser Ala Val Ala Arg Ile Gln Glu Phe Phe Arg Arg Arg Lys
        755                 760                 765

Glu Arg Lys Glu Met Glu Glu Leu Asp Thr Leu Asn Ile Arg Arg Pro
    770                 775                 780
```

```
Leu Val Lys Met Val Tyr Lys Gly His Arg Asn Ser Arg Thr Met Ile
785                 790                 795                 800

Lys Glu Ala Asn Phe Trp Gly Ala Asn Phe Val Met Thr Gly Ser Glu
                805                 810                 815

Cys Gly His Ile Phe Ile Trp Asp Arg His Thr Ala Glu His Leu Met
            820                 825                 830

Leu Leu Glu Ala Asp Asn His Val Val Asn Cys Leu Gln Pro His Pro
        835                 840                 845

Phe Asp Pro Ile Leu Ala Ser Ser Gly Ile Asp Tyr Asp Ile Lys Ile
    850                 855                 860

Trp Ser Pro Leu Glu Glu Ser Arg Ile Phe Asn Arg Lys Leu Ala Asp
865                 870                 875                 880

Glu Val Ile Thr Arg Asn Glu Leu Met Leu Glu Glu Thr Arg Asn Thr
                885                 890                 895

Ile Thr Val Pro Ala Ser Phe Met Leu Arg Met Leu Ala Ser Leu Asn
            900                 905                 910

His Ile Arg Ala Asp Arg Leu Glu Gly Asp Arg Ser Glu Gly Ser Gly
        915                 920                 925

Gln Glu Asn Glu Asn Glu Asp Glu Glu
    930                 935
```

<210> 5
<211> 21
<212> DNA
<213> Homo sapiens

<400> 5

```
caggacacag atgacagtga t                                              21
```

<210> 6
<211> 20
<212> DNA
<213> Homo sapiens

<400> 6

```
agagccttct gatctgtcac                                                20
```

<210> 7
<211> 9
<212> PRT

<213> Homo sapiens

<400> 7

Lys Thr Pro Ser Ala Glu Glu Arg Arg
1 5

<210> 8
<211> 8
<212> PRT
<213> Homo sapiens

<400> 8

Arg Ala Arg Pro Glu Ser Glu Arg
1 5

<210> 9
<211> 7
<212> PRT
<213> Homo sapiens

<400> 9

Arg Met Ser Asp Met Ser Arg
1 5

<210> 10
<211> 9
<212> PRT
<213> Homo sapiens

<400> 10

Asn Glu Lys Leu Ser Pro Lys Pro Gly
1 5

<210> 11
<211> 35
<212> PRT
<213> Homo sapiens

<400> 11

Pro Asn Val Ser Leu Met Gln Arg Met Ser Asp Met Leu Ser Arg Trp
1 5 10 15

Phe Glu Glu Ala Ser Glu Val Ala Gln Ser Asn Arg Gly Arg Gly Arg
20 25 30

Ser Arg Pro
35

<210> 12
<211> 25
<212> PRT
<213> Homo sapiens

<400> 12

Val Pro Ser Ser Pro Asp Leu Glu Val Ser Glu Thr Ala Met Glu Val
1 5 10 15

Asp Thr Pro Ala Glu Gln Phe Leu Gln
20 25

<210> 13
<211> 25
<212> PRT
<213> Homo sapiens

<400> 13

Pro Val Leu Ser Leu His Tyr Ser Thr Glu Gly Thr Thr Thr Ser Thr
1 5 10 15

Ile Lys Leu Asn Phe Thr Asp Glu Trp
20 25

<210> 14
<211> 24
<212> PRT
<213> Homo sapiens

<400> 14

Glu Thr Lys Ala Pro Glu Glu Ser Ser Glu Asp Val Thr Lys Tyr Gln
1 5 10 15

Glu Gly Val Ser Ala Glu Asn Pro
20

<210> 15
<211> 25
<212> PRT

<213> Homo sapiens

<400> 15

Glu Asn His Ile Asn Ile Thr Gln Ser Asp Lys Phe Thr Ala Lys Pro
1 5 10 15

Leu Asp Ser Asn Ser Gly Glu Arg Asn
20 25

<210> 16
<211> 24
<212> PRT
<213> Homo sapiens

<400> 16

Asn Thr Asn Pro Glu Pro Gln Phe Gln Thr Glu Ala Thr Gly Pro Ser
1 5 10 15

Ala His Glu Glu Thr Ser Thr Arg
20

<210> 17
<211> 60
<212> PRT
<213> Homo sapiens

<400> 17

Asp Arg Arg Ser Ala Val Ala Arg Ile Gln Glu Phe Phe Arg Arg Arg
1 5 10 15

Lys Glu Arg Lys Glu Met Glu Glu Leu Asp Thr Leu Asn Ile Arg Arg
20 25 30

Pro Leu Val Lys Met Val Tyr Lys Gly His Arg Asn Ser Arg Thr Met
35 40 45

Ile Lys Glu Ala Asn Phe Trp Gly Ala Asn Phe Val
50 55 60

<210> 18
<211> 37
<212> PRT
<213> Homo sapiens

<400> 18

```
Glu Cys Gly His Ile Phe Ile Trp Asp Arg His Thr Ala Glu His Leu
 1               5                  10                  15

Met Leu Leu Glu Ala Asp Asn His Val Val Asn Cys Leu Gln Pro His
            20                  25                  30

Pro Phe Asp Pro Ile
        35
```

<210> 19
<211> 60
<212> PRT
<213> Homo sapiens

<400> 19

```
Leu Ala Ser Ser Gly Ile Asp Tyr Asp Ile Lys Ile Trp Ser Pro Leu
 1               5                  10                  15

Glu Glu Ser Arg Ile Phe Asn Arg Lys Leu Ala Asp Glu Val Ile Thr
            20                  25                  30

Arg Asn Glu Leu Met Leu Glu Glu Thr Arg Asn Thr Ile Thr Val Pro
        35                  40                  45

Ala Ser Phe Met Leu Arg Met Leu Ala Ser Leu Asn
    50                  55                  60
```

<210> 20
<211> 11
<212> PRT
<213> Homo sapiens

<400> 20

```
Ser Gly Gln Glu Asn Glu Asn Glu Asp Glu Glu
 1               5                  10
```

Claims

1. A recombinant polynucleotide comprising a nucleotide sequence encoding polypeptide Repro-EN-1.0 of SEQ ID NO: 2, wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis.

2. The polynucleotide of claim 1 wherein the nucleotide sequence is selected from the Repro-EN-1.0 sequence of SEQ ID NO: 1.

3. The polynuclectide of claim 1 wherein the nucleotide sequence is a native Repro-EN-1.0 nucleotide sequence.

4. The polynucleotide of claim I wherein the nucleotide sequence is identical to nucleotides 176 to 2755 of SEQ ID NO: 1.

5. A recombinant polynucleotide composing a nucleotide sequence encoding polypeptide IB1 of SEQ ID NO: 4, wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis.

6. The polynucleotide of claim 5 wherein the nucleotide sequence is selected from the IB1 sequence of SEQ ID NO: 3.

7. The polynucleotide of claim 5 wherein the nucleotide sequence is a native IB1 nucleotide sequence.

8. The polynucleotide of claim 5 wherein the nucleotide sequence is identical to nucleotides 176 to 2986 of SEQ ID NO: 3.

9. The polynucleotide of any one of claims 1 to 8 further comprising an expression control sequence operatively linked to the nucleotide sequence.

10. A polynucleotide primer pair which amplifies a nucleotide sequence encoding polypeptide Repro-EN-1.0 of SEQ ID NO:2 or polypeptide IB1 of SEQ ID NO:4, wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis, the pair comprising:

(1) a 3' primer of at least 7 nucleotides that specifically hybridizes to a 3' end of the nucleotide sequence or downstream from the sequence, and
(2) a 5' primer of at least 7 nucleotides that specifically hybridizes to the 3' end of the complement of the nucleotide sequence or downstream from the complement of the sequence.

11. The polynucleotide primer pair of claim 10 wherein the 3' primer has a sequence complementary to a nucleotide sequence selected from Repro-EN-1.0 cDNA of SEQ ID NO: 1, and the 5' primer has a sequence identical to nucleotide sequence selected from Repro-EN-1.0 cDNA of SEQ ID NO: 1.

12. The polynucleotide primer pair of claim 10 wherein the 3' primer has a sequence complementary to a nucleotide sequence selected from IB1 cDNA of SEQ ID NO: 3, and the 5' primer has a sequence identical to nucleotide sequence selected from IB1 cDNA of SEQ ID NO: 3.

13. The polynucleotide primer pair of claim 10 wherein the pair of polynucleotides are peptide nucleic acids.

14. A recombinant cell comprising a recombinant polynucleotide according to claim 9.

15. A purified, recombinant Repro-EN-1.0 polypeptide whose amino acid sequence is identical to that of SEQ ID NO: 2, or an allelic variant of SEQ ID NO: 2.

16. A purified, recombinant IB1 polypeptide whose amino acid sequence is identical to that of SEQ ID NO: 4, or an allelic variant of SEQ ID NO: 4.

17. A purified polypeptide comprising Repro-EN-1.0 of SEQ ID NO: 2 or IB1 of SEQ ID NO:4, wherein the polypeptide specifically binds to antibodies from subjects diagnosed with endometriosis.

18. A composition consisting essentially of an antibody that specifically binds to Repro-EN-1,0 polypeptide of SEQ ID NO: 2 or an antibody that specifically binds to IB1 polypeptide of SEQ ID NO: 4.

19. The composition of claim 18 wherein the antibody is monoclonal.

20. A method for diagnosing endometriosis in a subject comprising the steps of:

(a) detecting a test amount of an antibody that specifically binds to Repro-EN-1.0 polypeptide whose amino acid sequence is identical to that of SEQ ID NO:2 or an allelic variant of SEQ ID NO:2 or a IB1 polypeptide whose amino acid sequence is identical to that of SEQ ID NO:4 or an allelic variant of SEQ ID NO:4 in a sample from the subject; and
(b) comparing the test amount with a normal range of the antibody in a control sample from a subject who does not suffer from endometriosis,

whereby a test amount above the normal range provides a positive indication in the diagnosis of endometriosis.

21. The method of claim 20 wherein the step of detecting comprises capturing the antibody from the sample with an immobilized Repro-EN-1.0 or IB1 and detecting captured antibody.

22. The method of claim 20 wherein the step of detecting comprises capturing the antibody from the sample with an immobilized anti-immunoglobulin antibody and detecting captured antibody.

23. The method of claim 21 wherein the step of detecting captured antibody comprises contacting the captured antibody with a detectable antibody that specifically binds immunoglobulins and detecting binding between the captured antibody and the detectable antibody.

24. The method of claim 22 wherein the step of detecting captured antibody comprises contacting the captured antibody with Repro-EN-1.0 or IB1 and detecting binding between the captured antibody and the Repro-EN-1.0 or IB1.

25. The method of any one of claims 20 to 24 wherein the sample comprises blood serum.

26. The method of any one of claims 20 to 25 wherein the antibody is an IgG immunoglobulin, an IgE immunoglobulin or an IgG_4 immunoglobulin.

27. A method for use in following the progress of endometriosis in a subject comprising the steps of:

(a) detecting first and second amounts of an antibody that specifically binds Repro-EN-1.0 polypeptide whose amino acid sequence is identical to that of SEQ ID NO:2 or an allelic variant of SEQ ID NO:2 or a IB1 polypeptide whose amino acid sequence is identical to that of SEQ NO:4 or an allelic variant of SEQ ID NO:4 in samples from the subject at a first and a second time, respectively; and
(b) comparing the first and second amounts,

whereby an increase between the first and second amounts indicates progression of the endometriosis and a decrease between the first and second amounts indicates remission of the endometriosis

28. A screening method for determining whether a compound increases or decreases the expression of Repro-EN-1.0 of SEQ ID NO: 2 or IB1 of SEQ ID NO: 4 or an allelic variant thereof in a cell, comprising contacting the cell with the compound and determining whether the production of Repro-EN-1.0 or IB1 mRNA or polypeptide is increased or decreased.

29. A method of detecting polymorphic forms of Repro-EN-1.0 of SEQ ID NO:1 or 2 or IB1 of SEQ ID NO:3 or 4 comprising the steps of

(a) determining the identity of a nucleotide or amino acid at a selected position within the sequence of a test Repro-EN-1.0 gene or polypeptide or a test IB1 gene or polypeptide;
(b) determining the identity of the nucleotide or amino acid at the corresponding position of native Repro-EN-1.0 gene or polypeptide of SEQ ID NO: 1 or 2 or native IB1 gene or polypeptide of SEQ ID NO: 3 or 4 and
(c) comparing the identity from the test gene or polypeptide with the identity of the native gene or polypeptide, whereby a difference in identity indicates that the test gene or polypeptide is a polymorphic form of Repro-EN-1.0 or IB1.

Patentansprüche

1. Rekombinantes Polynukleotid, umfassend eine Nukleotidsequenz, die das Polypeptid Repro-EN-1,0 der SEQ ID Nr. 2 codiert, wobei das Polypeptid spezifisch an Antikörper von Patienten mit der Diagnose Endometriose bindet.

2. Polynukleotid nach Anspruch 1, worin die Nukleotidsequenz gewählt ist aus der Repro-EN-1,0-Sequenz der SEQ ID Nr. 1.

3. Polynukleotid nach Anspruch 1, worin die Nukleotidsequenz eine native Repro-EN-1,0-Nukleotidsequenz ist.

4. Polynukleotid nach Anspruch 1, worin die Nukleotidsequenz identisch zu Nukleotiden 176 bis 2755 der SEQ ID Nr. 1 ist.

5. Rekombinantes Polynukleotid, umfassend eine Nukleotidsequenz, die das Polypeptid IB1 der SEQ ID Nr. 4 codiert, wobei das Polypeptid spezifisch an Antikörper von Patienten mit der Diagnose Endometriose bindet.

6. Polynukleotid nach Anspruch 5, wobei die Nukleotidsequenz gewählt ist aus der IB1-Sequenz von SEQ ID Nr. 3.

7. Polynukleotid nach Anspruch 5, worin die Nukleotidsequenz eine native IB1-Nukleotidsequenz ist.

8. Polynukleotid nach Anspruch 5, worin die Nukleotidsequenz identisch zu Nukleotiden 176 bis 2986 der SEQ ID Nr. 3 ist.

9. Polynukleotid nach einem der Ansprüche 1 bis 8, außerdem umfassend eine Expressionskontrollsequenz, die operativ an die Nukleotidsequenz geknüpft ist.

10. Polynukleotid-Primerpaar, welches eine Nukleotidsequenz amplifiziert, die das Polypeptid Repro-EN-1,0 der SEQ ID Nr. 2 oder das Polypeptid IB1 der SEQ ID Nr. 4 codiert, wobei das Polypeptid spezifisch an Antikörper von Patienten mit der Diagnose Endometriose bindet, welches Paar umfasst:

(1) einen 3'-Primer von mindestens 7 Nukleotiden, der spezifisch an ein 3'-Ende der Nukleotidsequenz oder stromabwärts von der Sequenz hybridisiert, und

(2) einen 5'-Primer von mindestens 7 Nukleotiden, der spezifisch an das 3'-Ende des Komplements der Nukleotidsequenz oder stromabwärts vom Komplement der Sequenz hybridisiert.

11. Polynukleotid-Primerpaar nach Anspruch 10, worin der 3'-Primer eine Sequenz aufweist, die komplementär zu einer Nukleotidsequenz ist, gewählt aus Repro-EN-1,0-cDNA der SEQ ID Nr. 1, und der 5'-Primer eine Sequenz aufweist, die identisch zur Nukleotidsequenz ist, gewählt aus Repro-EN-1,0-cDNA der SEQ ID Nr. 1.

12. Polynukleotid-Primerpaar nach Anspruch 10, worin der 3'-Primer eine Sequenz aufweist, die komplementär zu einer Nukleotidsequenz ist, gewählt aus IB1-cDNA der SEQ ID Nr. 3, und der 5'-Primer eine Sequenz aufweist, die identisch zur Nukleotidsequenz ist, gewählt aus IB1-cDNA der SEQ ID Nr. 3.

13. Polynukleotid-Primerpaar nach Anspruch 10, wobei das Paar der Polynukleotide Peptid-Nukleinsäuren sind.

14. Rekombinante Zelle, umfassend ein rekombinantes Polynukleotid nach Anspruch 9.

15. Gereinigtes, rekombinantes Repro-EN-1,0-Polypeptid, dessen Aminosäuresequenz identisch zu der von SEQ ID Nr. 2 oder einer Allel-Variante der SEQ ID Nr. 2 ist.

16. Gereinigtes, rekombinantes IB1-Polypeptid, dessen Aminosäuresequenz identisch zu der von SEQ ID Nr. 4 oder einer Allel-Variante der SEQ ID Nr. 4 ist.

17. Gereinigtes Polypeptid, umfassend Repro-EN-1,0 der SEQ ID Nr. 2 oder IB1 der SEQ ID Nr. 4, wobei das Polypeptid spezifisch an Antikörper von Patienten mit der Diagnose Endometriose bindet.

18. Zusammensetzung, die im wesentlichen aus einem Antikörper besteht, der spezifisch an Repro-EN-1,0-Polypeptid der SEQ ID Nr. 2 bindet, oder einem Antikörper, der spezifisch an IB1-Polypeptid der SEQ ID Nr. 4 bindet.

19. Zusammensetzung nach Anspruch 18, worin der Antikörper monoklonal ist.

20. Verfahren zur Diagnose von Endometriose bei einem Patienten, welches die folgenden Schritte umfasst:

(a) Nachweisen einer Testmenge eines Antikörpers, der spezifisch an Repro-EN-1,0-Polypeptid bindet, dessen Aminosäuresequenz identisch zu der von SEQ ID Nr. 2 oder einer Allel-Variante von SEQ ID Nr. 2 ist, oder IB1-Polypeptid, dessen Aminosäuresequenz identisch zu der von SEQ ID Nr. 4 oder einer Allel-Variante von SEQ ID Nr. 4 ist, in einer Probe von dem Patienten; und

(b) Vergleichen der Testmenge mit einem normalen Bereich des Antikörpers in einer Kontrollprobe von einem Patienten, der nicht an Endometriose leidet,

wobei eine Testmenge oberhalb des normalen Bereichs eine positive Indikation in der Diagnose der Endometriose erbringt.

21. Verfahren nach Anspruch 20, wobei der Schritt des Nachweisens das Einfangen des Antikörpers aus der Probe mit einem immobilisierten Repro-EN-1,0 oder IB1 und das Nachweisen des eingefangenen Antikörpers umfasst.

22. Verfahren nach Anspruch 20, wobei der Schritt des Nachweisens das Einfangen des Antikörpers aus der Probe mit einem immobilisierten Anti-Immunglobulin-Antikörper und das Nachweisen des eingefangenen Antikörpers umfasst.

23. Verfahren nach Anspruch 21, wobei der Schritt des Nachweisens des eingefangenen Antikörpers das Zusammenbringen des eingefangenen Antikörpers mit einem nachweisbaren Antikörper, der Immunglobuline spezifisch bindet, und das Nachweisen der Bindung zwischen dem eingefangenen Antikörper und dem nachweisbaren Antikörper umfasst.

24. Verfahren nach Anspruch 22, wobei der Schritt des Nachweisens des eingefangenen Antikörpers das Zusammenbringen des eingefangenen Antikörpers mit Repro-EN-1,0 oder IB1 und das Nachweisen der Bindung zwischen dem eingefangenen Antikörper und dem Repro-EN-1,0 oder IB1 umfasst.

25. Verfahren nach einem der Ansprüche 20 bis 24, wobei die Probe Blutserum umfasst.

26. Verfahren nach einem der Ansprüche 20 bis 25, wobei der Antikörper ein IgG-Immunglobulin, ein IgE-Immunglobulin oder ein IgG_4-Immunglobulin ist.

27. Verfahren zur Anwendung beim Verfolgen des Verlaufs einer Endometriose in einem Patienten, welches die folgenden Schritte umfasst:

(a) Nachweisen erster und zweiter Mengen eines Antikörpers, der spezifisch Repro-EN-1,0-Polypeptid bindet, dessen Aminosäuresequenz identisch zu der von SEQ ID Nr. 2 oder einer Allel-Variante von SEQ ID Nr. 2 ist, oder ein IB1-Polypeptid, dessen Aminosäuresequenz identisch zu der von SEQ ID Nr. 4 oder einer Allel-Variante von SEQ ID Nr. 4 ist, in Proben von dem Patienten von einem ersten bzw. einem zweiten Zeitpunkt; und

(b) Vergleichen der ersten und zweiten Mengen;

wobei ein Anstieg zwischen den ersten und zweiten Mengen ein Fortschreiten der Endometriose indiziert und eine Abnahme zwischen den ersten und zweiten Mengen eine vorübergehende Besserung der Endometriose indiziert.

28. Screening-Methode zum Bestimmen, ob eine Verbindung die Expression von Repro-EN-1,0 der SEQ ID Nr. 2 oder IB1 der SEQ ID Nr. 4 oder einer Allel-Variante davon in einer Zelle vermehrt oder vermindert, umfassend das Zusammenbringen der Zelle mit der Verbindung und das Bestimmen, ob die Produktion von Repro-EN-1,0- oder IN1-mRNA oder -Polypeptid vermehrt oder vermindert ist.

29. Verfahren zum Detektieren polymorpher Formen von Repro-EN-1,0 der SEQ ID Nr. 1 oder 2 oder von IB1 der SEQ ID Nr. 3 oder 4, welches die folgenden Schritte umfasst:

(a) Bestimmen der Identität eines Nukleotids oder Aminosäure an einer ausgewählten Position innerhalb der Sequenz eines Test-Repro-EN-1,0-Gens oder - Polypeptids oder eines Test-IB1-Gens oder -Polypeptids; und

(b) Bestimmen der Identität des Nukleotids oder Aminosäure an einer entsprechenden Position des nativen Repro-EN-1,0-Gens oder -Polypeptids der SEQ ID Nr. 1 oder 2 oder des nativen IB1-Gens oder -Polypeptids der SEQ ID Nr. 3 oder 4; und

(c) Vergleichen der Identität des Test-Gens oder -Polypeptids mit der Identität des nativen Gens oder Polypeptids, wobei ein Unterschied in der Identität anzeigt, dass das Test-Gens oder-polypeptids eine polymorphe Form von Repro-EN-1,0 oder IB1 ist.

Revendications

1. Polynucléotide recombinant comprenant une séquence nucléotidique codant pour un polypeptide Repro-EN-1.0 de SEQ ID NO : 2, le polypeptide se liant spécifiquement à des anticorps de sujets chez qui une endométriose est diagnostiquée.

2. Polynucléotide de la revendication 1 dans lequel la séquence nucléotidique est choisie à partir de la séquence Repro-EN-1.0 de SEQ ID NO :1.

3. Polynucléotide de la revendication 1 dans lequel la séquence nucléotidique est une séquence nucléotidique de Repro-EN-1.0 native.

4. Polynucléotide de la revendication 1 dans lequel la séquence nucléotidique est identique aux nucléotides 176 à 2755 de SEQ ID NO :1.

5. Polynucléotide recombinant comprenant une séquence nucléotidique codant pour un polypeptide IB1 de SEQ ID NO : 4, le polypeptide se liant spécifiquement à des anticorps de sujets chez qui une endométriose est diagnostiquée.

6. Polynucléotide de la revendication 5 dans lequel la séquence nucléotidique est choisie à partir de la séquence IB1 de SEQ ID NO : 3.

7. Polynucléotide de la revendication 5 dans lequel la séquence nucléotidique est une séquence nucléotidique IB1 native.

8. Polynucléotide de la revendication 5 dans lequel la séquence nucléotidique est identique aux nucléotides 176 à 2986 de SEQ ID NO : 3.

9. Polynucléotide de l'une quelconque des revendications 1 à 8 comprenant en outre une séquence de contrôle de l'expression liée de manière fonctionnelle à la séquence nucléotidique.

10. Paire d'amorces polynucléotidiques qui amplifie une séquence nucléotidique codant pour un polypeptide Repro-EN-1.0 de SEQ ID NO : 2 ou un polypeptide IB1 de SEQ ID NO : 4, le polypeptide se liant spécifiquement à des anticorps de sujets chez qui une endométriose est diagnostiquée, la paire comprenant :

    (1) une amorce 3' d'au moins 7 nucléotides qui s'hybride spécifiquement avec l'extrémité 3' de la séquence nucléotidique ou en aval d'une séquence, et
    (2) une amorce 5' d'au moins 7 nucléotides qui s'hybride spécifiquement avec l'extrémité 3' du complément de la séquence nucléotidique ou en aval du complément de la séquence.

11. Paire d'amorces polynucléotidiques de la revendication 10 dans laquelle l'amorce 3' a une séquence complémentaire d'une séquence nucléotidique choisie à partir de l'ADN_c Repro-EN-1.0 de SEQ ID NO :1 et l'amorce 5' a une séquence identique à la séquence nucléotidique choisie à partir de l'ADN_c Repro-EN-1.0 de SEQ ID NO :1.

12. Paire d'amorces polynucléotidiques de la revendication 10 dans laquelle l'amorce 3' a une séquence complémentaire de la séquence nucléotidique choisie à partir de l'ADN_c IB1 de SEQ ID NO :3 et l'amorce 5' a une séquence identique à la séquence nucléotidique choisie à partir de l'ADN_c IB1 de SEQ ID NO :3.

13. Paire d'amorces polynucléotidiques de la revendication 10 dans laquelle la paire de polynucléotides est constituée d'acides nucléiques peptidiques.

14. Cellule recombinante comprenant un polynucléotide recombinant selon la revendication 9.

15. Polypeptide Repro-EN-1.0 recombinant purifié dont la séquence d'acides aminés est identique à celle de SEQ ID NO : 2, ou un variant allélique de SEQ ID NO : 2.

16. Polypeptide IB1 recombinant purifié dont la séquence d'acides aminés est identique à celle de SEQ ID NO : 4, ou un variant allélique de SEQ ID NO : 4.

17. Polypeptide purifié comprenant Repro-EN-1.0 de SEQ ID NO : 2 ou IB1 de SEQ ID NO : 4, le polypeptide se liant spécifiquement à des anticorps de sujets chez qui une endométriose est diagnostiquée.

18. Composition constituée essentiellement d'un anticorps qui se lie spécifiquement au polypeptide Repro-EN-1.0 de SEQ ID NO : 2 ou un anticorps qui se lie spécifiquement au polypeptide IB1 de SEQ ID NO : 4.

19. Composition de la revendication 18 dans laquelle l'anticorps est monoclonal.

20. Méthode de diagnostic d'une endométriose chez un sujet comprenant les étapes consistant à :

(a) détecter une quantité test d'un anticorps qui se lie spécifiquement au polypeptide Repro-EN-1.0 dont la séquence d'acides aminés est identique à celle de SEQ ID NO : 2 ou un variant allélique de SEQ ID NO : 2 ou un polypeptide IB1 dont la séquence d'acides aminés est identique à celle de SEQ ID NO : 4 ou un variant allélique de SEQ ID NO : 4 dans un échantillon du sujet ; et

(b) comparer la quantité test avec une gamme normale de l'anticorps dans un échantillon témoin chez un sujet qui ne souffre pas d'endométriose,

dans laquelle une quantité test au-dessus de la gamme normale fournit une indication positive dans le diagnostic de l'endométriose.

21. Méthode de la revendication 20 dans laquelle l'étape de détection comprend le fait de capturer l'anticorps à partir de l'échantillon avec un Repro-EN-1.0 ou IB1 immobilisé et de détecter l'anticorps capturé.

22. Méthode de la revendication 20 dans laquelle l'étape de détection comprend le fait de capturer l'anticorps à partir de l'échantillon avec un anticorps anti-immunoglobuline immobilisé et de détecter l'anticorps capturé.

23. Méthode de la revendication 21 dans laquelle l'étape de détection de l'anticorps capturé comprend le fait de mettre en contact l'anticorps capturé avec un anticorps détectable qui se lie spécifiquement à des immunoglobulines et de détecter la liaison entre l'anticorps capturé et l'anticorps détectable.

24. Méthode de la revendication 22 dans laquelle l'étape de détection de l'anticorps capturé comprend le fait de mettre en contact l'anticorps capturé avec Repro-EN-1.0 ou IB1 et de détecter la liaison entre l'anticorps capturé et Repro-EN-1.0 ou IB1.

25. Méthode de l'une quelconque des revendications 20 à 24 dans laquelle l'échantillon comprend du sérum sanguin.

26. Méthode de l'une quelconque des revendications 20 à 25 dans laquelle l'anticorps est une immunoglobuline IgG, une immunoglobuline IgE ou une immunoglobuline IgG4.

27. Méthode utilisable pour suivre l'évolution d'une endométriose chez un sujet comprenant les étapes consistant à :

(a) détecter des première et seconde quantités d'un anticorps qui se lie spécifiquement à un polypeptide Repro-EN-1.0 dont la séquence d'acides aminés est identique à celle de SEQ ID NO : 2 ou un variant allélique de SEQ ID NO : 2 ou un polypeptide IB1 dont la séquence d'acides aminés est identique à celle de SEQ ID NO : 4 ou un variant allélique de SEQ ID NO : 4 dans des échantillons du sujet une première et une deuxième fois respectivement ; et
(b) comparer les première et seconde quantités,

dans laquelle une augmentation entre les première et seconde quantités indique une progression de l'endométriose et une diminution entre les première et seconde quantités indique une rémission de l'endométriose.

28. Méthode de criblage pour déterminer si un composé augmente ou diminue l'expression de Repro-EN-1.0 de SEQ ID NO : 2 ou IB1 de SEQ ID NO : 4 ou un variant allélique de ceux-ci dans une cellule, comprenant le fait de mettre en contact la cellule avec le composé et de déterminer si la production d'ARNm ou de polypeptide Repro-EN-1.0 ou IB1 augmente ou diminue.

29. Méthode de détection de formes polymorphes de Repro-EN-1.0 de SEQ ID NO : 1 ou 2 ou de IB1 de SEQ ID NO :

3 ou 4 comprenant les étapes consistant à :

(a) déterminer l'identité d'un nucléotide ou acide aminé en une position choisie dans la séquence d'un gène ou polypeptide test Repro-EN-1.0 ou d'un gène ou polypeptide test IB1 ;
(b) déterminer l'identité du nucléotide ou acide aminé à la position correspondante d'un gène ou polypeptide Repro-EN-1.0 natif de SEQ ID NO : 1 ou 2 ou d'un gène ou polypeptide IB1 natif de SEQ ID NO : 3 ou 4 ; et
(c) comparer l'identité du gène ou polypeptide test avec l'identité du gène ou polypeptide natif, ce par quoi une différence dans l'identité indique que le gène ou polypeptide test est une forme polymorphe de Repro-EN-1.0 ou IB1.

FIG. 1.

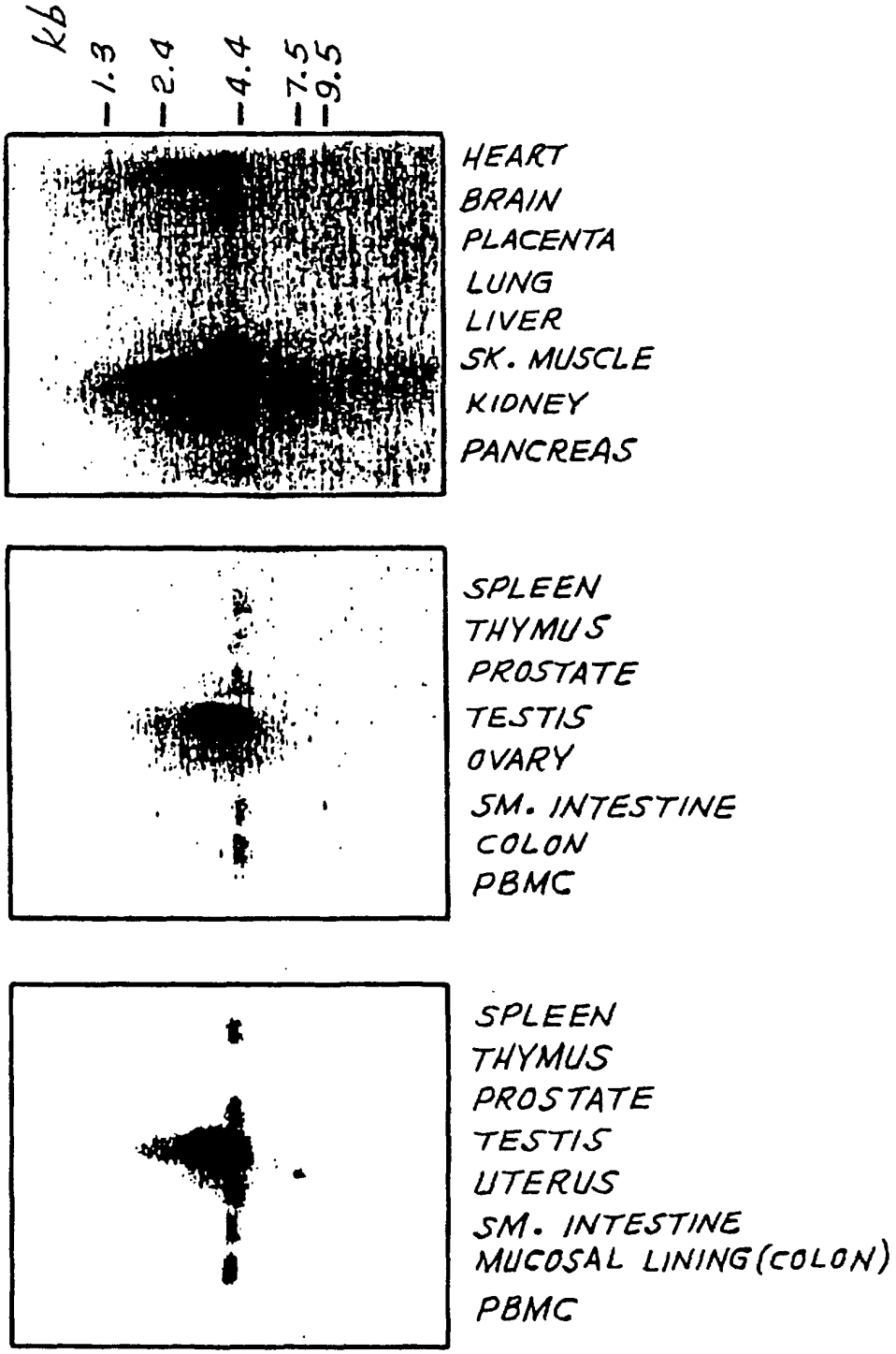

kb
1.3
2.4
4.4
7.5
9.5
HEART
BRAIN
PLACENTA
LUNG
LIVER
SK. MUSCLE
KIDNEY
PANCREAS
SPLEEN
THYMUS
PROSTATE
TESTIS
OVARY
SM. INTESTINE
COLON
PBMC
SPLEEN
THYMUS
PROSTATE
TESTIS
UTERUS
SM. INTESTINE
MUCOSAL LINING (COLON)
PBMC

FIG. 2.

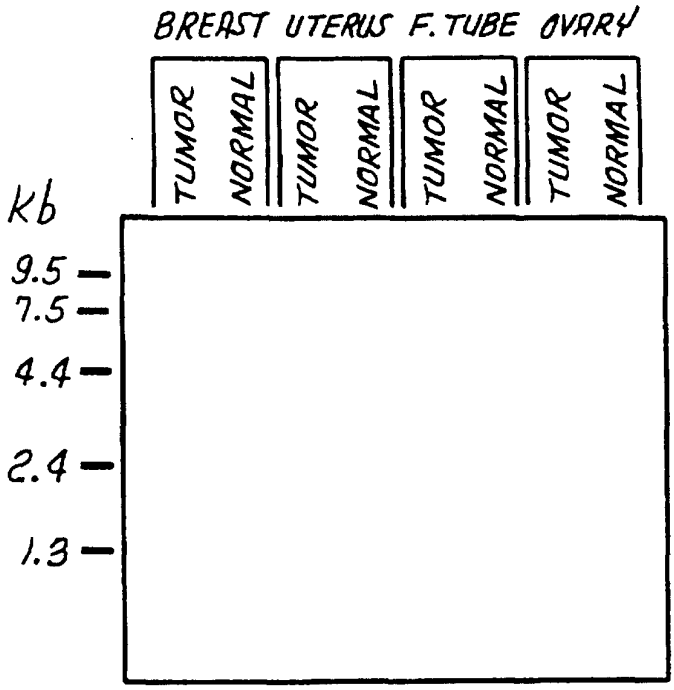
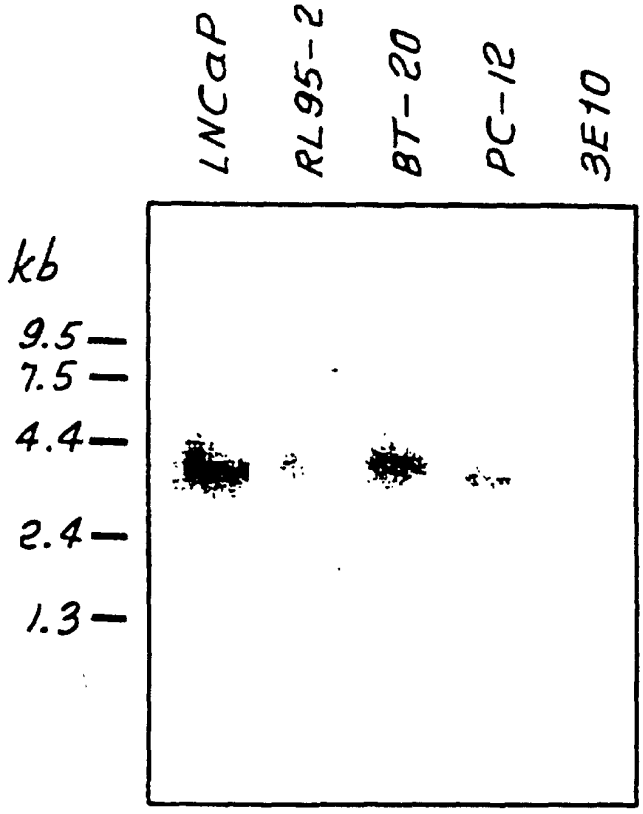

BREAST
UTERUS
F. TUBE
OVARY
TUMOR
NORMAL
TUMOR
NORMAL
TUMOR
NORMAL
TUMOR
NORMAL
kb
9.5
7.5
4.4
2.4
1.3

FIG. 3.



专利名称(译)	编码与子宫内膜异位症相关的自身抗原的多核苷酸		
公开(公告)号	EP1106690B1	公开(公告)日	2005-06-01
申请号	EP2000310408	申请日	2000-11-23
[标]申请(专利权)人(译)	诊断产品		
申请(专利权)人(译)	诊断产品有限公司		
当前申请(专利权)人(译)	诊断产品有限公司		
[标]发明人	EL SHAMI A SAID MENON SURENDRA NATH FRENCH CYNTHIA K		
发明人	EL SHAMI, A. SAID MENON, SURENDRA NATH FRENCH, CYNTHIA K.		
IPC分类号	A61K38/00 C07K14/47 C12N15/12 A61K38/17 C07K16/18 C07K19/00 C12N5/10 C12Q1/68 G01N33/53 G01N33/557		
CPC分类号	C07K14/4713 A61K38/00 G01N2800/364		
优先权	09/447399 1999-11-23 US		
其他公开文献	EP1106690A3 EP1106690A2		
外部链接	Espacenet		

摘要(译)

本发明提供编码Repro-EN-1.0和IB1的多核苷酸，与子宫内膜异位症相关的多肽。在诊断患有子宫内膜异位症的受试者中发现了针对Repro-EN-1.0和IB1的自身抗体。本发明还提供了使用该多核苷酸和多肽的方法。