

METHODS AND PROBES FOR DIAGNOSING A GYNAECOLOGICAL CONDITION

FIELD OF THE INVENTION

[0001] The present invention relates to DNA sequences useful in the diagnosis of gynaecological conditions such as endometriosis. More specifically, the present invention relates to probes and methods the utilisation of the differential expression of certain genes in the diagnosis of endometriosis and related conditions.

BACKGROUND OF THE INVENTION

[0002] Pelvic endometriosis is a prevalent disease in women of reproductive age. It has a propensity to run a chronic and recurrent course after treatment, leading to debilitating chronic pelvic pain and infertility. The etiology and pathogenesis of endometriosis is controversial. A long held belief postulates that endometrial cells from retrograde menstruation are the origin of the disease. The molecular and cellular events that lead to the implantation and growth of these ectopic endometrial cells and development of endometriosis in these women have yet to be determined. There appear to be some aberrant expressions of growth factors, angiogenic factors and adhesion molecules, and the deficiency of immune system have been postulated to play important roles in development and progression of endometriosis. Endometriosis is most probably a multifactorial/polygenic disorder involving dysregulation of multiple genes in the ectopic endometrial cells.

[0003] Current methods for diagnosing endometriosis have a number of drawbacks. Diagnosis by physical examination is invasive and may be painful since it is generally performed during early menses, when implants are likely to be largest and most tender. The physician palpates for a fixed, retroverted uterus, adnexal and uterine tenderness, pelvic masses or nodularity along the uterosacral ligaments. A rectovaginal examination is also required to identify uterosacral, cul-de-sac or septal nodules.

[0004] However, most women with endometriosis have normal pelvic findings, and laparoscopy is necessary for definitive diagnosis.

[0005] Pelvic ultrasonography, computed tomography and magnetic resonance imaging are occasionally used to identify individual lesions, but these modalities are not helpful in assessing the extent of endometriosis. Even with direct visualization, diagnosis of endometriosis can be difficult. Lesions appear in multiple guises that are at times difficult to interpret. This diagnostic challenge is compounded by the unreliable correlation between clinical manifestations and surgical findings. A patient who is asymptomatic or has very mild symptoms may have extensive disease, whereas an infertile patient may have very few implants.

[0006] Thus, there exists a need for a more reliable and less invasive method of diagnosing endometriosis. The present invention overcome or alleviates a problem of the prior art by providing a genetic based method of diagnosis.

[0007] The discussion of documents, acts, materials, devices, articles and the like is included in this specification solely for the purpose of providing a context for the present invention. It is not suggested or represented that any or all of these matters formed part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

SUMMARY OF THE INVENTION

[0008] Throughout the description and claims of this specification, the word “comprise” and variations of the word, such as “comprising” and “comprises”, is not intended to exclude other additives, components, integers or steps.

[0009] The subject specification contains nucleotide sequence information prepared using the program PatentIn Version 3.2, presented herein after the bibliography. Each nucleotide sequence is identified in the sequence listing by the numeric indicator <210> followed by the sequence identifier (e.g. <210>1, <210>2, etc). The length, type of sequence (DNA, etc) and source organism for each nucleotide sequence are indicated by information provided in the numeric indicator fields <211>, <212> and <213>, respectively. Nucleotide sequences

referred to in the specification are defined by the information provided in numeric indicator field <400> followed by the sequence identifier (eg. <400>1, <400>2, etc).

[0010] The applicants have identified a number of DNA sequences that are over expressed or under expressed in cells of a patient having a gynaecological condition such as endometriosis, adenomyosis or endometrioma. Accordingly, in one aspect the present invention provides a DNA molecule or genetic product thereof for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the DNA molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, and 45 or equivalents or fragments thereof.

[0011] In another aspect the present invention provides primers for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the primers comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, or equivalents or fragments thereof.

[0012] In another aspect the present invention provides a DNA molecule or genetic product thereof for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the DNA molecule comprising a nucleotide sequence selected from the group consisting of GenBank accession numbers NM_001904, AF126110, NM_000094, NM_001028, NM_007104, NM_021104, X51346, BC006226, NM_006791, NM_013293, AK094591, NM_001533, NM_005839, NM_003011, NM_020306, NM_004450, BC018111, AF317228, NM_001798, NM_005168, NM_003224, NM_005000, AY007096, NM_000196, NM_025233, X84075, BC004275,

NM_018269, AK001814, XM_031397, NM_014179, NM_152350, NM_009280, L27560, AK001278, NM_030939, AK021534, BC014498, BC011980, BC010281, AK026200, AK091133, or equivalents or fragments thereof.

[0013] Some of these sequences have a high identity with genes with known functions such as Plm-2 oncogenes, IGFBP-5, ribosomal protein L41, propsapasin, fibulin-1, DLX5, 11 β hydroxysteroid dehydrogenase type 2, SET, and RhoE.

[0014] In another aspect the present invention provides a method for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition in a subject, the method comprising determining the expression level of a gene comprising a nucleotide sequence described herein in the subject, and comparing the expression level of the gene to the expression level of the same or a similar gene obtained from a reference sample, wherein a positive diagnosis is made if the expression level in the gene is statistically different to that found in the reference sample.

[0015] In another aspect the present invention provides a method for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the method comprising the detection of a mutation in a gene, the mutation capable of producing a protein with a higher or lower biological activity than a protein from a non-mutated gene.

[0016] In yet a further aspect the present invention provides a probe or primer for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition that is capable of hybridising to a DNA molecule or genetic product thereof as described herein.

[0017] Also provided is a kit for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition comprising a probe or primer as described herein.

[0018] Still a further aspect of the present invention provides a method of screening for a compound having efficacy in the treatment or prevention of a gynaecological condition, the method including determining whether the candidate compound is capable of normalizing the expression levels of a gene comprising a DNA molecule as described herein.

[0019] Another aspect of the present invention provides a method of screening for a compound having efficacy in the treatment or prevention of a gynaecological condition, said method including determining whether the candidate compound is capable of acting as an agonist or antagonist to the protein product of a gene comprising a DNA molecule as described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] Figure 1 shows a schematic protocol for cDNA subtractive hybridisation. Subtractive hybridization was performed between 1st-strand cDNA of driver and 2nd-strand cDNA of tester.

[0021] Figure 2 shows subtraction efficiency by depletion of human β -actin and enrichment of the doped gene CAT in the tester cDNA population. The tester cDNA and the eluted cDNA samples were amplified by PCR with β -actin and/or CAT-specific primers before and after subtractions. PCR products were run out on a 1% agarose gel. Cycle 0, 1, 2, and 3 represent the cDNA samples before subtraction and after 1st, 2nd, and 3rd subtraction, respectively. The high level of β -actin cDNA in the eluted solution (lane 6) demonstrates that the subtractive effect, not degradation, reduced the β -actin cDNA level in the tester.

[0022] Figure 3 shows screening procedures and numbers of cDNA clones screened and identified. a, the false positive clones or equally expressed cDNA clones; b, equally expressed or duplicate cDNA clones; c, duplicate clones or cDNA clones containing only human repeat sequence inserts, such as human Alu repeat sequence.

[0023] Figure 4 shows Northern blot and real-time PCR analysis of the cDNA clones selected from the subtractive cDNA libraries, confirming a high agreement between Northern blot hybridization and real-time PCR technique. A, Overexpressed candidate genes in endometriosis vs. the paired uterine endometrium confirmed by Northern hybridization. Lane a: 4 µg of total RNA samples from uterine endometrium; Lane b: 4 µg of total RNA samples from endometriosis; β-actin was used as control for normalization. B, Underexpressed candidate genes in endometriosis vs. the paired uterine endometrium confirmed by Northern hybridization. C, Linear regression analysis of gene expression data determined by real-time PCR and by Northern blot analysis using the same paired RNA samples, as shown in (A) and (B).

[0024] Figure 5 shows analysis of differential expression data of 76 candidate genes in 15 paired tissue samples from human endometriosis and autologous uterine endometrium generated by real-time PCR. A, Scatter plot of expression data of 76 genes in 15 paired samples after log₂-transformation. The red line represents that $y = 2x$, and the green line represents that $y = 0.5x$. B, Hierarchical clustering of gene expression data of 76 candidate genes (rows) in 15 pairs of clinical samples (columns). Each square represents a ratio of gene expression level in endometriosis to that in uterine endometrium after log₂-transformation. Red, green and black colours represent over-, under- and equal-expression, respectively, of gene in endometriosis vs. the paired uterine endometrium. C, A zoom-in picture for the 14 best candidate genes selected by employing a combination of both the criteria of mean fold-change of 2.0 and $P \leq 0.01$ in 15 cases, showing the consistent patterns of their differential expression. D, Analysis of 15 cases by re-clustering based on the expression data of three immediate-early genes EGR1, JUN and JUND.

DETAILED DESCRIPTION OF THE INVENTION

[0025] Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes V* published by Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew et al (eds.), *The Encyclopedia of Molecular*

Biology, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8). The nomenclature for DNA bases as set forth at 37 CFR .s.1.822 are used herein.

[0026] The applicants have identified a number of DNA sequences that are over expressed or under expressed in cells of a patient having a gynaecological condition such as endometriosis, adenomyosis or endometrioma. Accordingly, in one aspect the present invention provides a DNA molecule or genetic product thereof for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the DNA molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, and 45 or equivalents or fragments thereof.

[0027] It has been shown that the expression of certain genes is altered in a patient suffering endometriosis. The identification of the DNA sequences described herein allows *inter alia* for the development of diagnostic probes, diagnostic methods, and screening assays to identify compounds useful in treating gynaecological disorders.

[0028] The skilled person understands that it is not necessary to use the exact sequences defined by SEQ ID NO: 1 to 45. If limited alterations are made to the DNA sequences described herein, the DNA molecules and genetic products thereof are still able to fulfil their functions in the context of the present invention. For example, it is known that many genes are present in alternate forms (ie alleles) in different individuals. A probe generated against one allele of a gene is very often able to detect a different allele. Accordingly, the present invention includes within its scope DNA molecules "equivalent" to those described in SEQ ID NO: 1 to 45.

[0029] Preferably an equivalent molecule will be at least 90% identical to the sequences described herein. More preferably, an equivalent molecule will be at least 95% identical, and even more preferably it will be at least 97% identical to the sequences described herein.

[0030] Another indication that two nucleic acid molecules are equivalent is that the two molecules hybridize to each other under stringent conditions when one molecule is used as a hybridization probe, and the other is present in a biological sample. Specific hybridization means that the molecules hybridize substantially only to each other and not to other molecules that may be present in the genomic material. Stringent conditions are sequence dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5° C. to 2° C. lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. In a preferred form of the invention an equivalent sequence will hybridise to a sequence described herein under stringent conditions.

[0031] The degeneracy of the genetic code also provides flexibility in the sequences that will be useful in the context of the present invention. Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequence that all encode substantially the same protein.

[0032] The genetic product may be a RNA transcript, or a cDNA derived from the RNA transcript. The genetic product may be a protein. The ability to interchange amino acids in a protein without materially affecting structure or function also provides flexibility in biological systems. For example, substituting a hydrophobic amino acid residue for another hydrophobic amino acid residue will have little effect on the properties of the resultant protein.

[0033] As a result of the degeneracy of the genetic code and the ability for a protein to tolerate conservative substitutions, a gene coding for a protein in one person can be quite different to that in another. However, there will still be a certain percentage identity between allelic forms of a gene in a given species. It is for this reason that the sequences described herein as well as equivalent sequences are useful in the context of the present invention.

[0034] It is also understood that fragments of the sequences described herein will be useful in the context of the present invention. Clearly, it is not necessary to utilize the entire sequence to practice the present invention. To be useful as a probe a sequence may be as little as 10 nucleotides long.

[0035] In a preferred form of the invention the DNA molecule comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 2, 6, 8, 11, 14, 17, 20, 24, 27, 28, 31, 34, 43 and 44 or equivalents or fragments thereof. Applicants have found that genes containing these sequences displayed a particularly high level of dysregulation in subjects suffering from endometriosis.

[0036] After identifying sequences 1 to 45 as relevant to the pathogenesis and diagnosis of gynaecological disorders, the applicants performed further sequence analysis to identify the genes from which SEQ ID NO:1 to 45 originate. In screening the GenBank database, applicants have shown that many of SEQ ID Nos 1 to 45 demonstrate identity with known genes and sequences. The present invention therefore also includes sequences that comprise these genes showing identity to SEQ ID NO 1 to 45. The present invention therefore also provides a DNA molecule or genetic product thereof for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the DNA molecule comprising a nucleotide sequence selected from the group consisting of GenBank accession numbers NM_001904, AF126110, NM_000094, NM_001028, NM_007104, NM_021104, X51346, BC006226, NM_006791, NM_013293, AK094591, NM_001533, NM_005839, NM_003011, NM_020306, NM_004450, BC018111, AF317228, NM_001798, NM_005168, NM_003224, NM_005000, AY007096, NM_000196, NM_025233, X84075, BC004275, NM_018269, AK001814, XM_031397, NM_014179, NM_152350, NM_009280,

L27560, AK001278, NM_030939, AK021534, BC014498, BC011980, BC010281, AK026200, AK091133, ,or equivalents or fragments thereof.

[0037] or example, SEQ ID NO:17 shows identity with the Pim-2 oncogene, which encodes a serine/threonine kinase. It also shows a considerable identity to Pim-1 oncogene, another close-related member of the family. Pim-2 mRNA is highly expressed in various human cancer cells, suggesting the possible involvement of Pim-2 in the transformation processes. Pim-1 and Pim-2 have been identified as common proviral insertion sites in lymphomas induced by MuLV. Pim kinases exhibit strong synergy with c-myc in tumorigenesis, cell proliferation and anti-apoptosis. In human diseases, Pim-1 has recently been shown to be a prognostic biomarker and a co-transcriptional oncogenic factor to myc in prostate cancer. Without wishing to be limited by theory the applicant's finding of over-expression of Pim-2 in over 90% of endometriosis patients provides the molecular basis for the possible synergistic transforming effect to take place in the ectopic tissue, conferring them the survival and progression properties.

[0038] SEQ ID NO:34 shows identity with a mRNA sequence L27560 (GenBank accession number L27560), which was initially misnamed as IGFBP-5. This gene was closely linked to two members of the IGFBP family and was mapped to chromosome 2q35 between the loci of IGFBP-5 and IGFBP-2 genes. Actually, the 5'-end of the reference mRNA sequence is partially overlapped with the 3'-end (exon 4) of IGFBP-5 gene. Besides structural overlapping, IGFBP-5 also exhibited high over-expression in endometriosis, which follows the same pattern as EA30 gene in the 15 cases with a correlation coefficient of 0.86 (data not shown). Their expression regulation appears to be governed by a similar mechanism in endometriosis, indicating that EA30 may be functionally related to IGFBP-5 gene.

[0039] SEQ ID NO:6 has identity with the gene encoding a highly basic ribosomal protein L41, which has been identified as a cellular factor capable of interacting with the CK2 β and regulating the CK2 activity. Protein Ser/Thr kinase

CK2 is one of the key cellular signals for cell survival, growth and proliferation. CK2 participates in a complex series of cellular functions by modulating the stability and/or activity of various important cellular proteins, such as DNA topoisomerase II, p53, NF- κ B, I κ B, β -catenin and proapoptosis protein Bid. It has been demonstrated that CK2 exhibits the elevated expression in various cancers and is implicated in tumorigenesis. RPL41 protein has been shown to stimulate the phosphorylation of DNA topoisomerase II α by CK2 and to enhance the autophosphorylation of CK2 α . Without wishing to be limited by theory, applicants propose that over-expression of RPL41 may lead to the enhanced activity of CK2 in endometriosis. The up-regulation of RPL41 may play an important role in pathogenesis of endometriosis *via* modulating the CK2 activity.

[0040] SEQ ID NO:27 shows identity with Prosaposin, the glycoprotein precursor of saposins A, B, C, and D, which activate lysosomal hydrolysis of sphingolipids. Prosaposin exists in various tissues and body fluids and is especially abundant in the nervous system. Genetic defects in prosaposin have been associated with human lysosomal sphingolipid storage disorders. A recent study on Globoid cell leukodystrophy in a saposin A^{-/-} mouse model demonstrated that pregnancy dramatically alleviated the clinical and pathological phenotype of the affected mice. The estrogen supplementation produced the similar protective effects to pregnancy, indicating that estrogen may, to certain extent, complement the deficiency of saposin A in the mice. Prosaposin was found to be stimulated dose-dependently by estrogen and was secreted by several breast cancer cells. It has been shown that prosaposin mRNA was highly expressed in the adult and embryonic gonads of both male and female mice, suggesting an important role in the reproductive systems. In human breast and ovarian cancer cells, prosaposin interacted with procathepsin D intracellularly and extracellularly, suggesting an involvement in tumor invasiveness and metastasis. In addition, prosaposin treatment of pheochromocytoma cells (PC12) induced extracellular signal-regulated kinases (ERKs) activity, stimulated DNA synthesis, and prevented cell death. Without wishing to be limited by theory, applicants propose that the elevated expression of prosaposin in endometriosis may contribute to the

survival, proliferation and invasiveness of endometrial cells on the ectopic locations.

[0041] SEQ ID NO:2 shows identity with the Fibulin-1 protein which is a multifunctional extracellular matrix protein that strongly interacts with fibronectin and suppresses fibronectin-regulated cell adhesion and motility. Over-expression of fibulin-1 in endometriosis suggests the similar mechanism may be involved in the pathogenesis of this benign and estrogen-dependent disease.

[0042] SEQ ID NO:8 shows identity with the DLX5 gene which encodes a homeodomain transcription factor. It is one of six known human DLX homeobox genes, which are required for neurogenesis, skeleton and appendage development, and bone formation. The functions of different members of the DLX family appear to be partially redundant and the targets of DLX regulation include the DLX genes themselves. It has recently been shown that a homeotic transformation of lower jaws to upper jaws occurred in the *Dlx5/6*^{-/-} mutant mice. Accumulating evidences suggest that DLX5 may play a crucial role in bone development and formation by modulating the gene expression of osteocalcin, COL1A1 and bone sialoprotein. Over-expression of DLX5 in the chicken calvarial osteoblasts stimulated osteoblastic differentiation that does not normally happen *in vitro*. Ectopic expression of DLX5 in rat osteosarcoma cells also induced up-regulation of fibronectin, suggesting an involvement in cell adhesion. The marked down-regulation of DLX5 expression in all 15 endometriosis cases by up to 92 folds, shown in this study, suggests that dysregulated DLX5 gene may be involved in this pathological process and withdrawal of the effect of DLX5 product may facilitate persistence of endometriosis cells in the proliferative state on the ectopic locations by failure of entering end stage differentiation and/or by interference of normal cell adhesion.

[0043] SEQ ID NO:24 shows identity with the HSD11B2 gene which encodes 11 β -hydroxysteroid dehydrogenase type 2 (11 β -HSD2), which inactivates glucocorticoid by converting cortisol to cortisone. It plays an important role in modulating glucocorticoid action within a given tissue by pre-receptor regulation,

in collaboration with 11 β -hydroxysteroid dehydrogenase type 1 (11 β -HSD1). It has been shown that glucocorticoid was able to stimulate expression of P450 aromatase, which converts C19 steroids to estrogens, in human ovarian surface epithelial cell, leiomyoma and adipose tissue. Taken together, these findings indicate that the dysregulated expression of 11 β -HSD2 may play a role in the pathogenesis of endometriosis *via* pre-receptor modulation of glucocorticoid action within the ectopic tissue.

[0044] SEQ ID NO:14 shows identity with the SET gene which encodes the protein I₂^{PP2A}, a potent and specific inhibitor of protein phosphatase 2A (PP2A). PP2A is a major protein Ser/Thr phosphatase expressed in all eukaryotic cells. It is involved in diverse cellular processes including cytokine signaling, transcription and translation, and is an important regulator of cell growth and apoptosis. It has been shown that the activity of phosphoprotein phosphatases including PP2A was required for the cAMP-induced expression of steroidogenic acute regulatory (StAR) protein and steroidogenesis. PP2A and SET have also been identified as posttranslational regulators of androgen biosynthesis by Cytochrome P450c17. In addition, PP2A was a binding partner of estrogen receptor α (ER α) and modulated the estrogen-independent activation of ER α by kinase-mediated phosphorylation. More interestingly, it has been demonstrated that Pim kinases were regulated by PP2A at the posttranslational level and inhibition of PP2A activity by okadaic acid stabilized the Pim proteins. Without wishing to be limited by theory, applicants propose that by modulating the PP2A activity, down-regulated SET expression may positively contribute to the steroidogenesis and modulate the action of steroid hormones in the endometriotic lesion, whereas it may partially alleviate the transforming effect of up-regulation of Pim-2 mRNA (the first candidate gene discussed above) *via* affecting the protein stability in the pathogenesis of endometriosis.

[0045] SEQ ID NO:20 shows identity with the RhoE gene encodes RhoE/Rnd3 protein, a member of the Rho family of Ras-related GTPases, which regulate the organization of the actin cytoskeleton in response to extracellular growth factors. RhoE binds GTP, but has no intrinsic GTPase activity and is found constitutively

in the activated GTP-bound form. Expression of RhoE in mammalian cells inhibits the formation of actin stress fibers, membrane ruffles, and integrin-based focal adhesions, and induces loss of cell-substrate adhesion leading to cell rounding. RhoE may act to inhibit signaling downstream of RhoA and possess a RhoA antagonistic cell function. It has been shown that RhoA protein is over-expressed in breast cancers as compared to the paired normal tissues and the protein level correlates with tumor progression. Without wishing to be limited by theory, applicants propose that down-regulated RhoE contributes to the pathogenesis of endometriosis by altering cell adhesions and/or probably by mimicking the effect of enhanced RhoA protein expression in breast cancers

[0046] Similarities with other sequences have been found, and are detailed in Table 2. The applicants have defined not only cDNA sequences, but also proteins, and classes of proteins that are involved in the pathogenesis of endometriosis. The present invention therefore includes the use of any sequence encoding a protein that exhibits identity with any one of SEQ ID NO: 1 to 45.

[0047] In another aspect the present invention provides a method for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition in a subject, the method comprising determining the expression level of a gene comprising a nucleotide sequence according to claim 1 in the subject, and comparing the expression level of the gene to the expression level of the same or a similar gene obtained from a reference sample, wherein a positive diagnosis is made if the expression level in the gene is statistically different to that found in the reference sample.

[0048] The method may be used to diagnose a new gynaecological condition in a subject, or provide further information on an existing gynaecological condition. In both cases the method can be used to provide prognostic information for use by the clinician to assist in the management of the condition.

[0049] The method could use any one or more of a number of biological samples from the subject as a source for determining the expression level of the gene. In a preferred embodiment of the method the biological sample is tissue biopsy

material from the vagina, uterus, cervix, fallopian tube or ovary. The biological sample may also be a fluid such as blood, saliva, urine, or a secretion of the reproductive tract. Preferably the sample is a biopsy of uterine endometrial tissue.

[0050] Where determination of the expression level involves isolating or purifying a polynucleotide or protein from the biological sample, the skilled artisan will be able to select an appropriate method. As an example, suitable techniques are described by Maniatis, T., Fritsch, E., F. and Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL, 2nd Ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor. N.Y. (1989)

[0051] The expression level of the gene may be determined by any method familiar to the skilled person. One way of determining the expression level of a gene is to quantitate the amount of mRNA that has been transcribed from the relevant gene. This can be achieved by Northern blotting techniques for example. The quantitation could be determined on an absolute scale, or a relative scale.

[0052] The method of diagnosis requires that a judgement as to whether the gene is under expressed, or over expressed by comparison to a reference value for the same or similar gene. The reference value could be obtained from another sample from the subject that would be unlikely to be involved in a gynaecological condition. In a preferred form of the invention the reference value is obtained from eutopic endometrium taken from the subject.

[0053] Alternatively, the reference value could be an expressing level or average of a number of expression levels determined using biological samples taken from one or more subjects not suffering a gynaecological condition.

[0054] In a preferred embodiment of the method a positive diagnosis is made if at least a two-fold difference is noted between the expression level of the test sample and the reference sample. Based on further experience with the diagnostic method the skilled person may arrive at other cut off values. However,

determining more appropriate cut off values is well within the ability of the skilled person.

[0055] In a preferred embodiment of the invention the method considers the expression level of a gene that is over expressed by at least a factor of two (ie SEQ ID NO: 2, 6, 11, 17, 27, 28, 34, 43, or 44). In a further preferred embodiment, the method considers the expression of a gene that is under expressed by a factor of at least two (ie SEQ ID NO: 8, 24, 14, or 20).

[0056] In a preferred form of the invention the expression level is determined by analysis of a gene transcript, such as Northern blotting, quantitative PCR or sequencing.

[0057] In a further preferred form of the invention the expression level is determined by analysis of a protein encoded by the gene. In a more highly preferred embodiment the analysis of the protein is by Western blot, ELISA, surface plasmon resonance, or amino acid sequencing.

[0058] The skilled person will understand that the expression of more than one gene is likely to provide a more definitive picture of the gynaecological condition. Accordingly, in a preferred method the expression levels of at least 5 genes are considered. In a more highly preferred method the expression levels of at least 10 genes are considered.

[0059] Another aspect of the present invention provides a method for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the method comprising the detection of a mutation in a gene, the mutation capable of producing a protein with a higher or lower biological activity than a protein from a non-mutated gene.

[0060] It is possible that while a certain gene involved in a gynaecological condition is not under expressed or over expressed, the protein product has an amino acid sequence that affords it a higher or lower than normal biological activity. Thus, the expression level may appear normal, but a gynaecological

condition will still result because of the abnormally high or low biological activity of the protein.

[0061] In another aspect the present invention provides a probe or primer for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition that is capable of hybridising to a DNA molecule or genetic product thereof as described herein.

[0062] Nucleic acid probes and primers may readily be prepared based on the nucleic acids provided by this invention. A probe may comprise an isolated polynucleotide attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, e.g., in Sambrook et al. (1989).

[0063] Primers are short nucleic acids, preferably DNA oligonucleotides 15 nucleotides or more in length. Primers may be annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, and then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods known in the art.

[0064] Methods for preparing and using probes and primers are described, for example, in Sambrook et al. (1989). PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, Whitehead Institute for Biomedical Research, Cambridge, Mass.). One of skill in the art will appreciate that the specificity of a particular probe or primer increases with its length. Thus, for example, a primer comprising 20 consecutive nucleotides of the gene of interest will anneal to a target sequence with a higher specificity than a corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers may be selected that comprise 20, 25, 30, 35, 40, 50 or more consecutive nucleotides.

[0065] Given the DNA sequences provided in the instant specification, the skilled person is adequately enabled to produce a probe or primer useful in the methods described herein.

[0066] In a preferred form of the invention the primers comprise nucleotide sequences selected from the group consisting of SEQ ID NO: 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 214, 125, 126, 127, 128, 129, 30, 131, 132, 133, 134, 135, or equivalents or fragments thereof.

[0067] Furthermore, a skilled person could deduce the protein that would be translated from the transcript, synthesise the protein (or fragment of the protein), and use it to generate antibody probes specific for that protein.

[0068] The present invention contemplates the use of one or a plurality of sequences. In a preferred embodiment of the invention a plurality of sequences are screened for abnormal expression in the patient. A clinician would be more convinced that a patient was in fact suffering from a gynaecological condition if abnormal expression levels were seen for a number of candidate genes.

[0069] In another aspect the present invention provides a kit for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition comprising a probe or primer as described herein.

[0070] In another aspect the present invention provides a method of screening for a compound having efficacy in the treatment or prevention of a gynaecological condition, the method including determining whether the candidate compound is capable of normalizing the expression levels of a gene comprising a DNA molecule as described herein.

[0071] The present invention is not limited to the diagnosis or prognosis of gynaecological conditions, but will have use in the treatment of such conditions. The sequences identified by the applicants as important in pathogenesis point to a number of genes that are involved in the disease process. This allows further investigations into the identification of compounds that are able to modulate expression of the relevant genes allowing the expression levels to return to normal. For example, if a gene is over expressed in endometriosis, the aim would be to decrease expression. This could be accomplished by gene therapy methods (such as antisense technology). If a gene is under expressed, gene therapy could be used to augment production of the gene product.

[0072] Similarly, it would be possible to identify compounds that are capable of modulating the biological activity of a protein gene product encoded by a gene described herein. Another aspect of the present invention therefore provides a method of screening for a compound having efficacy in the treatment or prevention of a gynaecological condition, said method including determining whether the candidate compound is capable of acting as an agonist or antagonist to the protein product of a gene comprising a DNA molecule as described herein.

[0073] The invention will be applicable to many diseases and conditions of gynaecological origin that relate to the ectopic or uncontrolled growth of uterine tissue. Preferably, the uterine tissue is endometrium.

[0074] The gynaecological condition may be related to a disorder of any pelvic tissue such as uterine, fallopian, ovarian, cervical or vaginal tissue. It will be understood that tissues and other than the reproductive tissues may also be involved. For example, the ectopic growth of uterine tissue may occur in other organs or structures of the pelvis such as the ureters, bladder, urethra, kidney, peritoneum, mesentery, abdominal wall, diaphragm, stomach, liver, pancreas, intestine, rectum, blood vessel, or connective tissue. Conditions affecting these organs or structures are therefore included in the scope of the present invention.

[0075] In a preferred form of the invention the gynaecological conditions to which the present invention applies includes but is not limited to endometriosis, adenomyosis and endometrioma.

[0076] The invention will now be further described in the following non-limiting examples.

EXAMPLES

EXAMPLE 1: Materials and Methods.

[0077] *Tissue specimens.* Tissue biopsies of pelvic endometriosis and normal uterine endometrium were collected in pair from 15 patients undergoing laparoscopy or hysterectomy for endometriosis (Table 1). Portions of the tissue specimens were stored in liquid nitrogen immediately for this study while the remaining were used for histo-pathological diagnosis. The patients either had never received any hormonal treatment or had ceased any hormonal medication for at least 6 months before surgery. The patients were categorized by menstrual phases and severity of endometriosis. The proliferative (days 4-14) and secretory (days 16-28) phases of menstruation was determined by the patients' menstrual history and histo-pathological assessment of the endometrium. The severity or stage of evolution of endometriosis was based on surgical pathological findings according to the revised American Fertility Society classification system (8). For analysis of gene expression profiles, the 15 pairs of samples were classified into two groups: (A) early disease (clinical stage I & II) and (B) advanced disease (clinical stage III & IV).

[0078] *RNA preparation.* The fresh-frozen tissue was weighted, rinsed with ice-cold PBS solution and minced on ice before RNA purification. Total RNA was prepared from 7-50 mg fresh-frozen tissue specimens using TRIzol Reagent according to the protocol recommended by the manufacturer (Gibco BRL).

[0079] *1st-strand and 2nd-strand cDNA synthesis.* mRNA was purified from 30-40 µg of total RNA using Dynabeads Oligo (dT)₂₅ (DYNAL) and 1st-strand cDNA was synthesized (42 °C, 50 min; 50 °C, 30 min) with SuperScript II Reverse

Transcriptase (Gibco BRL, 400 units) using Dynabeads Oligo (dT)₂₅ as RT primer according to the protocol recommended by the manufacturer (DYNAL). All the reactions involving Dynabeads were performed on a roller mixer. After cDNA synthesis, the unprimed Oligo (dT)₂₅ on the Dynabeads were removed by T4 DNA polymerase (New England Biolabs, 3 units, room temperature, 1 hour) (9). 1st-strand cDNA on Dynabeads was tailed with poly(dA)_n by Terminal Deoxynucleotidyl Transferase (Gibco BRL, 15 units, room temperature, 1 hour). A 2nd-strand cDNA was then synthesized with Taq DNA polymerase (Qiagen) using an one-base anchor oligonucleotide HT₂₆V (AAGCTTTTTTTTTTTTTTTTTTTTTTTTTTIV, V: A, G or C) as primer (94°C, 3 min; 55 °C, 2 min and 68 °C, 15 min). The 2nd-strand cDNA was then eluted in 30 µl TE buffer with 0.6 µg carrier tRNA (Gibco BRL) from the 1st-strand cDNA [0080] Dynabeads at 95 °C, 2 min. The 1st-strand cDNA Dynabeads were then removed with a magnet (DynaL MPC-E-1).

[0081] *Subtractive cDNA library construction.* Subtractive hybridization was performed in a small volume ($\leq 30 \mu\text{l}$) based on the solid-phase cDNA using a similar protocol as described by Lönneborg *et al.* (10), with several modifications. Figure 1 shows the schematic protocol for cDNA preparation and subtractive library construction used in this study. Briefly, the main modifications include: (a) 2nd-strand cDNA was used as tester instead of mRNA; (b) subtractive progress was monitored by PCR amplification of β -actin gene in tester solution after each cycle of subtraction. Efficacy of the protocol was evaluated by duplex PCR of β -actin and CAT genes, representing common and differential sequences, respectively, in the differential cDNA model system, which consisted of a driver - 1st-strand cDNA Dynabeads and a tester - 2nd-strand cDNA from the same RNA doped with an extra CAT cDNA. PCR primers used for β -actin and CAT genes were as follows: (a) β -actin forward primer, 5'-ATGGATGATGATATCGCCGC-3'; (b) β -actin reverse primer, 5'-CTAGAAGCATTGCGGTGGA-3'; (c) CAT forward primer, 5'-GACATGGAAGCCATCACAGAC-3'; and (d) CAT reverse primer, 5'-CGACCGTTCAGCTGGATATTAC-3'.

[0082] After subtractive hybridization using cDNA samples prepared from endometriosis and paired uterine endometrium, the remaining 2nd-strand tester cDNA with a poly(dA)_n-tail in the hybridization solution was isolated by using Dynabeads Oligo (dT)₂₅. The subtracted cDNA molecules were globally amplified by PCR using HT₂₆V universal primer. The PCR product was purified with the GFX PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech). The purified PCR product was then cloned into a pCR-XL-TOPO TA vector (Invitrogen), and transformed into TOP10 competent cells to generate a subtractive cDNA library.

[0083] *Screening of subtractive cDNA libraries by differential hybridization.* 738 colonies were random-selected from the subtractive libraries for differential screening with colony and/or dot blot hybridization. About 400 colonies were prescreened by colony hybridization prior to dot blot hybridization. The rest colonies were directly screened by PCR and dot blot hybridization. For dot blot

hybridization, cDNA inserts were amplified by colony-PCR using the vector-specific primers (T7 primer: 5'-TAATACGACTCACTATAGGG-3'; M13 Reverse primer: 5'-GGAAACAGCTATGACCATG-3'). PCR product was mixed with equal volume of 0.6 N NaOH to denature dsDNA. 2 µl of denatured DNA sample was transferred to Hybond-N+ nylon membrane (Amersham Pharmacia Biotech) with a micropipettor. A 1128-bp human β -actin cDNA (coding region) and a vector without insert were used as the positive and negative controls, respectively. The blotted membranes were then neutralized with 0.5 M Tris-HCl (pH 7.5) and cross-linked using an UV Stratalinker (Stratagene). The cDNA probes prepared from paired endometriosis and endometrium samples were labeled with Digoxigenin-11-dUTP (Boehringer Mannheim). Colony and dot blot hybridization with the DIG-labeled probes and their chemiluminescent detection were performed according to the protocol recommended by the manufacturer (Boehringer Mannheim).

[0084] DNA sequencing. DNA sequencing was performed with ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit and ABI Prism 310 Genetic Analyzer (PE Applied Biosystems). Cycle sequencing reaction was carried out with GeneAmp PCR System 9700 (PE Applied Biosystems) using T7 and/or M13 Reverse sequencing primers.

[0085] Northern analysis. Total RNA samples (4-8 µg) were electrophoresed on 1.2% formaldehyde denaturing agarose gel prior to overnight capillary transfer to BM nylon membranes (positively charged, Boehringer Mannheim). The previously isolated cDNA clones were purified with Wizard plus SV Minipreps DNA Purification System (Promega). DIG-DNA probes were then prepared from the purified cDNA clones by PCR using the vector-specific primers. The RNA blots were hybridized with DIG-DNA probes in DIG Easy Hyb solution (Boehringer Mannheim) at 50°C overnight, followed by washes of 2x 5 min in 2x SSC, 0.1% SDS at room temperature and 2x 15 min in 0.1x SSC, 0.1% SDS at 68°C under constant agitation. Chemiluminescent detection of the hybridized blots was performed according to the protocol recommended by the manufacturer (Boehringer Mannheim). Northern blots were quantified with densitometry and the expression data for target genes were then normalized to β -actin mRNA levels for

further analysis. The extent of differential expression of these genes was determined by the fold-change of expression levels in endometriosis versus the paired uterine endometrium after normalization.

[0086] *Gene expression profiling by real-time PCR.* 4 µg of total RNA from each samples were treated with RNase-free DNase I (Gibco BRL) to eliminate the contaminant genomic DNA. The mixtures were directly used for reverse transcription with random hexamer primers and SuperScript First-Strand Synthesis System (Gibco BRL). Real-time PCR was performed with SYBR Green Master Mix on an ABI Prism 7700 Sequence Detection System according to the manufacturer's protocol (PE Applied Biosystems). The cDNAs were amplified by incubation for 10 minutes at 95 °C to activate the Hot Start AmpliTaq Gold DNA polymerase, followed by 45 cycles of denaturation at 94 °C for 30 seconds, annealing at 56 °C for 1 minutes and extension at 68 °C for 1 minutes. 15 ng of cDNA (total RNA equivalent) was used for each PCR reaction performed in 20 µl with Optical 96-Well Reaction Plates (PE Applied Biosystems). Melting curves were generated after amplification to check the PCR specificity. Amplicon size and reaction specificity were further confirmed by electrophoresis on a 1.5% agarose gel and a single PCR product with expected size should be observed. The changes in fluorescence of the SYBR Green I dye in each cycle were monitored by ABI 7700 system, and the threshold cycle (C_T), which is defined as the cycle number at which the amount of amplified target reaches a fixed threshold, was obtained for each gene in each sample. Both primer sets for β -actin and 18S rRNA were included in each plate of PCR reactions as endogenous control and β -actin was used to normalize the quantity of RNA used. The C_T value of β -actin was subtracted from that of each target gene to obtain a ΔC_T value. The relative mRNA level of each target gene in each pair of samples was determined by $2^{-\Delta\Delta C_T}$, where

$$\begin{aligned}\Delta\Delta C_T &= \Delta C_T[\text{Gene}(x)_{\text{endometriosis}}] - \Delta C_T[\text{Gene}(x)_{\text{endometrium}}], \\ \Delta C_T[\text{Gene}(x)_{\text{endometriosis}}] &= C_T[\text{Gene}(x)_{\text{endometriosis}}] - C_T[(\beta\text{-actin})_{\text{endometriosis}}], \\ \text{and } \Delta C_T[\text{Gene}(x)_{\text{endometrium}}] &= C_T[\text{Gene}(x)_{\text{endometrium}}] - C_T[(\beta\text{-actin})_{\text{endometrium}}].\end{aligned}$$

[0087] The C_T values of each gene in endometriosis and the paired endometrium samples were determined in triplicate experiments and the mean C_T values were used for analysis. The real-time PCR reactions were repeated if the coefficient of variation (CV) for any C_T values was more than 4%.

[0088] *Primers used for real-time PCR.* Primers for target genes were designed based on the sequences of the cDNA clones and synthesized by OPERON (USA). The nucleotide sequences for 78 selected cDNA clones (EA01-EA78) and 76 pairs of primers (corresponding to clone EA01-EA76) have been submitted to the GenBank (GenBank accession number: BU197985-BU198062) and will be available at the Web site: <http://www.ncbi.nlm.nih.gov>. The primer sequences for β -actin mRNA and 18S rRNA are as follows: (a) β -actin forward primer: 5'-CCAGCACAATGAAGATCAAGATCA-3'; (b) β -actin reverse primer: 5'-GGGCCGGACTCGTCATACT-3'; (c) 18S rRNA forward primer: 5'-TCGCTACTACCGATTGGATGGT-3'; and (d) 18S rRNA reverse primer: 5'-CACCTACGGAAACCTTGTTACGA-3'.

[0089] *Data analysis.* The differential gene expression measured by the relative quantitative values (fold changes) given by $2^{-\Delta\Delta CT}$ for each target gene in each pair of samples were analyzed and displayed by hierarchical clustering algorithm after logarithmic ($\log_2$) transformation using the Cluster and TreeView software, kindly provided by Dr. Esien (16). The genes and clinical samples were grouped on the basis of similarities of gene expression patterns. Statistical analysis for difference of gene expression data between normal endometrium and endometriotic tissue samples was performed with paired Student's t test. Analyses with a P value of 0.05 or less were considered to be statistically significant.

[0090] *Chromosome mapping.* The candidate genes of interest were mapped to chromosomes by searching human genome, LocusLink and UniGene databases at the Web site: <http://www.ncbi.nlm.nih.gov>.

EXAMPLE 2: Subtractive hybridization, library construction and screening with RNA samples from paired endometrium and endometriosis tissues.

[0091] The efficacy of the subtractive protocol was confirmed experimentally (Figure 2). The subtractive hybridization in two directions to enrich the genes over-expressed or underexpressed in endometriosis tissue generated a large number of cDNA clones. 738 colonies were screened by colony PCR and/or differential hybridization to remove those false positive clones, and duplicate or equally expressed genes. 108 cDNA clones were identified for further study with DNA sequencing (Figure 3).

EXAMPLE 3: DNA sequencing analysis, gene identities and chromosome mapping.

[0092] DNA sequencing analysis through the GenBank databases for 108 cDNA clones isolated from the libraries revealed that 78 contained different cDNA inserts while the other 30 were duplicate clones or had the same inserts with different orientation and were therefore excluded. Of 78 cDNA clones, 48 were identical with or very similar to known genes in the GenBank databases, and 25 were matched to unknown genes, including uncharacterized human or mouse mRNA and hypothetical proteins. The remaining 5 clones matched human genomic sequences with no similarity with any known cDNA or mRNA, i.e. novel genes. The 48 known genes included 5 genes for extracellular matrix/cell adhesion proteins, 8 for ribosomal proteins, 6 for transcription regulators, 6 for RNA processing and pre-RNA splicing factors, 8 for signaling intermediates, 2 for cell cycle, 3 for GDP/GTP binding proteins, 4 for metabolism and 6 for other cellular functions. All these 78 genes were mapped to chromosomes by searching human genome, LocusLink and UniGene databases (Table 2).

EXAMPLE 4: Northern analysis.

[0093] To verify the differential expression of the genes, those cDNA clones generating strong signals during differential screening were analyzed with Northern blot hybridization on the same pair of original RNA samples used for subtractive hybridization. This confirmed the differential expression of 18 genes in endometriosis versus the paired uterine endometrium (Figure 4). Of these 18

genes, 8 were over-expressed (Panel A, Figure 4) and 10 were under-expressed (Panel B, Figure 4) in the endometriotic sample. Notably, there was a marked over-expression in EA26 (EGR1), EA40 (JUN) and EA62 (PIM2) genes and under-expression in EA18 (TSAP19), EA27 (SET), EA35 (CTBP1) and EA61 (RPL13A) genes. Northern blots were quantified with densitometry and the expression data for target genes were then normalized to β -actin mRNA levels (shown in Panel A, Figure 4) for further analysis.

EXAMPLE 5: Gene expression profiling by real-time PCR.

[0094] The CYBR Green-based real-time PCR technique was used for gene expression profiling in multiple clinical samples from patients (Table 1) with endometriosis at different stages (I-IV) after an evaluation study confirming a high agreement between real-time PCR and Northern blot hybridization ($r^2 = 0.986$) (Figure 3C). Some 8,000 real-time PCR reactions were performed to analyze the expression profiles of 76 out of 78 candidate genes in 15 pairs of clinical tissue samples.

[0095] To obtain a general idea of differential regulation of 76 genes in 15 paired samples gene expression data generated by real-time PCR were displayed on a scatter plot (Figure 5A). Of 1140 data points, 649 lay between the red and green lines and therefore their expression ratios in endometriosis samples vs. the paired uterine endometrium are within the range from 0.5 to 2.0. The other 491, representing 43% of the total data points, lying outside this range indicate that the genes were differentially expressed by over 2 fold in these paired clinical samples.

[0096] Two methods were used to identify differentially regulated genes:

(1) Using an arbitrary cutoff criteria. Of the 76 candidate genes, 30 (39%) showed consistent differential expression in more than 70% of cases (at least 11 out of 15). In fact, consistent differential expression was seen in 15 (100%) cases for 2 genes, in 14 (93%) cases for 3 genes, in 13 (87%) cases for 8 genes, in 12 (80%) cases for 9 genes, and in 11 (73%) cases for another 8 genes. Raising the

cutoff threshold to 2-fold difference between endometriosis and autologous endometrium tissues, 14 (18%) of 76 genes were consistently differentially expressed in at least 9 (60%) of 15 cases.

(2) *Using significant statistical difference as criteria.* A quantitative assessment of the differential gene expression in terms of a relative expression ratio allowed the paired *t*-test to be used to test for statistically significant differential expression between endometriosis and the autologous endometrium tissues for 76 candidate genes (table 2). Employing a combination of both the criteria of mean fold-change of 2.0 and statistically significant $\Delta\Delta C_T$ with $P \leq 0.01$ in the 15 cases (Table 1), 14 best candidate genes were identified, including 10 genes over-expressed by 2.0 to 5.6 folds and 4 genes under-expressed by 2.2 to 15.2 folds in endometriosis. 12 of these 14 best candidates concurred with 12 of the 14 genes identified by using the arbitrary 2-fold cutoff method. These 14 candidates were therefore chosen for further investigation.

EXAMPLE 6: Identification of preferred genes.

[0097] The more detailed information including gene identity and chromosome mapping on the 14 best candidates was listed in Table 3. Briefly, 10 over-expressed genes include Pim-2 oncogene (Xp11.23), Ribosomal protein L41 (12q13), Prosaposin (10q22.1), Fibulin 1 (21q13), SIPL protein (2p25.3), three uncharacterized mRNAs: L27560 (2q35), HSPC157 protein (1p36.12) and FLJ37272 (19p13.3), and two novel genes EA08 (12q15) and EA20 (2p21). 4 under-expressed genes include Distal-less homeobox 5 protein (7q21), 11 beta-hydroxysteroid dehydrogenase type 2 (16q22), protein phosphatase 2A-specific inhibitor (SET) (9q34), and ras homolog gene family, member E (RhoE) (2q23.3).

EXAMPLE 7: Clustering of gene expression.

[0098] Cluster analysis was performed to identify the specific gene groups of interest and to display the global picture of differential regulation of 76 candidate genes in 15 pairs of clinical samples. The genes and clinical samples were grouped on the basis of similarities of gene expression patterns by hierarchical clustering algorithm using the Cluster software (16). The result was displayed by

TreeView software (16) and shown in Figure 5B. Based on the gene expression patterns, the 76 candidate genes were generally divided in four different groups, i.e. b1, b2, b3 and b4 (Figure 5B). The group b1 included the genes consistently over-expressed in most endometriosis samples while the genes in the group b3 were consistently under-expressed. The group b2 included most genes whose expression patterns exhibited a transition from over-expression to under-expression in endometriosis across the figure 4B from the left to right. The group b4 included three immediate-early genes and one uncharacterized gene, showing a unique expression pattern that is totally different from those of the other 72 genes.

[0099] A zoom-in picture for the 14 best candidate genes (Table 3) were similarly displayed in Figure 5C to demonstrate how these best candidate genes from the group b1 and b3 (Figure 5B) were differentially regulated in each paired samples. It was clearly demonstrated that the 10 candidate genes on the top part in Figure 5C were consistently over-expressed (red) in the 15 cases while the other 4 genes on the bottom were consistently under-expressed (green).

EXAMPLE 6: Demonstration of Clinical Co-relations.

[0100] As shown in Figure 5B, based on the similarities of gene expression patterns by hierarchical clustering, the majority of cases with early disease (stage I & II) clustered on the left-hand side of the figure, whereas the cases with advanced disease (stage III & IV) aggregated on the right-hand side. It is noteworthy that moving across the figure 5B from the left to right, the expression patterns of the genes in the group b2 exhibited a transition from over-expression to under-expression in endometriosis when compared to the paired endometrium tissue, indicating that some genes may be differentially regulated at different stages of the disease. In order to identify the genes with stage-specific regulation, statistic analyses for 76 candidate genes were performed in two groups of patients, namely, early disease (stage I & II) and advanced disease (stage III & IV), respectively. Employing the same criteria of mean fold-change of 2.0 and statistically significant $\Delta\Delta C_T$ with $P \leq 0.01$, 45 genes were differentially expressed in the paired samples in at least one group of patients. Of these 45 genes, 34

were specific to the early disease, 4 specific for the advanced disease and the remaining 7 genes were differentially expressed in both groups. The detailed information was shown in Table 4.

EXAMPLE 7: Differential regulation of several immediate-early genes.

[0101] As stated above, three of four genes in the group b4 (Figure 5B) are known to be the immediate-early genes, including EA22 (jun-D), EA26 (Egr-1) and EA40 (c-jun). These three genes behaved very similarly in terms of the relative expression ratio in each pair of samples (b4, Figure 5B). But they exhibited very distinct expression patterns in the 15 cases. Based on the expression data of these three genes, the 15 paired samples were re-grouped by clustering and displayed in Figure 5D. It was clearly demonstrated that two opposite expression patterns, d1 and d3, and one transient state, d2, existed in these 15 cases. These three different patterns appear independent of the disease stages as well as the menstrual cycle (Table 1).

[0102] It is to be understood that the embodiments and variations shown and described herein are for illustration only and that various modifications may be implemented by those skilled in the art without departing from the scope and spirit of the invention.

SEQUENCE LISTING

<110> Singhealth
 <120> Methods and Probes for Diagnosing a Gynaecological Condition
 <130> TBA
 <160> 135
 <170> PatentIn version 3.2
 <210> 1
 <211> 463
 <212> DNA
 <213> Homo sapiens

<400> 1
 cgcagtaact gttttttaag tctctcgtag tgttaagtta tagtgaatac tgctacagca 60
 atttctaatt ttttaagaatt gagtaatggt gtagaacact aattaattca taatcactct 120
 aattaattgt aatctgaata aagtgtaca attgtgtagc ctttttgtat aaaatagaca 180
 aatagaaaat ggtccaatta gtttcotttt taatatgctt aaaataagca ggtggatcta 240
 tttcatgttt ttgatcaaaa actatttggg atatgtatgg gtagggtaaa tcagtaagag 300
 gtgttatttg gaaccttggt ttggacagtt taccagttgc cttttatccc aaagttgttg 360
 taacctgctg tgatcagatg cttcaagaga aaatgcggtt ataaaaaatg gttcagaatt 420
 aaacttttaa ttcattcaaa aaaaaaaaaa aaaaaaaaaa aaa 463

<210> 2
 <211> 313
 <212> DNA
 <213> Homo sapiens

<400> 2
 ggccaatcag attttctatg ttctaaggac atggctgctg tagaatagca cagacgtgga 60
 tgataaatta tccccagaag cagcatgaca gaatgcctcg gggagcactt ggaagggaaa 120
 ttgcagttct gttgaaatag aggaaaatcc cttggttaagg acacagcctg ttaggctcgt 180
 gtgggcctcc agtatgttca ccaggggaat ggctgggatt tctcggcact ctgcatcatc 240
 catcttttct tataggtggg aaaataaaca actttgtgat cctcctgaaa aaaaaaaaaa 300
 aaaaaaaaaa aaa 313

<210> 3
 <211> 225
 <212> DNA
 <213> Homo sapiens

<400> 3

tttttttttt tttttttttt ttttttgctt tttaaaacag cagctttaat gccccccaga	60
atcagcacca tgtcatcaca ggcttgggtc aaggggcggg tcagacgcca gtcacatccg	120
ctcactgccc acagccaccc cccacagtg agtcatctgc cagggccgct ccagagtcta	180
tacacagggg tggctctctc aataaagaag tgttgattag acctg	225

<210> 4
 <211> 511
 <212> DNA
 <213> Homo sapiens

<400> 4	
tttttttttt tttttttttt ttttttgatt taataaagtt ttatttttcc aaatgtacag	60
ctgggtggac ctattcatgc atcttcacca gcagctggag catctccacc ottgggtattt	120
ctgggtgtaa ttacttgagc tctgtgcttt gaaaccagtt tgataagtcc ttactaagg	180
agctcctgaa gggctgccct ggccagggag cctcgaatct tcagtctctc agagaccaca	240
gctgggggta taagtttata gttgggaact tccttacaga gtttatcata ggtagctttg	300
tcaaacaaga ctaagttatt gagcttgctc cgaactttgc ctttggacca cttcttcttt	360
ttggccttgc ccccgatttt gttcactggg tctttgtctt tcttggccga ctttccagcg	420
tccttcttct tcttgtcgtc cttaggcggc attgcaagc tcggagaata gcagcagaca	480
ccgcagcctc gtcaagatgt cggacaaaaa g	511

<210> 5
 <211> 501
 <212> DNA
 <213> Homo sapiens

<400> 5	
ctcttttccg gttagcgcg cgtgagaagc catgagcagc aaagtctctc ggcacaccct	60
gtacgaggcg gtgcgggaag tcctgcacgg gaaccagcgc aagcgccgca agttcctgga	120
gacggtggag ttgcagatca gcttgaagaa ctatgatccc cagaaggaca agcgcttctc	180
gggcaccgtc aggccttaagt ccactccccg ccctaagttc tctgtgtgtg tcctggggga	240
ccagcagcac tgtgacgagg ctaaggccgt ggatatcccc cacatggaca tcgaggcgct	300
gaaaaaactc aacaagaata aaaaactggt caagaagctg gccagaagt atgatgcgtt	360
tttggcctca gagtctctga tcaagcagat tccacgaatc ctcggcccag gtttaaataa	420
ggcaggaaag ttcccttccc tgctcacaca caacgaaaac atggtggcca aagtggatga	480
ggtgaagtcc acaatcaagt t	501

<210> 6
 <211> 331

<212> DNA
 <213> Homo sapiens

<400> 6
 tttttttttt tttttttttt ttttttaatt cttcagctaa aacagcggaa gaggtgattt 60
 attatatggt tgttacctc ggccacaaat aaacacagaa atagtccaga atgtcacagg 120
 tccagggcag aggaccaaca tgggcatttt gtttatgagc aaggtgggtc tcagaggtga 180
 tcggcgatca gagggcgatg aagttctaga tccattgaga caagctctag acagtagcat 240
 gcagtccac aacttgtacc agcatcccca gcgtctggca ttccatgttt ctgctcctgt 300
 ggcctccacg gtgcaacaag ctagcggggg g 331

<210> 7
 <211> 320
 <212> DNA
 <213> Homo sapiens

<400> 7
 cttttggaag agagaagaac agagtgttcg attctgccct atttatgttt ctactcggga 60
 acaaacgttg gttgtgtgtg tgtgtgtttt cttgtgttg ttttttaaag aaatgggaag 120
 aagaaaaaaa aattctccgc ccttttctc gatctcgctc ccccttcggt ttctttcgac 180
 cgggtccccc tccctttttt gttctgtttt gttttgtttt gctacgagtc cacattcctg 240
 tttgtaatcc ttggttcgcc cgggtttctg ttttcagtaa agtctcgta cgccaaaaaa 300
 aaaaaaaaaa aaaaaaaaaa 320

<210> 8
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 8
 tttttttttt tttttttttt ttttttcatg gtcgaaataa ttttatttat ccagaatata 60
 cagtttaatt cctctatcta cacttattta catggctaaa ataacattga aaaaagtctt 120
 ttgaaaagtt gaggtcatag atttcaaggc accattgaaa gtgtccacag ttgcgcaaaa 180
 aaagtcctct gtaaaaaaag ggggggtctt ttgaaatgca ataacttaca tgcaaaaaaa 240
 agctttacac atgaatcttt ttcagttttc cgaacttccc catatgaatt cctttcttta 300
 tgattttcta gaacagcaaa acacagtagt ccc 333

<210> 9
 <211> 440
 <212> DNA
 <213> Homo sapiens

<400> 9
 tttttttttt tttttttttt ttttttgaaa ttttaaagaaa aaattttattg aagatctgaa 60
 aaacaattcc taaaagattg acttttccag aaaactagct acacaatgca ttgcgtctat 120
 catgttataaa cgtgcattag acacaaatac aaaaaccatg aaacaagcca ccattcttca 180
 acaatttgag ccaagataaa atgcctaagt aacaacatgg atgacttgca aaggatgggc 240
 tctttacttt aagcaccata aaaaaaaaaac aaaagcacia atggatgagt gtgttcagtt 300
 atatacactg aattgaacct ttggcactag gaatcagagc attttgtcat atagcattaa 360
 catgtattat aaaagtgcgt agtgtcaaag gaatagaacc accagcattc aaaagcagct 420
 ttgtcaacta ggcaataaaa 440

<210> 10
 <211> 369
 <212> DNA
 <213> Homo sapiens

<400> 10
 gggaaagtaa cacatgaagt agttcagtca tgtcaggttt gtctgggggtg gaatggaaca 60
 gtcaggtagt tgaaagtttt ttttttagaga tgaaaagcct gtgaactcct gtaaaacatg 120
 ctgtatttga aatacatctg ttaaaactta aaaactaaag tgtgattttt tttgttggtg 180
 ttaagtttat tgaagtttat tcccttaatc agtgaaggac aacctttatt tattacctag 240
 agcagttgta taattcgtg ttagaaattt tggattggg acagtggata agcaggttat 300
 cacatagcat agaattctta aaagatacat gtggggaaaa gtagtctcaa ataaaagcta 360
 atttctttg 369

<210> 11
 <211> 234
 <212> DNA
 <213> Homo sapiens

<400> 11
 tttttttttt tttttttttt ttttttgaag cttctcacat ctttattgga aaggcacagc 60
 taagcccacc cttgatacag ctttttccac ttctctctgt ggagatcaga cgattgaaca 120
 cgagataagc tgactatata cagacacagg tgtgggtggt gctttttttt aaaaacactt 180
 ttgtcttttt tttttttaat atcccccttc ttaaaagaca agctagtata ctgg 234

<210> 12
 <211> 237
 <212> DNA
 <213> Homo sapiens

<400> 12

tttttttttt tttttttttt ttttttgtct	gggtttattt tccaaacaat tttattgaaa	60
tgtgcccaaga gtacatgggc agcacaatg	tatgaacagg aaaaaaaaa atcacatgta	120
caataatttt ttaaaagtga aggttaatct	tgggagataa cagttcccct ctccctcccc	180
ctgcagattt ccagcgtttc cattaagggt	aagtaagcct ctacaaacct agcattc	237

<210> 13
 <211> 302
 <212> DNA
 <213> Homo sapiens

<400> 13		
cttttgtaaa ggcaatgagc tagtcccaga	aaggatcctt cagttacata caatttgttt	60
aatgaaatgt catggctctg ttcatacttt	tgtcttggtt ttccaattgg tatatacaac	120
tttcagagcc tcttgatatt ggaaggctgg	aagggccac actttggaat agcgtcttgg	180
tttcaactgtt tttgctttga tttttttttt	gttttgattt tttttaaact aaagccatat	240
aaagcttggtg gattaaacag aataaatttc	taaattttaa aaaaaaaaa aaaaaaaaa	300
aa		302

<210> 14
 <211> 410
 <212> DNA
 <213> Homo sapiens

<400> 14		
ccctcttggtg ctcaagtgcg ctgttcttga	ggtctctttt ctctactcca tggttctcaa	60
tttatattggg gggaaatacc ttgagcagaa	tacaatggga aaagagtctc tacccttttc	120
tgttcgaagt tcatttttat cccttcctgt	ctgaacaaaa actgtatgga atcaacacca	180
ccgagctctg tgggaaaaaa gaaaaacctg	ctcccttcgc tctgctggaa gctggagggt	240
gctaggcccc tgtgtagtag tgcatagaat	tctagctttt ttctctcttt ctctgtatat	300
tgggctcaga gagtacactg tgtctctatg	tgaatatgga cagttagcat ttaccaacat	360
gtatctgtct actttctctt gttcaaaaaa	aaaaaaaaa aaaaaaaaa	410

<210> 15
 <211> 220
 <212> DNA
 <213> Homo sapiens

<400> 15		
gaaaagtcac aattttataa aaatggtttt	tcttacattc ttttgagaac tgtttcactc	60
atacatcac ccacacacc cactcaacct	tgtatcaaat tccaaaagt taactaaagt	120
ataagaatat catgactagt taaaagatag	caaataccat aagggtacaag ttcaagtatt	180
agtataacaa gtatctgagt aacaaatgtc	cttggaatg	220

<210> 16
 <211> 223
 <212> DNA
 <213> Homo sapiens

<400> 16
 tttttttttt tttttttttt ttttttgact aaaggacatt tttttatttt taccaagact 60
 ttatcttgag gacatggcta aactgcattt ccacccccca ccccatcttg aagaagggtt 120
 agtctgtgaa tgttataaag tagatatgaa cagttgaagg actggaacca acattaagtg 180
 acgaagaaca acttccatct aaatcatcat aaaaatgttt aag 223

<210> 17
 <211> 397
 <212> DNA
 <213> Homo sapiens

<400> 17
 ggggtgagggg accctactct gttatcccaa gtgctcttat tctggtgaga agaaccttac 60
 ttccataatt tgggaaggaa tggaagatgg acaccaccgg acaccaccag acactaggat 120
 gggatggatg gttttttggg ggatgggctg ggggaaataa ggcttgctgt ttgttctcct 180
 ggggcgctcc ctccaacttt tgcagattct tgcaacctcc tctgagccg ggatcggtcca 240
 attactaaaa tgtaaataat cacgtattgt ggggagggga gttccaagtg tgccctcctc 300
 tcttctcctg cctggattat ttaaaaagcc atgtgtggaa acccactatt taataaaagt 360
 aatagaatca gaaaaaaaaa aaaaaaaaaa aaaaaaa 397

<210> 18
 <211> 319
 <212> DNA
 <213> Homo sapiens

<400> 18
 tccttggaat tgctgcgctc tttgtgtttt tccatcatta gtgcagtttg aatgaatgtg 60
 tataggtcag aggtcttcgt gttcacattt taaaattagg taaatgacct catctttcaa 120
 gcttgaattc atttttaatt ttaattttat ttatataaat gtgtagacag ttccctgttc 180
 tctgcattta gaagtataca caatataaat ctgttaattc tgtaagtaat ttttataatt 240
 atgatataac tctatcttat ccttaaaaaca tttaaaataa accctttatg tgcaaaaaaa 300
 aaaaaaaaaa aaaaaaaaaa 319

<210> 19
 <211> 190
 <212> DNA
 <213> Homo sapiens

<400> 19
aactaggcac atttttttaa tgagctacaa actaaattac aaataacgtt aatagcagga 60
gcactcaagg acaaggggtga cagagacaaa tcataaaaaa aaaatgcttt tttttttagc 120
actgtcatta caaaacccca aggtgcacat ggcaacaaat ccaccacta tgaccaaaaa 180
aaaaaaaaa 190

<210> 20
<211> 261
<212> DNA
<213> Homo sapiens

<400> 20
gttataatct attcatggat ctccactttg atttaatttt taaatgtttt aatctccttt 60
acaaaaagta tacgttaata taccgtcctc aagggggaac tggcactggg accttagcat 120
ttagttttct agaggatgtg atctaatttc tttctagctc atcattaaaa aggaaattgt 180
atcaggaccc atgggatata tccagaggca aactttatga ggctttgaaa tcttgccttc 240
ctgaagatag ctgagtggga g 261

<210> 21
<211> 77
<212> DNA
<213> Homo sapiens

<400> 21
tttttttttt tttttttttt ttttttcaca ctgcccgtct ttattttgca ctcaccctgg 60
gtgacactgg gcaggcc 77

<210> 22
<211> 401
<212> DNA
<213> Homo sapiens

<400> 22
cgaaaacaca tgtaatttta tttttgttta accttaagct tgccaacttc tctccctgaa 60
cagcatttgt cttgttttga taccaccta cacttatatt agaaacgtac tgcaaactat 120
ttagtgactc cacttttaat ttatggtcgt atgactaaga tgttttgaat ataagcctat 180
cggtatgcc a gatataaatt tcagggaactc ttaatgtctc ttcttttaca gaaactctac 240
ctctatcctg atgtaattag gttgattaag aaaattaaag gccacgggcc agaaaactaa 300
cttcacaaa tgcaacaaaa ggaaaatttc tccaataacc tgtatgcaat gggaaatctt 360
tcttttaagt accaagtcaa tttctttaac tgtacaatga t 401

<210> 23

<211> 408
 <212> DNA
 <213> Homo sapiens

<400> 23
 gagacaacgt ctggctctgt caccagggct ggagcgcagt ggcacaatct cggccactg 60
 cagcctcgac ctctggggct caagggatcc tcatgcctca gcctcatatg tagctggggc 120
 cacagacatg caccaccaca cctggctcga ggccatttta gttctgaggt tgagcagatc 180
 aggatccggc tccagcacgg tgctgtgttt gtgaaacaca gaaaggggac ccccgaggac 240
 ccagacagg gccttaggac tctcatatct tacttgtctt ctccatctgg tggctcttgc 300
 tctctggttt tctttacctt ttcatggccc cagaatccat atgcagtaaa ggaactctct 360
 ggaatataat taaagtcttc ccaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 408

<210> 24
 <211> 232
 <212> DNA
 <213> Homo sapiens

<400> 24
 cagtggaaaa taacttttat tgagaccccc tcaggattgt ttctcggcac cagtgcctca 60
 gtgctgcaat tgagggctaa atccaagtg tctcttgact ggctcaagaa ttagggcccc 120
 aactacacac cccaagcca caggaagca tgtactgtac ttccaattg ccacatttta 180
 aataaagaca aatttttatt tcttctaaaa aaaaaaaaaa aaaaaaaaaa aa 232

<210> 25
 <211> 129
 <212> DNA
 <213> Homo sapiens

<400> 25
 tttttttttt tttttttttt ttttttcaga acattcacct ttaattctyl lacaggttcc 60
 agtatttgcg ctggccttag tcagctgtgg gtaagagtca cccatgggca agagtgggga 120
 ggggactgc 129

<210> 26
 <211> 178
 <212> DNA
 <213> Homo sapiens

<400> 26
 acgtcattat tggtcgaact ttcctcacta tctgcttcat ccgccaacta atatttcact 60
 ttacatccaa acatcacttt ggcttcgaag ccgccgctg atactggcat tttgtagatg 120
 tggtttgact atttctgtat gtctccatct attgatgagg gtctcaaaaa aaaaaaaaaa 178

<210> 27
 <211> 127
 <212> DNA
 <213> Homo sapiens

<400> 27
 ccatccaatc cacacccagc agacccttcg gcatgccgcc ctctaccagg aagccagagg 60
 cctaggagct cgccatccat atttatttga aaaggtcaaa aggagcatct atgagacaag 120
 ggagggg 127

<210> 28
 <211> 264
 <212> DNA
 <213> Homo sapiens

<400> 28
 caggtcgttt ttaaagtatt tcttttttta actggatgaa aaatcttcat gttaggatta 60
 attttcttaa tcacctccac actgtacaga ggaaactcaa gccttaaatg tttaagtaaa 120
 ctctgtctca gttttaggat taaaataccc accggtgggtg tgatgatgcc atataccgca 180
 gggettgett ctgtcaagtg tgactctatc tcagtaatta aaataagtgc tgctctacaa 240
 aaaaaaaaaa aaaaaaaaaa aaaa 264

<210> 29
 <211> 391
 <212> DNA
 <213> Homo sapiens

<400> 29
 gagacagcgt ctcaactccct tgcccagggt ggaatgcagt ggcgtgatct tggcccactg 60
 caacctccat ctcccatggt caagtaattc tcctacctca gcctcccaag tagctgggat 120
 tacaggcatg caccaccacg ccagctaatt ttttgtattc ctagtagaaa cagggtttca 180
 ccatgttggc caggettggtc tcgaactcct gacctcaagt gacccgctg cctcggcctc 240
 ccaaagtgct gagattactg gcgtgagcca ccatgcctgg ctgataggaa tatttttgaa 300
 agaccttttc atggtacttt acatcctgaa aggaatggct ttttcttctc tcaggagaac 360
 atgtaagatt aacaagtatt tagactaaag c 391

<210> 30
 <211> 379
 <212> DNA
 <213> Homo sapiens

<400> 30
 gggcttttca taccacacac atatatttcc tatttctcag gaatggtgct gtgtcattac 60
 tatgcacaaa atcggattct tcaagaaaat tgtttttttag gccgagcatg gtggctcacg 120

cctgtaatcc cagcactttg ggaggccatg gcagggtgat cagcagggtca gtagttcgag	180
accagcctgg ccaacatggt gaaatcccat ctctactaaa aatacaaaaa ttagccgagc	240
gtggtggcac acgcgtgtaa tcccagctac tcagaaggct gaggcagaag aatcgcttga	300
acctgggaga cggaggttgc agtgagccga gatcgcacca ttgcaactca gcctgggtga	360
cagggaaga ctccatctc	379

<210> 31
 <211> 196
 <212> DNA
 <213> Homo sapiens

<400> 31	
gcactgtgtg acatgcttta atatatatat tttttatgta tccaaagagg taatcaactt	60
taactttcag aattcatgaa tcttggtttg gtggcatgaa tgaaaggta tgtaacccca	120
ccagtaactc atttgggggtg cccttgaaat attaaagttt ggtcctcaat aaaaaaaaaa	180
aaaaaaaaaa aaaaaa	196

<210> 32
 <211> 473
 <212> DNA
 <213> Homo sapiens

<400> 32	
tttttttttt tttttttttt ttttttcaag aattgacacc caagaccatc ctttattgta	60
gtattagtgc atggttaactg catgaaaaaa catttcagga ggaatttaca atttccagct	120
taaagaactt gccaccaaac ataaccaatt tatgaaagtc aattcattaa aaggtataga	180
acctcttggt gggcatgatg gcaagggaca aagctacaac ttggcctgtg cctttggaag	240
ctgaggcagg aggaccatct gagcccagga gcctgagacc agcctgggca acatagagaa	300
tccgtctcaa caaaaaaaaa attttagcca ggtgtgctgt gagctgtagt ccagctaca	360
agggtgggagg attgcttagg cctgggtgat tgaggatgca atgagctgtg attgtgccac	420
cacactccag cctgggcaat acagcaagac tgtctcaaaa aaaaaaaaaa aaa	473

<210> 33
 <211> 493
 <212> DNA
 <213> Homo sapiens

<400> 33	
tttttttttt tttttttttt ttttttaaaa tcacaaataa tttattcttc cgtttagagt	60
gatacagttt taaatagcaa caaagtacac agtgggtggaa aattggcaca gaacctaca	120
gcatttcaat cacaataagg ctgtccatga tttatctgtt caaacttgat gtccattttc	180

tactgatcaa gttgaaagaa aaaacctatt ctggccaggc acggtggctc acgctgtaa	240
tcccagcact ttggaaggcc gaggggggtg gatcacctga ggtcaggagt tcgagaccag	300
cctggccaac atggcgaaac ctggtctcta ctaaaagtae aaaaattagc tgggtgtggt	360
ggcgggtgcc tgtcatocca gctactcagg aggtgaggc aggagaactg cttgaacctg	420
ggagggggag gttgcagtga gctgagattg cgccaccgac tccagcctgg ggcacaagag	480
caagactcca tcc	493

<210> 34

<211> 207

<212> DNA

<213> Homo sapiens

<400> 34

tttttttttt tttttttttt ttttttttca caaactgggt gcatttgttg atttactttt	60
ccatcctttt caccagtaaa gattatcatc gtttattttt gtttttaaag tcgaaaacaa	120
aaaagaggag agaacagagt tatggtatga atgtatgtaa aactatagag aactacagta	180
atatgtactg tggttattgc tgtcttc	207

<210> 35

<211> 195

<212> DNA

<213> Homo sapiens

<400> 35

tttttttttt tttttttttt ttttttgttg ctcaataaac ttttattggt ttgttattta	60
tttttattca aaagaatttg tttccctatc acaaaaatta gataacacca ccaagagcag	120
caatgtccat ggagttacat tgatgggtga gagttggctg ggtattccca atgtcccaaa	180
ggtcataaag gaggg	195

<210> 36

<211> 373

<212> DNA

<213> Homo sapiens

<400> 36

caaattccaa ccagaagcta aatacaattg gaaactggta agcactagtt ttactccaaa	60
ggagtaggat cattcagatt tactccaata aaagtatgca acccttaagc aaagcttttt	120
ttcattttaa tggagaaaaa aaacaaccta tacagtagtc tttccttatg ttcattgcac	180
aaaatgagtt ctgcttttag aactttgaca ctcaatgggt aattttacaa tttaagattc	240
caactttata accttttttc tactccaaaa cacccttgta aagtttttct ttaggatggt	300

gtaaaaacca gcatttctgc acaattcact ggaatTTTTt tctttgtaat aaaaatctct 360
tctctgtaaa acc 373

<210> 37
<211> 246
<212> DNA
<213> Homo sapiens

<400> 37
gctggccatg agaattagac ttcatattag cttcaattaa aaagctcatt ccattttatt 60
aaaatgtttg taattgcaat tttatgttta atgtttactt tgccttcctt ttcttttagca 120
tttaggtgac tgttcggcag tttgatatat ggccattgtc ggccttaaac tgcactagta 180
acaagcatgg tatTTatctt gtattgaaaa taaaacaaac ttttgagaaa aaaaaaaaaa 240
aaaaaa 246

<210> 38
<211> 507
<212> DNA
<213> Homo sapiens

<400> 38
gagacggagt cttgctctgt tgcccaggca ggagtgcagt ggcgcattct ggctcactgc 60
aacctctgtc tcaagcgatt ctctgcctc agcctcctga gtagctggga ttacagggtgc 120
atgccaccag gccccgctga tttttgtatt tttagtagag atagggtttc accatgttgg 180
tcaggctggc cgtgaactcc tgacctcatg atctgcccgc ctcaacctcc caaagtgtctg 240
ggattatagg cgtgagccac cgcggccagc tgattccagg cgaattcttc tctgatgggc 300
gggcgagggt gtgttcgtgt gatgggttgg agtgtgtgtg tctgagtaca cagatgatgt 360
gttttcctt cagcttctta cgtttctga gcatccgttg tgccttaaca ttttctgctt 420
gtcctttggg acaaagcagt attttactca ttctttgaat gttctcattc ttttgtatca 480
tgtgacttat taaaatcagt ttctaac 507

<210> 39
<211> 217
<212> DNA
<213> Homo sapiens

<400> 39
gggcaggccc atctcagaat taacgctttg atggcatcac cgcgtcggga atccctgggg 60
atggcgcttct ccaccgtcaa gacctttgag ccgcctgagc gactaactcc tgcgccccctg 120
aggtcacagc tacttggaag gctgagatgt gaagactact tgagctcaga agttccactg 180
cactccagcc tgggttacag agggaaacct tgtctct 217

```

<210> 40
<211> 204
<212> DNA
<213> Homo sapiens

<400> 40
cataagaaaa ctacgttaat ttaccttac ttccattgta aataagcctg tcttcctatc      60
tggatTTTTT gtgtgcatac atattctact gattaactac ttttgcagct ttaatcctgt    120
attatttctt ctactttgtt ttgtgtaaaa ggggaaaaaa taaaaaagct ggaatcttca    180
aaaaaaaaaa aaaaaaaaaa aaaa                                           204

<210> 41
<211> 257
<212> DNA
<213> Homo sapiens

<400> 41
tttttttttt tttttttttt ttttttgaag cttttatcaa ctttaattga tataattata      60
attggtatgc aatatgtctt acaacaaatg tggatacaat gacatcaaat accaaagaat    120
tttccacctt ttccatcttt aaatgacttc agctaactta cccaagaga aaaatttggc    180
attgatctct tggttagcat tattaccaat atcaactagc actattaaaa tcaaagttga    240
aatggtacat tcatttg                                           257

<210> 42
<211> 322
<212> DNA
<213> Homo sapiens

<400> 42
gtaattatat gaagaaagtc cagttctcat aaatattgat cacttaaaaa acttactctt      60
tcttgaagag gtacacatgt aaaatttagg aaaatagcta aagtaggggc tggaaccata    120
agaatgttta tccgcacggt catttattat ttggatttg gaacttggct ttgtttttca    180
atagtgacaa gaatggttca gttctaggaa tgttctggaa gatgctgcta attttacttt    240
aaaatgagaa tctggtgtta ctgtatttta tcgttttcaa taaaacttct taagtgtttt    300
ggaaaaaaaa aaaaaaaaaa aa                                           322

<210> 43
<211> 491
<212> DNA
<213> Homo sapiens

<400> 43
gccactgcac tccagcctgg gtgacacagc aagaccttgt ccctaaataa ttaaataaat      60
acatctgtgc tatctcatc tagagctact agtcacatgg ggttactgag cagtggaaat    120

```

gtggctagtc taaattgaga tatgctctac atgcataata tacttcagat tttaaagact	180
tggtatgaaa caaagaatgt aaaatatctc cttaatagtt tttatattga ttataagttg	240
aaatattttg gatacattag attaaataat aaagtgagaa ttaactttac ttgcttttta	300
cctttttaat gtagctaaaa gaaaatgtga aataaaatgt gtggcttatg tttctattaa	360
accatgctgt tttaaataaa aaatccataa tgcctcagtat agatacagtg attaacacaa	420
aagtataggc tgaaagagct taggaatatt tctcactggg aacgtggtgt ataaatttat	480
attaagaata t	491

<210> 44
 <211> 367
 <212> DNA
 <213> Homo sapiens

<400> 44	
cagggatcag atttttatct atctttatcc ccagcacttg atgtggcact tggatatacag	60
taagtactca gtaaagattt ggatggatgg gccgggcaca gtggctcaag cctgtaatcc	120
cagcactatg ggaggccgag gtgggaggat cacaaagtca ggagtttgag accaccctgg	180
ccaatatggt gaaaccctgt ctctactgaa atacaaaaat tagccaggca tggtggtggg	240
cgctgtagt tccagctact ccggaggctg aggcaggaga atcacttgaa cccgagaggc	300
ggaggttgca gtaagccgag attgtgccac cgaactccac ctgggtgaga aagcgagact	360
ccatctc	367

<210> 45
 <211> 342
 <212> DNA
 <213> Homo sapiens

<400> 45	
caaagttgac agataacatt tattaaaaca tgtcatacaa aagggcatga tctcttctat	60
aagaagaaaa tattaaacat taacattcaa ttaagtaaaa ccatgctgca cactgaagac	120
agcaatatat aaagacaaaa ctattcaact tttgcaacat aggcataaca tttcacaaaa	180
tatcaataag atatgcacct aatgataatc tgtttatcat caaatggttt tgtggcctaa	240
gagtcacctg acttgcataa ataaagtact tttctggacg aaacattggg aacctcacia	300
aaggaaacaa ggcagcagca cctaacatga cacaggcttc aa	342

<210> 46
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> synthetic pcr primer

 <400> 46
 gtgtaacaat tgtgtagcct t 21

 <210> 47
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 47
 ggcaactggt aaactgtcca 20

 <210> 48
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 48
 ggccaatcag attttctatg 20

 <210> 49
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 49
 ggaggatcac aaagttgttt 20

 <210> 50
 <211> 18
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 50
 ggtggctgag ggcagtga 18

 <210> 51
 <211> 22
 <212> DNA
 <213> Artificial

<220>
 <223> synthetic pcr primer

 <400> 51
 gcctttttaa acagcagctt ta 22

 <210> 52
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 52
 ctgcggtgtc tgctgctatt 20

 <210> 53
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 53
 tgagcttgct ccgaactttg 20

 <210> 54
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 54
 ttggcctcag agtctctgat 20

 <210> 55
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 55
 tgtggacttc acctcatcca 20

 <210> 56
 <211> 20
 <212> DNA
 <213> Artificial

 <220>

<223> synthetic pcr primer
 <400> 56
 gaacttcac gccctctgat 20

 <210> 57
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer
 <400> 57
 gttgttacac tcggccacaa 20

 <210> 58
 <211> 22
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer
 <400> 58
 gagagaagaa cagagtgttc ga 22

 <210> 59
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer
 <400> 59
 caggaatgtg gactcgtagc 20

 <210> 60
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer
 <400> 60
 ggaagttcgg aaaactgaaa a 21

 <210> 61
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

<400> 61	
tcatagattt caaggcacca t	21
<210> 62	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 62	
gcccatacctt tgcaagtcac	20
<210> 63	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 63	
cacaatgcat tgcgtctatc a	21
<210> 64	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 64	
gggaaagtaa cacatgaagt	20
<210> 65	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 65	
cagcatgttt tacaggagtt	20
<210> 66	
<211> 19	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	

<400> 66 gcaacaccca cacctgtgt	19
<210> 67 <211> 22 <212> DNA <213> Artificial	
<220> <223> synthetic pcr primer	
<400> 67 gcttctcaca tctttattgg aa	22
<210> 68 <211> 20 <212> DNA <213> Artificial	
<220> <223> synthetic pcr primer	
<400> 68 gagtacatgg gcagcacaaa	20
<210> 69 <211> 21 <212> DNA <213> Artificial	
<220> <223> synthetic pcr primer	
<400> 69 ccttaatgga aacgctggaa a	21
<210> 70 <211> 20 <212> DNA <213> Artificial	
<220> <223> synthetic pcr primer	
<400> 70 gcaatgagct agtcccagaa	20
<210> 71 <211> 20 <212> DNA <213> Artificial	
<220> <223> synthetic pcr primer	
<400> 71	

cccttcagc cttccaaata	20
<210> 72	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 72	
tatcccttcc tgtctgaaca a	21
<210> 73	
<211> 19	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 73	
gcctagcacc ctccagctt	19
<210> 74	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 74	
actcatacat acaccacac a	21
<210> 75	
<211> 22	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 75	
acttgaactt gtaccttatg gt	22
<210> 76	
<211> 22	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 76	
cttaatgttg gttccagtcc tt	22

<210>	77	
<211>	22	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	77	
	caagacttta tcttgaggac at	22
<210>	78	
<211>	25	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	78	
	gaaataaggc ttgctgtttg ttctc	25
<210>	79	
<211>	18	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	79	
	ccggctcagg aggaggtt	18
<210>	80	
<211>	20	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	80	
	cttggaatg ctgcgctctt	20
<210>	81	
<211>	22	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	81	
	aagcttgaaa gatgaggtca tt	22

<210>	82	
<211>	20	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	82	
	gggtggattt gttgccatgt	20
<210>	83	
<211>	26	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	83	
	gctacaaact aaattacaaa taacgt	26
<210>	84	
<211>	22	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	84	
	tattcatgga tctccacttt ga	22
<210>	85	
<211>	22	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	85	
	gatcacatcc tctagaaaac ta	22
<210>	86	
<211>	19	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	86	
	gcactcaccc tgggtgaca	19

<210>	87	
<211>	19	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	87	
	gtcccctgct cgggagtaa	19
<210>	88	
<211>	20	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	88	
	ctggcatacc gataggctta	20
<210>	89	
<211>	21	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	89	
	gttttgatac ccacctacac t	21
<210>	90	
<211>	19	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	90	
	gacagggcct taggactct	19
<210>	91	
<211>	21	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	synthetic pcr primer	
<400>	91	
	cagagagttc ctttactgca t	21

<210> 92	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 92	
gattgtttct cggcaccagt	20
<210> 93	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 93	
gtacagtaca tgcttcctg t	21
<210> 94	
<211> 24	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 94	
cagaacattc acctttaatt ctgt	24
<210> 95	
<211> 19	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 95	
gcagtccct cccactct	19
<210> 96	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 96	
ctgcttcac cccaactaa	20

<210> 97
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 97
 ccacatctac aaaatgccag t 21

 <210> 98
 <211> 18
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 98
 cccctccctt gtctcata 18

 <210> 99
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 99
 catccaatcc acacccagca 20

 <210> 100
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 100
 ctccacactg tacagaggaa 20

 <210> 101
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 101
 gagatagaga cacacttgac a 21

 <210> 102

<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 102	
gaaagacctt ttcatgggtca	20
<210> 103	
<211> 24	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 103	
gcttttagtct aaatacttgt taat	24
<210> 104	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 104	
catgctcggc ctaaaaaaca	20
<210> 105	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 105	
gggcttttca taccacacac a	21
<210> 106	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 106	
tgtggacttc acctcatcca	20
<210> 107	
<211> 22	

<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 107	
gggttacata acctttcatt ca	22
<210> 108	
<211> 20	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 108	
caaaggcaca ggccaagttg	20
<210> 109	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 109	
gaccatcctt tattgtagta t	21
<210> 110	
<211> 21	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 110	
ctttcaactt gatcagtaga a	21
<210> 111	
<211> 22	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 111	
cttccgtag agttgataca gt	22
<210> 112	
<211> 21	
<212> DNA	

<213> Artificial
 <220>
 <223> synthetic pcr primer
 <400> 112
 gacagcaata accacagtac a 21

<210> 113
 <211> 21
 <212> DNA
 <213> Artificial
 <220>
 <223> synthetic pcr primer
 <400> 113
 ttccatcctt ttcaccagta a 21

<210> 114
 <211> 20
 <212> DNA
 <213> Artificial
 <220>
 <223> synthetic pcr primer
 <400> 114
 ccctccttta tgacctttgg 20

<210> 115
 <211> 24
 <212> DNA
 <213> Artificial
 <220>
 <223> synthetic pcr primer
 <400> 115
 caaaagaatt tgtttcoccta tcac 24

<210> 116
 <211> 22
 <212> DNA
 <213> Artificial
 <220>
 <223> synthetic pcr primer
 <400> 116
 gtgtcaaagt tctaaaagca ga 22

<210> 117
 <211> 21
 <212> DNA
 <213> Artificial

<220>
 <223> synthetic pcr primer

 <400> 117
 ccaaaggagt aggatcatc a 21

 <210> 118
 <211> 24
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 118
 gcttcaatta aaaagctcat tcca 24

 <210> 119
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 119
 gtgcagttta aggccgacaa 20

 <210> 120
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 120
 gctgattcca ggcgaattct 20

 <210> 121
 <211> 22
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 121
 tcagaaaacg taagaagctg aa 22

 <210> 122
 <211> 19
 <212> DNA
 <213> Artificial

<220>
 <223> synthetic pcr primer

 <400> 122
 gcatcaccgc gtcgggaat 19

 <210> 123
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 123
 gagctcaagt agtcttcaca t 21

 <210> 124
 <211> 25
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 124
 gaaaactacg ttaattttac cttac 25

 <210> 125
 <211> 24
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 125
 caaagtagaa gaaataatac agga 24

 <210> 126
 <211> 22
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 126
 ggtaataatg ctaaccaaga ga 22

 <210> 127
 <211> 22
 <212> DNA

 <213> Artificial

<220>
 <223> synthetic pcr primer

 <400> 127
 ggtagtgaat atgtcttaca ac 22

 <210> 128
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 128
 gggctggaac cataagaaga 20

 <210> 129
 <211> 21
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 129
 cagtaacacc agattctcat t 21

 <210> 130
 <211> 23
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 130
 ccataatgct cagtatagat aca 23

 <210> 131
 <211> 20
 <212> DNA
 <213> Artificial

 <220>
 <223> synthetic pcr primer

 <400> 131
 atacaccacg ttcccagtga 20

 <210> 132
 <211> 21
 <212> DNA
 <213> Artificial

 <220>

<223> synthetic pcr primer	
<400> 132	
cccatccatc caaatcttta c	21
<210> 133	
<211> 23	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 133	
cagggatcag atttttatatt atc	23
<210> 134	
<211> 22	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 134	
catgtcatatc aaaagggcat ga	22
<210> 135	
<211> 22	
<212> DNA	
<213> Artificial	
<220>	
<223> synthetic pcr primer	
<400> 135	
gaaatgttat gcctatgttg ca	22

WE CLAIM:

1. A DNA molecule or genetic product thereof for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the DNA molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, and 45 or equivalents or fragments thereof.
2. A DNA molecule or genetic product thereof, the DNA molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 2, 6, 8, 11, 14, 17, 20, 24, 27, 28, 31, 34, 43 and 44 or equivalents or fragments thereof.
3. A DNA molecule or genetic product thereof for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the DNA molecule comprising a nucleotide sequence selected from the group consisting of GenBank accession numbers NM_001904, AF126110, NM_000094, NM_001028, NM_007104, NM_021104, X51346, BC006226, NM_006791, NM_013293, AK094591, NM_001533, NM_005839, NM_003011, NM_020306, NM_004450, BC018111, AF317228, NM_001798, NM_005168, NM_003224, NM_005000, AY007096, NM_000196, NM_025233, X84075, BC004275, NM_018269, AK001814, XM_031397, NM_014179, NM_152350, NM_009280, L27560, AK001278, NM_030939, AK021534, BC014498, BC011980, BC010281, AK026200, AK091133, , or equivalents or fragments thereof.
4. A method for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition in a subject, the method comprising
determining the expression level of a gene comprising a nucleotide sequence according to claim 1 in the subject, and
comparing the expression level of the gene to the expression level of the same or a similar gene obtained from a reference sample, wherein a positive diagnosis is made if the expression level in the gene is statistically different to that found in the reference sample.

5. A method according to claim 4 wherein the expression level is determined by analysis of a gene transcript.
6. A method according to claim 5 wherein analysis of the transcript is by Northern blot, quantitative PCR or sequencing
7. A method according to claim 4 wherein the expression level is determined by analysis of a protein encoded by the gene.
8. A method according to claim 7 wherein the analysis of the protein is by Western blot, ELISA, surface plasmon resonance, or amino acid sequencing.
9. A method according to claim 4 wherein the expression levels of at least 5 genes are considered.
10. A method according to claim 4 wherein the expression levels of at least 10 genes are considered.
11. A method for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition, the method comprising the detection of a mutation in a gene, the mutation capable of producing a protein with a higher or lower biological activity than a protein from a non-mutated gene.
12. A probe or primer for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition that is capable of hybridising to a DNA molecule or genetic product thereof according to claim 1.
13. A probe or primer according to claim 12 comprising nucleotide sequences selected from the group consisting of SEQ ID NO: 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 214, 125, 126, 127, 128, 129, 30, 131, 132, 133, 134, 135, or equivalents or fragments thereof.

14. A kit for monitoring the progression of, or diagnosing, or determining the severity of a gynaecological condition comprising a probe or primer according to claim 12.

15. A method of screening for a compound having efficacy in the treatment or prevention of a gynaecological condition, the method including determining whether the candidate compound is capable of normalizing the expression levels of a gene comprising a DNA molecule according to claim 1.

16. A method of screening for a compound having efficacy in the treatment or prevention of a gynaecological condition, said method including determining whether the candidate compound is capable of acting as an agonist or antagonist to the protein product of a gene comprising a DNA molecule according to claim 1.

17. A DNA molecule, method, probe, primer or kit according to any one of claims 1, 2, 3, 4, 11, 12, 13, 14, 15 or 16 wherein the gynaecological condition is related to a disorder of uterine, fallopian, ovarian, cervical or vaginal tissue.

18. A DNA molecule, method, probe, primer or kit according to claim 17 wherein the gynaecological condition is related to the uncontrolled growth or ectopic growth of uterine tissue.

19. A DNA molecule, method, or kit according claim 17 wherein the gynaecological condition is selected from the group consisting of endometriosis, adenomyosis or endometrioma.

Table 1. Characteristics of clinical tissue biopsies used for gene expression profiling by real-time PCR

Patient group	Patient No. & symbol	Age (yr)	Cycle phase	Stage of endometriosis
Early disease	1 (I-P1)	33	Proliferative	I
	2 (I-P2)	32	Proliferative	I
	3 (II-P1)	28	Proliferative	II
	4 (II-P2)	32	Proliferative	II
	5 (II-P3)	34	Proliferative	II
	6 (II-S1)	32	Secretory	II
Advanced disease	7 (III-P1)	50	Proliferative	III
	8 (III-P2)	43	Proliferative	III
	9 (III-P3)	35	Proliferative	III
	10 (III-N)	29	Not available	III
	11 (IV-P1)	46	Proliferative	IV
	12 (IV-P2)	41	Proliferative	IV
	13 (IV-S1)	34	Secretory	IV
	14 (IV-S2)	39	Secretory	IV
	15 (IV-S3)	36	Secretory	IV

Table 2. Characteristics and chromosome mapping of cDNA clones identified from subtractive cDNA libraries

cDNA clone	Size (bp)	Accession number	Map locus	Sequence homology	Description	Accession number
Extracellular matrix/cell adhesion proteins						
EA01	463	BU197985	3p21	CTNNB1	β -catenin (cadherin-associated protein)	NM_001904
EA05	396	BU197989	17q21	COL1A1	Collagen, type I, alpha 1	NM_000088
EA29	313	BU198013	21q13	FBLN1	Fibulin-1 isoform D precursor	AF126110
EA52	225	BU198036	3p21.1	COL7A1	Collagen, type VII, alpha 1	NM_000094
EA77	344	BU198061	17q21	COL1A1	Collagen, type I, alpha 1	NM_000088
Ribosomal proteins						
EA07	507	BU197991	8q13.2	RPL7	Ribosomal protein L7	NM_000971
EA09	506	BU197993	15q24	RPS17	Ribosomal protein S17	NM_001021
EA10	500	BU197994	4p13	RPL9	Ribosomal protein L9	NM_000661
EA12	587	BU197996	Xq13.1	RPS4X	Ribosomal protein S4, X-linked	NM_001007
EA15	511	BU197999	11q25	RPS25	Ribosomal protein S25	NM_001028
EA16	501	BU198000	6p21.3	RPL10A	Ribosomal protein L10a	NM_007104
EA19	331	BU198003	12q13	RPL41	Ribosomal protein L41	NM_021104
EA61	630	BU198045	19q13.3	RPL13A	Ribosomal protein L13a (a cell proliferation inhibitor)	NM_012423
Transcription regulators						
EA22	320	BU198006	19p13.1	JUN-D	jun D proto-oncogene	X51346
EA26	527	BU198010	5q31.1	EGR1	Early growth response 1	NM_001964
EA33	333	BU198017	7q21	DLX5	Distal-less homeo box 5	BC006226
EA35	438	BU198019	4p16	CTBP1	C-terminal binding protein 1	XM_042659
EA40	153	BU198024	1p32	JUN	c-jun proto-oncogene	NM_002228
EA63	440	BU198047	15q24	MRG15	MORF-related gene 15	NM_006791
RNA processing and pre-mRNA splicing factors						
EA14	369	BU197998	7p15	HTR2A	Similar to Homo sapiens transformer-2 α	NM_013293
EA39	567	BU198023	12q13	FLJ13467	Similar to hnRNP-E2	AK023529
EA53	234	BU198037	19p13.3	FLJ37272	Highly similar to GRG PROTEIN	AK094591
EA54	237	BU198038	19q13.1	HNRPI	Heterogeneous nuclear ribonucleoprotein L	NM_001533
EA59	302	BU198043	1p36.11	SRRM1	Serine/arginine repetitive matrix 1 (a coactivator of pre-mRNA splicing)	NM_005839
EA76	169	BU198060	1p35	HPRP8BP	U5 snRNP-specific 40 kDa protein	NM_004814
Signaling intermediates						
EA24	531	BU198008	17p13.3	YWHAE	14-3-3 protein, epsilon polypeptide	NM_006761
EA27	410	BU198011	9q34	SET	A heat-stable protein phosphatase 2A-specific inhibitor (I2PP2A)	NM_003011
EA28	292	BU198012	22q13	TOMM22	Translocase of outer mitochondrial membrane 22 homolog (yeast)	NM_020243
EA31	175	BU198015	8q22	YWHAZ	14-3-3 protein, zeta polypeptide	NM_003406
EA36	220	BU198020	2p25	Adam17	Rat disintegrin and metalloproteinase domain 17	NM_020306
EA43	235	BU198027	2p23	FLJ13786	Highly similar to Mus musculus mRNA for ubiquitin conjugating enzyme	AK023848
EA57	223	BU198041	14q24.1	ERH	Enhancer of rudimentary homolog (Drosophila)	NM_004450
EA62	397	BU198046	Xp11.23	PIM2	Pim-2 oncogene	XM_010208
Cell cycle						
EA37	319	BU198021	9q22.1-q22.3	SPIN	Spindlin 1	AF317228
EA50	190	BU198034	12q13	CDK2	Cyclin-dependent kinase 2	NM_001798

GDP/GTP binding proteins						
EA55	316	BU198039	14q21	ARF6	ADP-ribosylation factor 6	BC008918
EA58	261	BU198042	2q23.3	ARHE	Ras homolog gene family, member E	NM_005168
EA70	77	BU198054	20q13.3	ARFRP1	ADP-ribosylation factor related protein 1 (Ras-related GTPase)	NM_003224
Metabolism						
EA06	401	BU197990	7q31	NDUFA5	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5 (13kD, B13)	NM_005000
EA21	408	BU198005	19q13.33	LOC126133	Similar to aldehyde dehydrogenase 1 family, member A2; retinaldehyde dehydrogenase 2	XM_058991
EA60	232	BU198044	16q22	HSD11B2	11-β hydroxysteroid dehydrogenase 2	NM_000196
EA64	129	BU198048	17q21	NBP	Similar to nucleotide binding protein	NM_025233
Other cellular functions						
EA04	178	BU197988	11p11.2	MYBPC	Cardiac myosin binding protein-C	X84075
EA23	159	BU198007	9q22-q31	SEMA4D	Semaphorin 4D	NM_006378
EA34	361	BU198018	1q23	ATP1B1	ATPase, Na ⁺ /K ⁺ transporting, beta 1 polypeptide	NM_001677
EA41	127	BU198025	10q22.1	PSAP	Prosaposin	XM_045140
EA44	264	BU198028	2p25.3	SIPL	A hepatic factor supporting hepatitis C virus replication	NM_018269
EA78	301	BU198062	3p25	IMAGE: 429689	Similar to ARPC4	BC012596
Unknown function						
EA02	605	BU197986	10q21	IMAGE: 4082362	Uncharacterized	BC018658
EA03	330	BU197987	10p15	IMAGE: 4096273	Uncharacterized	BC015987
EA11	391	BU197995	2p22	FLJ10952	Uncharacterized	AK001814
EA13	379	BU197997	1p36.32	KIAA0495	Uncharacterized	XM_031397
EA17	196	BU198001	1p36.12	HSPC157	cDNA clone from CD34 ⁺ stem cells	NM_014179
EA18	473	BU198002	17p11.2	TSAP19	mRNA activated in tumor suppression	AJ012499
EA25	493	BU198009	18q11.2	Ss18	Similar to Mus musculus synovial sarcoma translocation, Chromosome 18	NM_009280
EA30	207	BU198014	2q35	LOC151361	Uncharacterized	XM_098048
EA32	290	BU198016	20q13.33	FLJ20602	Hypothetical protein	AK000609
EA38	195	BU198022	1q23.3	FLJ10416	Uncharacterized	AK001278
EA42	373	BU198026	6p22	FLJ12619	Hypothetical protein	NM_030939
EA45	229	BU198029	3p14	A430107N12	Uncharacterized, from RIKEN Mus full-length enriched library	AK020781
EA46	223	BU198030	7p22.2	IMAGE: 4128750	Uncharacterized	BC027714
EA49	294	BU198033	2q33.1	FLJ13448	Hypothetical protein	BC022453
EA51	246	BU198035	12q14	FLJ11472	Uncharacterized	AK021534
EA56	507	BU198040	11q12	IMAGE: 4856273	Uncharacterized	BC014498
EA65	135	BU198049	5q22	IMAGE: 3867502	Uncharacterized	BC013250
EA66	254	BU198050	1p34.1	SP192	Hypothetical protein SP192	NM_021639
EA67	420	BU198051	5q23	FLJ21409	Uncharacterized	AK025062
EA68	293	BU198052	5q32	HLCDGP1	Down-regulated in lung cancer	NM_018548
EA69	217	BU198053	12p13	IMAGE: 3860421	Uncharacterized	BC011980
EA71	204	BU198055	16p12	IMAGE: 3048642	Uncharacterized	BC010281
EA72	374	BU198056	2q34	KIAA0981	KIAA0981 protein	XM_028867
EA74	257	BU198058	4q21	FLJ22547	Uncharacterized	AK026200

EA75	322	BU198059	22q12	FLJ33814	Uncharacterized	AK091133
Novel genes (match human genomic sequence)						
EA08	491	BU197992	12q15	BAC clone	Novel	AC121761
EA20	367	BU198004	2p21	BAC clone	Novel	AC073082
EA47	222	BU198031	8q23	BAC clone	Novel	AP000427
EA48	288	BU198032	5p13	BAC clone	Novel	AC008768
EA73	342	BU198057	9q34	BAC clone	Novel	AL354944

Table 3. Summary of characteristics of the 14 best candidate genes consistently differentially regulated in 15 paired samples

cDNA clone	Gene symbol (accession No.)	Gene description	Mean fold-change	P value	Map locus	Chromosomal aberrations *
Overexpression in endometriosis						
EA30	LOC151361 XM_098048	Uncharacterized	5.6	0.0003	2q35	Gain of Chr.2 (CGH, 18)
EA62	PIM2 XM_010208	Pim-2 oncogene	3.4	0.0002	Xp11.23	
EA19	RPL41 NM_021104	Ribosomal protein L41	3.2	0.0017	12q13	
EA41	PSAP XM_045140	Prosaposin	3.1	0.0000	10q22.1	Trisomy 10 (R-banding, 20)
EA29	FBLN1 AF126110	Fibulin 1 isoform D precursor	3.0	0.0003	21q13	Gain of 21q (CGH, 18)
EA08	BAC clone AC121761	Novel gene	2.7	0.0000	12q15	
EA17	HSPC157 NM_014179	Uncharacterized cDNA clone (from CD34+ stem cells)	2.4	0.0018	1p36.12	
EA20	BAC clone AC073082	Novel gene	2.2	0.0038	2p21	Gain of Chr.2 (CGH, 18)
EA53	FLJ37272 AK094591	Uncharacterized (highly similar to GRG PROTEIN)	2.2	0.0055	19p13.3	Tetrasomy 19 (R-banding, 20)
EA44	SIPL NM_018269	SIPL protein (a hepatic factor supporting hepatitis C virus replication)	2.0	0.0041	2p25.3	Gain of Chr.2 (CGH, 18)
Underexpression in endometriosis						
EA33	DLX5 BC006226	Distal-less homeobox 5 protein	15.2	0.0000	7q21	
EA60	HSD11B2 NM_000196	11 beta-hydroxysteroid dehydrogenase type 2	3.2	0.0023	16q22	Loss of 16q (CGH, 17); Monosomy 16 (FISH, 19)
EA27	SET NM_003011	A heat-stable protein phosphatase 2A-specific inhibitor (I2PP2A)	2.9	0.0015	9q34	Loss of 9q (CGH, 17)
EA58	ARHE NM_005168	Ras homolog gene family, member E	2.2	0.0100	2q23.3	

* Common chromosome aberrations previously identified in endometriosis tissue or cell line by CGH, FISH and R-binding (references 17-20).

Table 4. Statistic analysis of differential gene expression

cDNA clone	Sequence homology	Mean fold change at stage I & II	P value	Mean fold change at stage III & IV	P value	Mean fold change in all (I-IV)	P value
Overexpression in endometriosis							
EA30	LOC151361	8.7	0.0108	4.2	0.0195	5.6	0.0003
EA62	PIM2	3.9	0.0009	3.1	0.0238	3.4	0.0002
EA19	RPL41	7.7	0.0015	1.8	0.1225	3.2	0.0017
EA41	PSAP	4.9	0.0008	2.3	0.0030	3.1	0.0000
EA29	FBLN1	5.0	0.0055	2.1	0.0205	3.0	0.0003
EA08	Novel	4.2	0.0001	2.0	0.0003	2.7	0.0000
EA17	HSPC157	4.9	0.0003	1.5	0.1549	2.4	0.0018
EA20	Novel	4.0	0.0006	1.5	0.2263	2.2	0.0038
EA53	FLJ37272	4.8	0.0009	1.3	0.3178	2.2	0.0055
EA44	SIPL	3.4	0.0012	1.4	0.2141	2.0	0.0041
EA74	FLJ22547	4.2	0.0040	-1.2	0.7114	1.6	0.1743
EA64	NBP	3.8	0.0008	1.2	0.6200	1.9	0.0283
EA57	ERH	3.7	0.0041	-1.1	0.7171	1.6	0.1605
EA42	FLJ12619	3.4	0.0020	1.2	0.3660	1.8	0.0085
EA75	FLJ33814	3.2	0.0058	1.0	0.9827	1.6	0.0740
EA14	HTR2A	3.2	0.0021	1.1	0.6222	1.7	0.0183
EA38	FLJ10416	3.1	0.0009	1.3	0.0508	1.9	0.0007
EA63	MRG15	3.1	0.0033	1.3	0.2364	1.8	0.0053
EA59	SRRM1	3.1	0.0024	1.1	0.6553	1.7	0.0265
EA21	LOC126133	3.0	0.0035	1.1	0.6367	1.7	0.0242
EA54	HNRPL	2.9	0.0024	1.0	0.9539	1.5	0.0577
EA04	MYBPC	2.8	0.0000	1.4	0.2096	1.8	0.0030
EA50	CDK2	2.8	0.0041	1.3	0.4116	1.8	0.0233
EA06	NDUFA5	2.8	0.0025	1.3	0.1868	1.7	0.0034
EA18	TSAP19	2.8	0.0097	1.0	0.9409	1.5	0.1553
EA73	Novel	2.7	0.0036	1.5	0.1061	1.9	0.0019
EA15	RPS25	2.4	0.0004	1.4	0.1348	1.7	0.0018
EA36	Adam17	2.4	0.0002	1.3	0.1679	1.7	0.0019
EA25	Ss18	2.4	0.0109	1.1	0.6457	1.5	0.0346
EA11	FLJ10952	2.2	0.0093	1.5	0.0107	1.7	0.0003
EA56	IMAGE: 4856273	2.2	0.0126	1.0	0.8324	1.3	0.1317
EA70	ARFRP1	2.2	0.0037	-1.2	0.3232	1.2	0.2734
EA69	IMAGE: 3860421	2.2	0.0043	-1.3	0.0146	1.2	0.3469
EA37	SPIN	2.2	0.0257	-1.5	0.0582	1.1	0.6988
EA51	FLJ11472	2.2	0.0165	-1.5	0.1262	1.1	0.8099
EA01	CTNNB1	2.1	0.0108	1.1	0.6785	1.5	0.0786
EA71	IMAGE: 3048642	2.1	0.0106	-1.3	0.2105	1.1	0.5708
EA13	KIAA0495	2.0	0.0040	-1.1	0.6953	1.3	0.1230
Underexpression in endometriosis							
EA33	DLX5	-3.7	0.0129	-38.7	0.0000	-15.2	0.0000
EA60	HSD11B2	-1.4	0.4760	-5.6	0.0010	-3.2	0.0023
EA27	SET	-3.0	0.0222	-2.9	0.0331	-2.9	0.0015
EA58	ARHE	-1.6	0.2675	-2.7	0.0248	-2.2	0.0100
EA52	COL7A1	2.0	0.0722	-5.4	0.0030	-2.1	0.0912
EA16	RPL10A	-1.3	0.2875	-2.3	0.0092	-1.8	0.0050
EA22	JUN-D	-3.3	0.0319	-1.1	0.8064	-1.7	0.1157