



(51) International Patent Classification:
C12Q 1/6886 (2018.01)

(21) International Application Number:
PCT/US2019/068098

(22) International Filing Date:
20 December 2019 (20.12.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/784,403 22 December 2018 (22.12.2018) US
62/931,173 05 November 2019 (05.11.2019) US

(71) Applicant: **MDNA LIFE SCIENCES INC.** [US/US];
Corporation Trust Centre, 1209 Orange Street, Wilmington,
Delaware 19801 (US).

(72) Inventors: **CREED, Jennifer**; 168-515 Wren Place,
Kelowna, British Columbia V1W 5H7 (CA). **MAGGRAH,**
Andrea; 404 O'Brien St. Box 577, Atikokan, Ontario P0T
1C0 (CA). **REGULY, Brian**; 5921 Matsqui St, Chill-

wack, British Columbia V2R 0G6 (CA). **HARBOTTLE,**
Andrew; 19 Love Lane, Newcastle Upon Tyne NE1
3DW (GB). **USHER, Robert**; 4 Garden Street, Newbottle,
Houghton-Le-Spring DH4 4HF (GB).

(74) Agent: **VORNDRAN, Charles** et al.; Smith Gambrell &
Russell, LLP, 1230 Peachtree St. N.E., Promenade, Suite
3100, Atlanta, Georgia 30309 (US).

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,

(54) Title: MITOCHONDRIAL DNA DELETIONS ASSOCIATED WITH ENDOMETRIOSIS

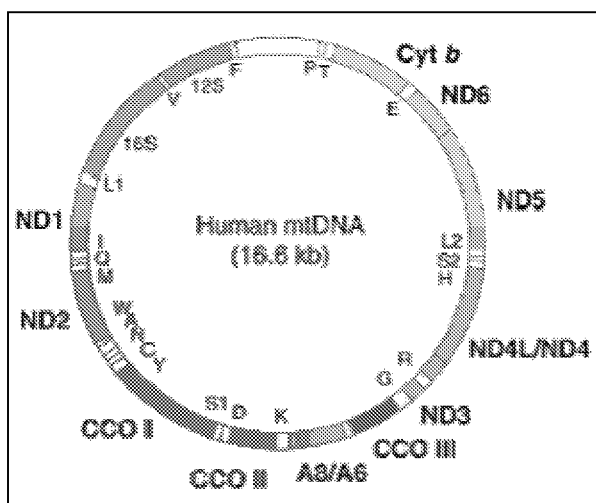


FIG. 1

(57) **Abstract:** Aberrant mitochondrial DNA (mtDNA) molecules having specific large-scale deletions and having an association with endometriosis are provided. The aberrant, or mutated, mtDNA may comprise the parent nucleic acid (i.e. the large sublimon), particularly when re-circularized, wherein adjacent nucleotides are fused following the deletion to form a junction site. Alternatively, the mtDNA may comprise the deleted strand (i.e. the small sublimon), also particularly when re-circularized to create a junction site. In addition, fusion transcripts resulting from such mutated mtDNA, and their putative protein products, are provided, where such transcripts and proteins are also associated with endometriosis. Hybridization probes and amplification primers and kits containing same are provided for detecting, diagnosing, or monitoring endometriosis.



GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

(88) Date of publication of the international search report:

23 July 2020 (23.07.2020)

MITOCHONDRIAL DNA DELETIONS ASSOCIATED WITH ENDOMETRIOSIS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims priority under the Paris Convention to US Application Number 62/784,403, filed December 22, 2018, and US Application Number 62/931,173, filed November 5, 2019. The entire contents of such prior applications are incorporated herein by reference.

STATEMENT REGARDING SEQUENCE LISTING

[0002] A Sequence Listing associated with this application is being filed concurrently herewith in ASCII format and is hereby incorporated by reference into the present specification. The text file containing the Sequence listing is titled "Sequence_Listing.txt", was created on December 17, 2019, and is approximately 119 kilobytes in size.

FIELD OF THE DESCRIPTION

[0003] The present description generally relates to novel biomarkers and methods for detecting/diagnosing and/or monitoring endometriosis. The description also relates to unique analytes and/or reagents that are useful in the subject methods.

BACKGROUND

[0004] Endometriosis is a burdensome disease that occurs in up to 5% to 10% of women of reproductive age and is a common cause of infertility [1–7, 58]. The disease is characterized by the presence of endometrial tissue (epithelial cells and stroma) growing outside of the uterus. Such ectopic endometrial tissue can be found on the pelvic peritoneum and Fallopian tubes, the ovaries, the bowel and bladder, and rarely more distal body sites [8–11]. Women with endometriosis frequently suffer from often debilitating symptoms including non-menstrual pelvic pain, painful menstrual cramps, pain during intercourse, fatigue, and infertility [12], which can lead to a substantial reduction in quality of life [13]. Given its high prevalence and significant morbidity, endometriosis results in a very significant economic cost globally, estimated to be in the hundreds of billions of Euros each year [14].

1 **[0005]** Unfortunately, diagnosis of endometriosis is often a lengthy process, resulting a delay in
2 treatment. The current “gold standard” for diagnosing endometriosis involves laparoscopic
3 surgery followed by histopathological confirmation of tissue samples [5,15]. Making a timely
4 diagnosis is further complicated by delayed reporting [16] and misinterpretation of symptoms
5 [17] and can be even further delayed if the patient is hesitant to undergo a costly and invasive
6 laparoscopic procedure. Indeed, the delay in diagnosing endometriosis can be more than a
7 decade [16]. Due to these delays, a majority of women develop moderate to severe symptoms
8 by the time a definitive diagnosis is made, which can result in increased morbidity, treatment
9 costs, and decreased quality of life [14]. There is a therefore a need for a reliable, non-invasive
10 test that can facilitate early detection of endometriosis and provide actionable real-time results.
11 However, no non-invasive methods for detecting endometriosis are currently available.

12 **[0006]** Molecular biomarkers have been widely used as tools to measure, detect, and predict
13 human disease [18–24]; however, the search for an endometriosis-specific biomarker has
14 proven difficult [25]. Some of the key challenges include non-standardized sample collection,
15 analysis methods, and data interpretation and the lack of biomarker specificity [17] though
16 recent efforts have been made to harmonize methods of collecting and storage of biological
17 specimen and reporting of endometriosis data, including the World Endometriosis Research
18 Foundation (WERF) EPHeC Protocols [26]. A variety of candidate biomarkers from blood,
19 tissue, and urine have been reported, but none have been successfully translated into clinical
20 use. Many of these candidates have specific limitations on sample collection such as biopsy
21 from diseased tissue, requirement for collection during a particular phase of menstruation, or
22 are dependent upon changes in regulatory patterns (e.g., gene expression, DNA methylation)
23 induced by inflammation, which can overlap with other gynaecological disorders [10,17] and
24 increase the likelihood of false positive detections. Thus, an ideal biomarker would be
25 detectable from healthy cells or body fluids and independent of transient disease, inflammation-
26 generated, or cyclical physiological changes.

27 **[0007]** The mitochondrial genome represents a less-explored biomarker repository. As shown
28 in FIG. 1, the mitochondrial genome codes for a complement of 24 genes, including 2 rRNAs
29 and 22 tRNAs that ensure correct translation of the remaining 13 genes which are vital to
30 electron transport. Mitochondrial DNA (mtDNA) targets are attractive from a diagnostic
31 perspective due to a high mutation frequency, limited DNA repair capability, presence in all
32 nucleated cells, and high copy number (thousands of genomes per cell) [27]. As a result, even

low frequency mutations or deletion events can be amplified reliably from heteroplasmic mitochondrial populations. Indeed, mtDNA mutations have been well-described as biomarkers for several cancers across multiple body sites including bone, brain, breast, lung, colorectal, gastric, ovarian, prostate, and endometrial tissues [28–37]. The mitochondrial (mt) genome is relatively small, having 16,569 nucleic acid base pairs, whereas the nuclear genome has over 3 billion base pairs. Furthermore, typically all mtDNA genomes in a given individual are identical given the clonal expansion of mitochondria within the ovum, once fertilization has occurred. The mt genome is also unusual in that it is a circular, intron-less DNA molecule interspersed with repeat motifs that flank specific lengths of sequences. Sequences between these repeats are prone to deletion under circumstances that are not well understood. Moreover, such deletions often include at least a portion of one or both of the flanking repeat sequences. As discussed further below, once the sequence constituting the deletion is removed, the remaining “parent” mtDNA re-circularizes to form a “large sublimon”. Similarly, the deleted sequence may also re-circularize to form a “small sublimon”. Given the number of repeats in the mt genome, there are many possible deletions. One of the best-known examples of these deletions is the 4977 bp “common deletion”, which has been associated with various disease states. Although the common deletion was also investigated as a marker for endometriosis [54], a lack of specificity did not suggest that such deletion would be an effective marker of the disease. Certain mitochondrial DNA deletions have been previously associated with some specific conditions and age-related disorders (see [59] - [64]). An 8686 bp deletion between nucleotides 5371 - 14058 of the mtDNA genome has also been published ([65]), but without any correlation with a disease state or condition.

[0008] In some cases, mtDNA deletions and other large-scale mtDNA rearrangements can result in a mutated mtDNA sequence that can be transcribed, resulting in a mitochondrial fusion transcript. Examples of associations between mitochondrial fusion transcripts and disease states have been described, for example, in the present Applicant’s previous application numbers: PCT/CA2006/000652; PCT/CA2007/001711; PCT/CA2009/000351; and PCT/CA2010/000423, the entire disclosures of which are incorporated herein by reference.

[0009] MtDNA alterations have been detected in the endometrium during investigations of endometrial cancers [37–40]. However, these studies did not reveal a consensus region within the mtDNA genome or a specific mtDNA alteration that correlated to endometrial disease. As a result, these studies did not suggest a conclusion that mtDNA alternations could be used as a

biomarker for detection of endometriosis. Further, no prior investigations are believed to have been conducted in relation to mitochondrial fusion transcripts and endometrial disease or state.

[0010] Thus, there exists a need for an accurate and/or more efficient means of detecting endometrial disease and/or condition that addresses at least one of the deficiencies in the known methods.

SUMMARY OF THE DESCRIPTION

[0011] In one aspect, the present description provides methods, reagents, and/or kits for detecting, diagnosing, and/or monitoring endometriosis in a subject. The description involves the use of mitochondrial DNA (mtDNA) biomarkers, fusion transcripts thereof and/or translated fusion proteins that have been identified herein as being associated with endometriosis. The present methods can be conducted using a biological sample obtained from a subject being screened. Such sample may comprise tissue (such as a biopsy tissue), menstrual fluid, circulatory blood, or blood derivatives such as serum or plasma. The presently described methods can be performed on samples obtained non-invasively from subjects that are suspected of having or developing endometriosis and serve as an effective means of determining whether further invasive diagnostic investigation is necessary.

[0012] In one aspect, there is provided a method of detecting, diagnosing, and/or monitoring endometriosis in a mammalian subject, the method comprising identifying, in a biological sample from the subject, an aberrant mitochondrial DNA, mtDNA, molecule having at least one deletion resulting in a junction point in the rejoined, or re-circularized mtDNA nucleotide sequence, wherein the junction point is at nucleotide pairs 8469:13447, 7992:15730, 9191:12909, 9188:12906, 10367:12829, 6260:12814, 7973:9023, 9086:10313, 9079:14988, 7260:15540, 8431:10841, 8984:13833, or 5362:14049 of the mtDNA nucleotide sequence of SEQ ID NO: 1.

[0013] In one aspect, the method comprises identifying the aberrant mtDNA by contacting a biological sample with DNA probes or primers designed to hybridize to the aberrant mtDNA.

[0014] In one aspect, the method comprises identifying fusion transcripts of the aberrant mtDNA molecule(s).

1 **[0015]** In another aspect, the method comprises identifying fusion proteins encoded by the
2 aberrant mtDNA molecule(s).

3 **[0016]** In one aspect, there is provided a method of identifying, in a biological sample from a
4 mammalian subject, an aberrant mitochondrial DNA, mtDNA, molecule having a deletion,
5 wherein the deletion comprises a nucleotide sequence between nucleotides 5362-14049; 8469-
6 13447; 7992-15730; 9191-12909; 9188-12906; 10367-12829; 6260-12814; 7973-9023; 9086-
7 10313; 9079-14988; 7260-15540; 8431-10841; or 8984-13833 of the mtDNA nucleotide
8 sequence of SEQ ID NO: 1, and wherein, once re-circularized, the mtDNA includes a junction
9 point.

10 **[0017]** In another aspect, there is provided a method of identifying, in a biological sample from a
11 mammalian subject, an aberrant mitochondrial DNA, mtDNA, molecule having a deletion,
12 wherein once re-circularized, the mtDNA includes a junction point consisting of first and second
13 nucleotides, and wherein, with respect to SEQ ID NO: 1:

14 a) the deletion includes nucleotides 5377-14048, the first nucleotide is between
15 nucleotides 5362-5377 and the second nucleotide is between nucleotides 14048-14063;

16 b) the deletion includes nucleotides 8483-13446, the first nucleotide is between
17 nucleotides 8469-8483 and the second nucleotide is between nucleotides 13446-13460;

18 c) the deletion includes nucleotides 7993-15722, the first nucleotide is between
19 nucleotides 7985-7993 and the second nucleotide is between nucleotides 15722-15730;

20 d) the deletion includes nucleotides 9196-12908, the first nucleotide is between
21 nucleotides 9191-9196 and the second nucleotide is between nucleotides 12908-12912;

22 e) the deletion includes nucleotides 9196-12905, the first nucleotide is between
23 nucleotides 9188-9196 and the second nucleotide is between nucleotides 12905-12913;

24 f) the deletion includes nucleotides 10368-12825, the first nucleotide is between
25 nucleotides 10364-10368 and the second nucleotide is between nucleotides 12825-12829;

26 g) the deletion includes nucleotides 6261-12813, the first nucleotide is between
27 nucleotides 6260-6271 and the second nucleotide is between nucleotides 12813-12824;

28 h) the deletion includes nucleotides 7984-9022, the first nucleotide is between
29 nucleotides 7973-7984 and the second nucleotide is between nucleotides 9022-9033;

30 i) the deletion includes nucleotides 9087-10303, the first nucleotide is between
31 nucleotides 9077-9087 and the second nucleotide is between nucleotides 10303-10313;

j) the deletion includes nucleotides 9086-14987, the first nucleotide is between nucleotides 9079-9086 and the second nucleotide is between nucleotides 14987-14904;
k) the deletion includes nucleotides 7261-15531, the first nucleotide is between nucleotides 7252-7261 and the second nucleotide is between nucleotides 15531-15540;
l) the deletion includes nucleotides 8440-10840, the first nucleotide is between nucleotides 8431-8440 and the second nucleotide is between nucleotides 10840-10849; or,
m) the deletion includes nucleotides 8994-13832, the first nucleotide is between nucleotides 8984-8994 and the second nucleotide is between nucleotides 13832-13842.

[0018] In another aspect, there are provided methods of detecting fusion transcripts and fusion proteins resulting from the aberrant mtDNA molecules or from the mtDNA deletions.

BRIEF DESCRIPTION OF THE FIGURES

[0019] The features of certain embodiments will become more apparent in the following detailed description in which reference is made to the appended figures wherein:

[0020] FIG. 1 is an illustration showing mitochondrial coding genes.

[0021] FIGs. 2A to 2J illustrate the detection of fusion transcripts 1, 4, 14, 16, 120, 122, 193, 400, 516, and 586 in endometrial tissues as discussed in Example 1. Scatterplots represent normalized results of endometrial control tissues and endometriosis positive tissues that were tested against probes specific to ten fusion transcripts, identified as transcript numbers: 1 (FIG. 2A); 4 (FIG. 2B); 14 (FIG. 2C); 16 (FIG. 2D); 120 (FIG. 2E); 122 (FIG. 2F); 193 (FIG. 2G); 400 (FIG. 2H); 516 (FIG. 2I); and 586 (FIG. 2J). The y-axis of each figure indicates the normalized Relative Luminescence Units, RLUs, ($\log_2\text{LOQProbe} - \log_2\text{LOQHK23}$) where HK23 is the nuclear housekeeper transcript Human beta-2-microglobulin. The x-axis of each figure indicates tissue diagnosis, as determined by physician's diagnosis upon laparoscopic exam, where: endometrial control = 0.0 and endometriosis positive = 1.0.

[0022] FIG. 3 depicts an mtDNA fusion transcript map illustrating the mtDNA genome of SEQ ID NO: 1, gene locations and locations of the 10 mtDNA deleted portions described herein (i.e., "probes" or "targets"), which are indicated by a line spanning the length of each deletion.

[0023] FIGs. 4A and 4B illustrate the diagnostic accuracy of the 1.2 kb and 3.7 kb Deletions of Example 2, comparing symptomatic control samples and samples from patients with confirmed

1 endometrial disease conditions. The 1.2 kb and 3.7 kb Deletions were evaluated for the ability
2 to distinguish between symptomatic patient specimens and specimens from patients with
3 confirmed endometriosis (all subtypes/stages combined). Receiver operator characteristic
4 curves were constructed and the areas under the curves were calculated. Abbreviations: CI =
5 confidence interval; ROC = receiver operator characteristic; Std = standard; vs = versus.

6 **[0024]** FIGs. 5A to 5D illustrate the diagnostic accuracy of the 1.2 kb Deletion of Example 2, in
7 differentiating between symptomatic control samples and samples of different endometrial
8 disease subtype. The 1.2 kb Deletion was evaluated for the ability to distinguish between
9 symptomatic patient specimens and specimens from patients stratified by subtype of
10 endometriosis (peritoneal, ovarian, deep infiltrating). FIG. 5A illustrates the distribution of
11 normalized 1.2 kb Deletion for specimens from symptomatic controls and patients with
12 peritoneal, ovarian or deep infiltrating endometriosis. Box boundaries represent the, 25th and
13 75th percentile, the line in the middle represents the median, and the whiskers represent the
14 90th (top) and 10th (bottom) percentiles. Dots represent outlier values (left). Descriptive
15 statistics are summarized for each group (right). In FIGS. 5B to 5D receiver operator
16 characteristic curves for the 1.2 kb Deletion were constructed and the areas under the curves
17 were calculated, showing diagnostic accuracy. Abbreviations: CI = confidence interval; Dev =
18 deviation; DIE = deep infiltrating endometriosis; N = number of specimens in each group; ROC
19 = receiver operator characteristic; Std = standard; vs = versus.

20 **[0025]** FIGS. 6A to 6D illustrate the diagnostic accuracy of the 3.7 kb Deletion of Example 2, in
21 differentiating between symptomatic control samples and samples of endometrial disease
22 subtype. The 3.7 kb Deletion was evaluated for the ability to distinguish between symptomatic
23 patient specimens and specimens from patients stratified by subtype of endometriosis
24 (peritoneal, ovarian, deep infiltrating). FIG 6A illustrates the distribution of normalized 3.7 kb
25 Deletion for specimens from symptomatic controls and patients with peritoneal, ovarian or deep
26 infiltrating endometriosis. Box boundaries represent the, 25th and 75th percentile, the line in the
27 middle represents the median, and the whiskers represent the 90th (top) and 10th (bottom)
28 percentiles. Dots represent outlier values (left). Descriptive statistics are summarized for each
29 group (right). In FIGS. 6B to 6D receiver operator characteristic curves for the 3.7 kb deletion
30 were constructed and the areas under the curves were calculated, showing diagnostic accuracy.
31 Abbreviations: CI = confidence interval; Dev = deviation; DIE = deep infiltrating endometriosis; N

= number of specimens in each group; ROC = receiver operator characteristic; Std = standard; vs = versus.

[0026] FIGs. 7A to 7C illustrate the diagnostic accuracy of the 1.2 kb Deletion of Example 2 in differentiating between symptomatic control samples and samples from patients with known disease stages. The 1.2 kb Deletion was evaluated for the ability to distinguish between symptomatic patient specimens and specimens from patients stratified by stage (low or high) of endometriosis. FIG. 7A illustrates the distribution of normalized 1.2 kb Deletion for specimens from symptomatic controls and patients with low (I/II) or high (III/IV) stages of endometriosis. Box boundaries represent the, 25th and 75th percentile, the line in the middle represents the median, and the whiskers represent the 90th (top) and 10th (bottom) percentiles. Dots represent outlier values (left). Descriptive statistics are summarized for each group (right). In FIGs. 7B and 7C receiver operator characteristic curves for the 1.2 kb Deletion were constructed and the areas under the curves were calculated, showing diagnostic accuracy. Abbreviations: CI = confidence interval; Dev = deviation; N = number of specimens in each group; ROC = receiver operator characteristic; Std = standard; vs = versus.

[0027] FIGs. 8A to 8C illustrate the diagnostic accuracy of the 3.7 kb Deletion of Example 2 in differentiating between symptomatic control samples and samples from patients with known disease stages. The 3.7 kb Deletion was evaluated for the ability to distinguish between symptomatic patient specimens and specimens from patients stratified by stage (low or high) of endometriosis. FIG. 8A illustrates the distribution of normalized 3.7 kb Deletion for specimens from symptomatic controls and patients with low (I/II) or high (III/IV) stages of endometriosis. Box boundaries represent the, 25th and 75th percentile, the line in the middle represents the median, and the whiskers represent the 90th (top) and 10th (bottom) percentiles. Dots represent outlier values (left). Descriptive statistics are summarized for each group (right). In FIGs. 8B and 8C, receiver operator characteristic (ROC) curves for the 3.7 kb Deletion were constructed and the areas under the curves were calculated, showing diagnostic accuracy. Abbreviations: CI = confidence interval; Dev = deviation; N = number of specimens in each group; ROC = receiver operator characteristic; Std = standard; vs = versus.

[0028] Fig. 9 is a scatterplot showing the difference in the 8.7kb deletion score between endometriosis positive samples, symptomatic controls samples and normal healthy control samples.

1 **[0029]** Fig .10 is a box and whisker plot showing the difference in the 8.7kb deletion score
2 between endometriosis positive samples, symptomatic controls samples and normal healthy
3 control samples.

4 **[0030]** Fig. 11 shows the ROC curves for the 8.7kb deletion comparing endometriosis positive
5 patients vs. healthy/normal controls.

6 **[0031]** Fig. 12 illustrates the diagnostic accuracy of the 8.7 kb deletion - symptomatic vs. all
7 endometrial disease. The 8.7 kb deletion was evaluated for its ability to distinguish between
8 samples from symptomatic patients and those from patients with confirmed endometriosis (all
9 subtypes/stages combined) by calculating the area under ROC curves. Abbreviations: CI =
10 confidence interval; ROC = receiver operator characteristic; Std = standard; vs = versus.

11 **[0032]** Figs. 13A to 13B further illustrate the diagnostic accuracy of the 8.7 kb deletion -control
12 vs. disease by subtype. These figures illustrate a study of whether the 8.7kb deletion assay
13 could distinguish between samples from symptomatic participants and those from participants
14 stratified by endometriosis subtype (peritoneal, ovarian, deep infiltrating). Fig. 13A shows the
15 normalized 8.7 kb deletion distribution for specimens from asymptomatic and symptomatic
16 controls, participants with peritoneal, ovarian or deep infiltrating endometriosis. The box
17 boundaries represent the 25th and 75th percentile, the line in the middle represents the median,
18 and the whiskers represent the 90th (top) and 10th (bottom) percentiles. The dots represent
19 outlier values (left). Descriptive statistics are also summarized for each group. Figs. 13B to
20 13D show the areas under the ROC curves, which were calculated to show diagnostic accuracy.
21 Abbreviations: As Con = asymptomatic controls; CI = confidence interval; Dev = deviation; DIE
22 = deep infiltrating endometriosis; N = number of specimens in each group; ROC = receiver
23 operator characteristic; Sym Con = symptomatic controls; Std = standard; vs = versus.

24 **[0033]** Figs. 14A to 14C further illustrate the diagnostic accuracy of the 8.7 kb deletion –
25 controls versus disease by stage. These figures illustrate whether the 8.7kb deletion assay
26 could distinguish between samples from symptomatic participants and those from participants
27 stratified by endometriosis stages I/II and III/IV. Fig. 14A shows normalized 8.7 kb deletion
28 distribution for specimens from symptomatic controls, participants with low (I/II) or high (III/IV)
29 stages of endometriosis. The box boundaries represent the 25th and 75th percentile, the line in
30 the middle represents the median, and the whiskers represent the 90th (top) and 10th (bottom)

percentiles. The dots represent outlier values (left). Descriptive statistics are summarized for each group. Figs. 14B and 14C show the areas under the ROC curves that were calculated to show diagnostic accuracy. Abbreviations: CI = confidence interval; Dev = deviation; N = number of specimens in each group; ROC = receiver operator characteristic; Sym Con = symptomatic controls; Std = standard; vs = versus.

[0034] Fig. 15 further illustrates the disease specificity of the 8.7 kb deletion for endometriosis. This figure summarizes the evaluation of the frequency of the 8.7 kb deletion in female cancers including endometrial cancer, ovarian cancer and breast cancer. Normalized 8.7 kb deletion distribution for specimens from endometrial cancer, ovarian cancer, breast cancer, symptomatic controls, and participants with peritoneal, ovarian or deep infiltrating endometriosis. The box boundaries represent the 25th and 75th percentile, the line in the middle represents the median, and the whiskers represent the 90th (top) and 10th (bottom) percentiles. The dots represent outlier values (left).

[0035] Fig. 16 is a scatterplot showing the difference in the 4.8kb deletion score between endometriosis positive samples, symptomatic controls samples and normal healthy control samples.

[0036] Fig. 17 is a box and whisker plot showing the difference in the 4.8kb deletion score between endometriosis positive samples, symptomatic controls samples and normal healthy control samples.

[0037] Fig. 18 illustrates a ROC for the 4.8kb deletion comparing data from endometriosis positive patients vs. symptomatic controls.

[0038] Fig. 19 illustrates a ROC for the 4.8kb deletion comparing data from endometriosis positive patients vs. healthy/normal controls.

[0039] Fig. 20 illustrates a deletion event according to the present description.

DETAILED DESCRIPTION

[0040] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although suitable materials and methods for the practice or testing of the present

invention are described below, other known materials and methods similar or equivalent to those described herein can be used.

[0041] The terms “deletion”, “deletion fragment”, or “deletion sequence” as used herein with respect to mtDNA will be understood to mean a nucleotide sequence or segment that is removed, or deleted, from the wild-type or naturally occurring mtDNA genome.

[0042] The terms “wild-type mtDNA” or “naturally occurring mtDNA” refer to the Revised Cambridge Reference Sequence (rCRS) (2001, GenBank accession number: NC_012920.1), which is provided herein as SEQ ID NO: 1. Although this sequence is identified as being 16569 bp in length, the actual number of nucleotides is 16568. As known in the art, this sequence includes a gap or placeholder nucleotide at position 3107.

[0043] The term “mutation” or “aberration”, as used herein with respect to mtDNA, will be understood to be synonymous with the term “deletion”.

[0044] The term “mutated mtDNA” or “aberrant mtDNA”, as used in the context of the present description, will be understood as meaning a mtDNA molecule having at least one deletion (as defined above) in its genome sequence.

[0045] The terms “junction” or “junction point” will be understood to mean the location in the nucleotide sequence of the re-circularized mtDNA molecule that includes the re-joined, or spliced, nucleotides of the remaining mtDNA genome sequence following removal of the deletion. As discussed further herein, the deletion event typically results in the creation of two new sequence fragments, consisting of a parent sequence, corresponding to the re-joined mtDNA molecule after removal of the deletion, and deleted sequence, corresponding to the deleted section. Generally, the parent sequence is longer than the deleted sequence. Often, and as discussed above, both the long and short fragments re-circularize to form what are known, respectively, as the large and small sublimons. As would be understood, both of the sublimons would have a unique junction point in their nucleotide sequence. Thus, the terms junction or junction point may be used to refer to either the large or small sublimons.

[0046] The phrase “having a deletion” will be understood to refer to a mtDNA molecule having a nucleotide sequence wherein a deletion sequence is removed. In other words, the phrase “a mtDNA having a deletion” refers to is the parent nucleic acid. Thus, a “mtDNA having a

common deletion” means a mtDNA molecule having a sequence that does not include the 4977 bp deletion sequence.

[0047] As used herein, the term “detecting” will be understood to mean determining or identifying and/or measuring or quantifying the presence in a biological sample of a particular feature. In one aspect, the term “detecting” will be used herein to refer to the identification of a mitochondrial DNA (mtDNA) sequence, more particularly, a mtDNA having a deletion. The term “detecting” may also be used to refer to the identification of a mitochondrial fusion transcript and/or a protein encoded by such mtDNA molecule. In the latter instance, the protein is referred to herein as a “fusion protein” and would include an amino acid sequence that results from the translation of the rejoined mtDNA following a deletion event. Such mtDNA may comprise the parent, or aberrant mtDNA or the deleted sequence.

[0048] As used herein, the term “diagnosing” will be understood to mean the identification of a disease condition or disease state or the determination of a higher, or increased probability of the existence of a disease condition or disease state. For example, with respect to the present description, a higher probability of the existence of a state or condition of endometriosis will be deemed to exist, or “diagnosed” when a mtDNA molecule or fusion transcript described herein is detected. It will be understood that the actual or clinical diagnosis of the state or condition will be made by a clinician upon examination of a biopsy sample or other such means. Thus, in some cases, the terms “detecting” and “diagnosing” may be used interchangeably herein.

[0049] As used herein, the term “biological sample” will be understood to refer to a tissue or bodily fluid containing cells or nucleic acids from which a molecule of interest can be obtained. The biological sample can be used either directly as obtained from the source or be initially subjected to a pre-treatment to modify the character of the sample. In one aspect, the biological sample is blood, in particular circulatory blood, it being understood that the term “blood” as used herein is intended to include blood derivatives, such as plasma and/or serum. In another aspect, the biological sample is menstrual fluid including menstrual blood. In another aspect, the biological sample is a tissue sample obtained from a subject. In one aspect, circulatory blood may be used as the biological sample. It will be understood that blood samples for the purpose of the present description may be drawn from any source on a subject’s body. This would include, without limitation, blood drawn from venous sources by syringe etc., collection of menstrual fluid samples, or capillary blood, such as blood drawn by finger pricks. Employing the

1 presently described methods using circulatory blood (including, as noted above, blood
2 derivatives), provides an effective means of detecting the presence of endometriosis in an
3 individual suspected of having such condition without having to unnecessarily undergo painful
4 and risky invasive procedures. As noted above, in situations where the presently described
5 methods indicate the presence of endometriosis, a diagnosis will still require a clinical
6 assessment, and perhaps laparoscopic examination/surgery or analysis of a biopsy sample.
7 Thus, it will be understood that, in one aspect, the presently described methods, in particular
8 when using circulatory blood (or one or more derivatives thereof, as described above) as the
9 biological sample, could be conducted on a sub-population of patients, including those
10 individuals that have one or more indications suggestive of the presence of endometriosis. It
11 will also be understood that the presently described methods may be conducted on general
12 population members as an initial phase of screening endometriosis. In other words, the
13 presently described methods may be performed on non-symptomatic subjects (i.e. individuals
14 who do not present with symptoms).

15 **[0050]** As used herein, the phrase "mitochondrial fusion transcript" or "fusion transcript" refers
16 to an RNA transcription product produced as a result of the transcription of a mtDNA sequence.

17 **[0051]** As used herein, the term "variant" refers to a nucleic acid sequence differing from a
18 naturally occurring sequence but retaining the essential or functional properties thereof. In one
19 aspect, the term "variant" may refer to a sequence that varies with respect to the wild-type
20 sequence. Generally, in the case of mtDNA, variants are overall closely similar, and, in many
21 regions, identical to a select mtDNA sequence. In the context of the present description,
22 variants may comprise at least one of the nucleotides of the junction point of the spliced genes
23 and may further comprise one or more nucleotides adjacent thereto. In one aspect, a variant
24 sequence is at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%,
25 97%, 98% or 99% identical to a given mtDNA sequence described herein, or its complementary
26 strand.

27 **[0052]** The phrase "substantially similar" as used herein refers to nucleic acids that are
28 functionally the same but differing in their respective nucleic acid sequences. In one aspect,
29 two sequences that are substantially similar to each other may be referred to as "variants".
30 Thus, two nucleic acid molecules may be considered substantially similar where a difference in
31 one or more nucleotides between the respective nucleic sequences does not alter their

functional properties or the functional properties of any polypeptides encoded by such nucleic acids. As would be understood, owing to the degeneracy of the genetic code, a base pair change can result in no change in the encoded amino acid sequence.

[0053] The phrase "substantial complementarity" refers to a sufficiently high degree of complementarity between the nucleotide sequences of nucleic acid molecules that allows hybridization there-between, but not necessarily 100% complementarity. For example, a primer or probe with substantial complementarity to a target sequence may have 80% to 99% sequence identity to the target sequence. In one aspect, substantial complementarity as used herein refers to at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity between sequences.

[0054] The term "fragment" as used herein refers to a nucleic acid sequence that is a portion of a given mitochondrial genomic sequence, or the complementary strand thereto. In one aspect, such "portion" includes at least two of the nucleotides comprising the junction point of spliced genes and may further comprise one or more nucleotides adjacent thereto. That is, the portion comprises the rejoined, or re-circularized DNA sequence after removal of a deletion. The fragments described herein are at least about 150 nucleotides (nt) in length, at least about 75 nt, at least about 50 nt, at least about 40 nt, at least about 30 nt, at least about 20 nt, or preferably at least about 15 nt in length. Although certain minimum nucleotide lengths are recited above, it will be understood, as described herein, that fragments of any size (e.g., 50, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 4000, 5000, 6000, 7000, 8000 or more nucleotides) are also contemplated.

[0055] In the context of sequence lengths, the term "about" as used herein, includes the particularly recited value or a value larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini.

[0056] As used herein, the term "probe" or "primer" refers to an oligonucleotide molecule that forms a duplex structure with, or "hybridizes" to, a target nucleic acid, due to complementarity of at least a portion of the nucleotide sequence of the probe/primer with a portion of the nucleotide sequence of the target molecule. The target nucleic acid molecule may in some cases be a fragment of a naturally occurring nucleic acid molecule. Probes described herein may be labeled according to methods known in the art. It will be understood that the probes or primers

described herein would be used under suitable hybridizing conditions as would be known to persons skilled in the art. The probes herein may also be referred to hybridizing probes. The probes and primers described herein may be of any length, as would be understood by persons skilled in the art. By way of example only, the presently described probes and primers may have lengths of about 150, 140, 130, 120, 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 15, or 10 nucleotides (nt). In one preferred aspect, the probes and/or primers described herein are about 12 to about 35 nt in length, or preferably about 18 to about 25 nt in length, and more preferably about 15 nt in length. As would be appreciated by persons skilled in the art, probes may have longer nucleotide lengths than primers. Thus, in some cases, the probes described herein may have lengths of about 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, or 2500 nucleotides. The present description is not limited to any particular probe or primer length.

[0057] The terms “comprise”, “comprises”, “comprised” or “comprising” may be used in the present description. As used herein (including the specification and/or the claims), these terms are to be interpreted as specifying the presence of the stated features, integers, steps or components, but not as precluding the presence of one or more other feature, integer, step, component or a group thereof as would be apparent to persons having ordinary skill in the relevant art. Thus, the term “comprising” as used in this specification means “consisting at least in part of”. When interpreting statements in this specification that include that term, the features, prefaced by that term in each statement, all need to be present but other features can also be present. Related terms such as “comprise” and “comprised” are to be interpreted in the same manner.

[0058] The term “and/or” can mean “and” or “or”.

[0059] Unless stated otherwise herein, the article “a” when used to identify any element is not intended to constitute a limitation of just one and will, instead, be understood to mean “at least one” or “one or more”.

[0060] As described herein, the present inventors have identified novel mtDNA deletions that are, in one aspect, associated with endometriosis and therefore constitute accurate diagnostic markers for such condition. The inventors have also identified novel mtDNA fusion transcripts that are, in one aspect, associated with endometriosis. Both of these aspects are discussed

further below. Translation products resulting from the fusion transcripts are also encompassed by the present description.

[0061] In one aspect, the present description relates to the inventors' hypothesis that endometrial cells shed in menstrual fluid during menorrhea would harbour the same genetic profile as endometrial-like cells present in ectopic and/or eutopic endometrial lesions. Using the knowledge gained from mapping the large-scale deletions of the human mitochondrial genome, the observation of high frequencies of these deletions, and the evidence in other disease types of transcriptionally active mutated mtDNA molecules, the inventors further hypothesized that mitochondrial deletions and fusion transcripts might be present in endometrial-like cells present in ectopic and/or eutopic endometrial lesions.

[0062] To test these hypotheses, 268 mitochondrial fusion transcripts were selected, based on predicted direct and indirect repeats throughout the mitochondrial genome, and screened for their use as biomarkers of endometriosis. A number of mtDNA deletions and corresponding fusion transcripts were identified by the inventors as being particularly useful in distinguishing samples with endometriosis from those without endometriosis. These deletions and fusion transcripts are discussed further below. These mtDNA molecules produce fusion sequences having open reading frames (ORFs) that can be transcribed by mitochondrial transcription machinery, resulting in fusion transcripts. Protein products, or fusion proteins, encoded by such fusion transcripts are also expected to be produced.

[0063] **1.0) mtDNA Deletions, Fusion Transcripts, and Translation Products**

[0064] **1.1) Mitochondrial DNA (mtDNA) Mutations**

[0065] As discussed above, mtDNA mutations generally comprise a deletion of a portion of the mtDNA wild-type sequence. The present description is based on associations between specific mtDNA mutations, in particular deletions of the mtDNA genomic sequence, and endometriosis.

[0066] According to the present description, to determine candidate genomic sequences, junction points resulting from sequence deletions were first identified. Sequence deletions were primarily identified by direct or indirect repetitive elements which flank the sequence to be deleted at the 5' and 3' end. The removal of a section of the nucleotides from the genome followed by the ligation of the remaining genome results in the creation of a novel junction point.

1 **[0067]** Upon identification of a junction point, the nucleotides of the genes flanking the junction
2 point were determined in order to identify a spliced gene. Typically, the spliced gene comprises
3 the initiation codon from the first gene and the termination codon of the second gene, and may
4 be expressed as a continuous transcript, i.e. one that keeps the reading frame from the
5 beginning to the end of both spliced genes. It is also possible that alternate initiation or
6 termination codons contained within the gene sequences may be used.

7 **[0068]** Large-scale deletions in the mitochondrial genome often result in two products arising
8 from the mutation process. These products are the result of the re-circularization of both parts
9 of the mtDNA genome: 1) a short sequence that may, in one aspect, correspond to the deleted
10 mtDNA sequence; and, 2) a long sequence that may, in one aspect, correspond to the
11 remaining mtDNA genomic sequence. It will be understood that depending on the size of the
12 deletion, the deletion may be larger than the remaining mtDNA. This situation would occur, for
13 example, when the deleted sequence is larger than about 8200 bp in length. Often, both the
14 short and long sequences re-circularize to form what are known, respectively, as the small and
15 large sublimons. In cases where the small component is of an insufficient number of
16 nucleotides, re-circularization is not possible, in which case the mutation process only results in
17 a large sublimon. As discussed herein, both large and small sublimons can be identified
18 thereby allowing both molecules to be used for detecting, diagnosing, and/or monitoring
19 endometriosis.

20 **[0069]** 1.2) Fusion Transcripts

21 **[0070]** Large-scale rearrangement mutations in the mitochondrial genome result in the
22 generation of fusion transcripts. Thus, it was expected that mtDNA rearrangements associated
23 with endometriosis would result in fusion transcripts that are also associated with endometriosis.
24 Thus, the use of mtDNA encoding such transcripts and probes directed thereto for the diagnosis
25 and monitoring of endometriosis are provided herein.

26 **[0071]** The present description provides the identification of fusion transcripts and associated
27 hybridization probes and primers useful in methods for predicting, diagnosing, and/or monitoring
28 endometriosis. One of skill in the art will appreciate that such molecules may be derived
29 through the isolation of naturally-occurring transcripts or, alternatively, by the recombinant
30 expression of mtDNA molecules isolated according to the methods of the invention. As

discussed, such mtDNA molecules typically comprise a spliced gene having the initiation codon from the first gene and the termination codon of the second gene. Accordingly, fusion transcripts derived therefrom comprise a junction points associated with the spliced genes.

[0072] 1.3) Translation Products

[0073] Based on the fusion transcripts described herein, the present description also provides amino acid sequences of putative proteins, i.e. "fusion proteins", resulting from the translation of the subject fusion transcripts. The description also provides translation products of at least a portion of the fusion transcripts, in particular the portion comprising the transcribed fusion site, or junction point of the mtDNA.

[0074] Fusion proteins of the description can be recovered and purified from a biological sample by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, hydrophobic charge interaction chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

[0075] Assaying fusion protein levels in a biological sample can occur using a variety of techniques. For example, protein expression in tissues can be studied with classical immunohistological methods (Jalkanen et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987)). Other methods useful for detecting protein expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (¹²⁵I, ¹²¹I), carbon (¹⁴C), sulfur (³⁵S), tritium (³H), indium (¹¹²In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

[0076] The polypeptides of the description can also be produced by recombinant techniques known in the art. Typically this involves transformation (including transfection, transduction, or infection) of a suitable host cell with an expression vector comprising a polynucleotide encoding the protein or polypeptide of interest.

[0077] Antibodies and protein binding agents

[0078] Protein specific antibodies for use in the assays of the present description can be raised against the wild-type or expressed fusion proteins described herein or an antigenic polypeptide fragment thereof, which may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. It will be understood that although antibodies are described, any other suitable binding agent, specific to identifying proteins, may also be used. In either case, the antibodies, or binding agents, are capable of identifying the fusion proteins described herein by means of specifically binding to a region of such proteins that is representative, or indicative, of the deletion. In one aspect, the fusions proteins have a unique amino acid profile that represents the translation of the junction point of the mtDNA molecule (either the large or small sublimon) after a deletion event.

[0079] As used herein, the term “antibody” (Ab) or “monoclonal antibody” (Mab) is meant to include intact molecules as well as antibody fragments, or antigen-binding fragments, thereof (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to, or having “specificity to”, a mitochondrial fusion protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding of an intact antibody (Wahl et al., J. Nucl. Med. 24:316-325 (1983)). Thus, these fragments are preferred.

[0080] The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing the mitochondrial fusion protein or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. In one method, a preparation of mitochondrial fusion protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

[0081] In a related method, the antibodies of the present description are monoclonal antibodies. Such monoclonal antibodies can be prepared using hybridoma technology (Kohler et al., Nature 256:495 (1975); Kohler et al., Eur. J. Immunol. 6:511 (1976); Kohler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., (1981) pp. 563-681). In general, such procedures involve immunizing an animal (preferably a mouse) with a mitochondrial fusion protein antigen or with a mitochondrial fusion protein-expressing cell.

[0082] In one aspect, the present description comprises immunological assays using antibodies or antigen-binding fragments having specificity to the fusion proteins described herein (as described above). Such immunological assays may be facilitated by kits containing the antibodies or antigen-binding fragments along with any other necessary reagents, test strips, materials, instructions etc.

[0083] Assays

[0084] Measuring the level of a translation product such as a fusion protein in a biological sample can determine the presence or progression of endometriosis in a subject. Therefore, in one aspect, the present description provides methods for predicting, diagnosing or monitoring endometriosis, comprising obtaining one or more biological samples, extracting mitochondrial fusion proteins from the samples, and assaying the samples for such molecules by: quantifying the amount of one or more molecules in the sample and comparing the quantity detected with a reference value. As would be understood by those of skill in the art, the reference value is based on whether the method seeks to predict, diagnose or monitor endometriosis. Accordingly, the reference value may relate to protein data collected from one or more control sample, or biological samples not positive for endometriosis, from one or more biological samples positive for endometriosis, and/or from one or more biological samples taken over time.

[0085] Techniques for quantifying proteins in a sample are well known in the art and include, for instance, classical immunohistological methods (Jalkanen et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987)). Additional methods useful for detecting protein expression include immunoassays such as the radioimmunoassay (RIA) and the enzyme linked immunosorbent assay (ELISA).

[0086] In one aspect, the description provides a method of detecting, diagnosing or monitoring endometriosis in a mammal, the method comprising assaying a tissue sample from the mammal for the presence of at least one mitochondrial fusion protein.

[0087] 2.0) Probes and Primers

[0088] 2.1) mtDNA Probes and Primers

1 **[0089]** Also described herein are mtDNA hybridization probes and/or primers capable of
2 hybridizing to aberrant mtDNA sequences under suitable hybridizing conditions. Any known
3 method of hybridization may be used.

4 **[0090]** Probes and/or primers may be generated directly against exemplary mtDNA fusion
5 molecules described herein (such as those listed in Table 1 below), or to a fragment or variant
6 thereof. For instance, the aberrant mtDNA sequences discussed herein can be used to design
7 primers or probes that will detect a nucleic acid sequence comprising a fusion nucleotide
8 sequence of interest. As would be understood by those of skill in the art, primers and/or probes
9 that hybridize to these nucleic acid molecules may do so under highly stringent hybridization
10 conditions or lower stringency conditions. Such conditions would be known to those skilled in
11 the art and are described, for example, in Current Protocols in Molecular Biology (John Wiley &
12 Sons, New York (1989)), 6.3.1-6.3.6.

13 **[0091]** In some aspects, the probes and primers described herein contain a sequence
14 complementary to at least a portion of the aberrant mtDNA comprising the junction point of the
15 spliced genes. As discussed above, this “portion” includes at least the two nucleotides
16 remaining in the mtDNA genome after removal of the deletion, thereby resulting in a junction
17 point, identified herein as A:B, where “A” and “B” represent the mtDNA genomic nucleotides on
18 opposite sides of the deleted sequence, but which are adjacent to each other after the
19 remaining sequence is re-circularized. The “portion” may further comprise one or more
20 nucleotides adjacent to the junction point. In this regard, the present description encompasses
21 any suitable targeting mechanism that will select a mtDNA molecule using the nucleotides
22 involved in and/or adjacent to the junction point A:B. It is further contemplated herein that
23 primer and probe sequences could be altered by one or more base pairs while still enabling
24 hybridization to the target sequence. Such primers or probes will be referred to as having
25 “substantial complementarity” to the target sequence. As discussed above, after a deletion
26 event, both large and small sublimons may result, both of such sublimons would have a
27 respective junction point, such as defined above, once the molecules are re-circularized.

28 **[0092]** Further, the present description encompasses primers that are designed, in one aspect,
29 to span the deletion junction or junction point A:B in the forward or reverse direction. In another
30 aspect, one or more primer may be designed to hybridize to a location on the target sequence
31 that is adjacent the junction point.

1 **[0093]** Various types of probes known in the art are contemplated for use in the present
2 description. For example, the probe may be a hybridization probe, the binding of which to a
3 target nucleotide sequence can be detected using a general DNA binding dye such as ethidium
4 bromide, SYBR® Green, SYBR® Gold and the like. Alternatively, the probe can incorporate
5 one or more detectable labels. Detectable labels are molecules or that can be detected directly
6 or indirectly and are chosen such that the ability of the probe to hybridize with its target
7 sequence is not affected. Methods of labelling nucleic acid sequences are well-known in the art
8 (see, for example, Ausubel et al., (1997 & updates) Current Protocols in Molecular Biology,
9 Wiley & Sons, New York).

10 **[0094]** Labels suitable for use with the probes of the present description include those that can
11 be directly detected, such as radioisotopes, fluorophores, chemiluminophores, enzymes,
12 colloidal particles, fluorescent microparticles, and the like. One skilled in the art will understand
13 that directly detectable labels may require additional components, such as substrates, triggering
14 reagents, light, and the like to enable detection of the label. The present description also
15 contemplates the use of labels that are detected indirectly.

16 **[0095]** As discussed above, the presently described probes and primers may be of any suitable
17 length as would be understood by persons skilled in the art. Nucleotide lengths of the probes
18 and primers of the present description were discussed above. As discussed above, the probes
19 and/or primers described herein may preferably be about 12 to about 25 nucleotides in length,
20 more preferably about 12 to about 15 nt in length. It will be understood that the primers and/or
21 probes described herein may preferably be of a length that is at least the size of the mtDNA
22 repeat (i.e. repeated) sequence. The present description is not limited to any particular primer
23 or probe length.

24 **[0096]** The probes described herein will preferably hybridize to nucleic acid molecules from the
25 biological samples described herein, thereby enabling the described methods. Accordingly, in
26 one aspect, there is provided a hybridization probe for use in the detection of endometriosis,
27 wherein the probe is complementary to, or substantially complementary to, at least a portion of
28 an aberrant mtDNA molecule described herein or a portion of a deleted sequence from the
29 mtDNA genome.

30 **[0097]** 2.2) Fusion Transcript Probes and Primers

1 **[0098]** Once a fusion transcript has been characterized, primers or probes can be developed to
2 target the transcript in a biological sample. Such primers and probes may be prepared using
3 any known method (as described above) or as set out in the examples provided below. A probe
4 may, for example, be generated for a fusion transcript, and detection technologies, such as
5 QuantiGene™ 2.0 by Panomics™, can be used to detect the presence of the transcript in a
6 sample. Primers and probes may be generated directly against exemplary fusion transcripts
7 described herein, or to a fragment or variant thereof. For instance, the sequences set forth
8 herein (such as those listed in Table 2 below) can be used to design probes or primers that will
9 detect an RNA sequence comprising a fusion sequence of interest.

10 **[0099]** As would be understood by those skilled in the art, probes and primers designed to
11 hybridize to the fusion transcripts described herein comprise sequences complementary, or
12 substantial complementary, to at least a portion of the transcript expressing the junction point of
13 the spliced genes. This portion includes at least two of the nucleotides complementary to the
14 expressed junction point and may further comprise one or more complementary nucleotides
15 adjacent thereto. In this regard, the present description encompasses any suitable targeting
16 mechanism that will select a fusion transcript that uses the nucleotides involved and adjacent to
17 the junction point of the spliced genes.

18 **[0100]** Various types of probes and methods of labelling known in the art are contemplated for
19 the preparation of transcript probes described herein. Some examples of such types and
20 methods have been described above with respect to the detection of genomic sequences. The
21 transcript probes of the present description are at least about 150 nt, at least about 75 nt, at
22 least about 50 nt, at least about 40 nt, at least about 30 nt, at least about 20 nt, or preferably at
23 least about 12-15 nt in length. A probe of "at least 20 nt in length," for example, is intended to
24 include 20 or more contiguous bases that are complementary to an mtDNA sequence of the
25 invention. Of course, larger probes (e.g., 50, 150, 500, 600, 2000 nucleotides) may be
26 preferable. As mentioned above, primers or probes of 18 to 25 nt are preferable.

27 **[0101]** In some aspects, there is provided one or more hybridization probes and/or primers for
28 use in the detection of endometriosis, wherein the one or more probes and/or primers is/are
29 complementary to, or substantially complementary to, at least a portion of a mitochondrial fusion
30 transcript described herein.

[0102] 3.0) Assays for detecting mtDNA deletions, fusion transcripts and protein products thereof

[0103] As mentioned above, the present description provides mitochondrial DNA biomarkers that are useful in detecting, diagnosing, and/or monitoring endometriosis in a subject using a biological sample from the subject. In particular, such biological sample is non-invasively collected menstrual fluid, circulatory blood, and/or tissue (such as biopsy tissue). The description therefore provides, in one aspect, a menstrual-fluid- or blood-based test that will enable the early and accurate detection of endometriosis, thereby preventing unnecessary initial and repeat surgical procedures. Thus, the methods described herein will reduce the need for unnecessary laparoscopic procedures when endometriosis is suspected but not detected. The present methods will also aid in determining whether endometriosis has recurred by allowing for the monitoring of endometriosis in a subject over time.

[0104] 3.1) Measurement of Aberrant mtDNA

[0105] According to the methods described herein, measuring the level of one or more aberrant mtDNA marker of the present invention in a biological sample can determine the presence or stage or progression of endometriosis in a subject. The present description, therefore, provides methods for detecting, diagnosing and/or monitoring endometriosis in a subject, comprising assaying a biological sample from the subject for one or more aberrant mtDNA biomarkers (or “markers”) described herein by measuring and/or quantifying the amount of the one or more aberrant mtDNA markers in the sample. Once quantified, the amount of the marker may be compared with a reference value (i.e., a control). The reference value may be based on whether the method seeks to detect, diagnose or monitor endometriosis. For example, in the case of detecting or diagnosing endometriosis, the reference value may comprise the amount of the aberrant mtDNA in a sample from a healthy subject, i.e. a subject not suffering from endometriosis. Such sample may be described herein as a “known non-endometriotic” (or “non-involved”) biological sample. Alternatively, the reference value may comprise the amount of aberrant mtDNA in a sample from a subject known to be suffering from endometriosis. Such sample may be described herein as a “known endometriotic” (or “involved”) biological sample. Where the control comprises a value, or amount, from a non-endometriotic source, it may be referred to herein as a “non-endometriotic amount”. In other aspects described herein, the control may comprise a reference value of another analyte from the same biological sample. In

1 some cases, and as described further herein, the amount of the aberrant mtDNA may be first
2 normalized against an amount of nuclear DNA, taken from the same subject, such as one that
3 codes for one or more housekeeping genes, such as those coding for rRNA. In one aspect the
4 nuclear DNA sequence used may code for the 18S rRNA. The normalized value of mtDNA
5 could then be compared to a threshold value. In the case of detecting or diagnosing
6 endometriosis, an increase in the amount of the subject aberrant mtDNA is indicative of
7 endometriosis. In the case of monitoring endometriosis, biological samples may be taken over
8 time from the subject and compared over a given time period. An increase in the amount of one
9 or more of the herein described aberrant mtDNA over time indicates the development,
10 recurrence, or advancement of endometriosis in the subject.

11 **[0106]** The presently described methods also encompass assaying a biological sample for a
12 panel of aberrant mtDNA markers described herein, wherein such panel comprises two or more
13 of the subject mtDNA markers. For example, such panel may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10,
14 or more of the presently described mtDNA markers.

15 **[0107]** In one aspect, there is provided herein a method of detecting endometriosis in a
16 mammal, the method comprising assaying a biological sample (such as blood, menstrual fluid, a
17 tissue sample etc.) from the mammalian subject for the presence of aberrant mtDNA by
18 hybridizing the sample with at least one hybridization probe that is capable of recognizing, or
19 hybridizing to, a mutant mtDNA sequence as described herein. In particular, and as described
20 herein, such probe is provided with a nucleotide sequence that is adapted to hybridize with a
21 portion of a mtDNA molecule of the sample, wherein such portion includes a junction point
22 described herein.

23 **[0108]** In some aspects, the present methods comprise assaying a biological sample from a
24 mammal by hybridizing the sample with at least two primers adapted to hybridize to an aberrant
25 mtDNA molecule as described herein. In one aspect, one of the primers may be designed with
26 a nucleotide sequence that is complementary to a portion of mtDNA having a junction point as
27 described herein. In another aspect, the primers may be provided with nucleotide sequences
28 that hybridize to regions adjacent the mtDNA junction point and adapted to overlap the junction
29 point.

1 **[0109]** In another aspect, the present description provides a method for detecting
2 endometriosis, wherein the assay comprises:

3 **[0110]** a) conducting a hybridization reaction using at least one of the probes described
4 herein to allow the at least one probe to hybridize to an aberrant mitochondrial DNA sequence
5 extracted from a biological sample;

6 **[0111]** b) quantifying the amount of the at least one aberrant mitochondrial DNA
7 sequence in the sample by quantifying the amount of the mitochondrial DNA hybridized to the at
8 least one probe; and,

9 **[0112]** c) comparing the amount of the mitochondrial DNA in the sample to at least one
10 known reference value, wherein:

11 **[0113]** - if the reference value comprises an amount of mtDNA not associated with
12 endometriosis, a higher amount of aberrant mtDNA in the sample indicates the presence of
13 endometriosis; or,

14 **[0114]** - if the reference value comprises an amount of mtDNA associated with
15 endometriosis, a lower amount of aberrant mtDNA in the sample indicates the absence of
16 endometriosis.

17 **[0115]** Methods and screening tools for diagnosing endometriosis by identifying specific
18 mitochondrial mutations are also herein contemplated. Any known method of hybridization may
19 be used to carry out such methods including, without limitation, probe- and/or primer-based
20 technologies including branched DNA and qPCR, both single-plex and multi-plex. Array
21 technology, which has oligonucleotide probes matching the wild type or mutated region, and a
22 control probe, may also be used. Commercially available arrays such as microarrays or gene
23 chips are suitable for use with the presently described methods.

24 **[0116]** Thus, by detecting the aberrant mtDNA molecules described herein in a biological
25 sample, it is possible to detect or diagnose endometriosis in a subject. Further, by measuring
26 and comparing, either qualitatively or quantitatively, the amount of aberrant mtDNA in
27 successive samples from a subject over time, the progression of endometriosis in such subject
28 can be monitored.

1 **[0117]** 3.2) Measurement of Fusion Transcripts

2 **[0118]** Measuring the level of the herein described mitochondrial fusion transcripts in a
3 biological sample can also determine the presence or stage or progression of endometriosis in a
4 subject. Thus, there is provided methods for detecting, diagnosing, and/or monitoring
5 endometriosis, comprising extracting mitochondrial RNA from one or more biological samples
6 obtained from a subject, and assaying the samples for fusion transcripts corresponding to the
7 aberrant mtDNA described herein. Such assaying may comprise quantifying the amount of one
8 or more fusion transcripts in the sample and comparing the amount detected with a reference
9 value. The reference value is based on whether the method seeks to diagnose or monitor
10 endometriosis. Accordingly, the reference value may relate to transcript data collected from one
11 or more known non-endometriotic biological samples, from one or more known endometriotic
12 biological samples, a population of known non-endometriotic or known endometriotic samples,
13 and/or from one or more biological samples taken from the subject over time.

14 **[0119]** In one aspect, the methods described herein encompass assaying one or more
15 biological samples from a subject for a panel of fusion transcript markers indicative of
16 endometriosis, wherein the panel comprises 2, 3, 4, 5, 6, 7, 8, 9 or 10 RNA markers described
17 herein.

18 **[0120]** Thus, in one aspect, there is provided a method of detecting endometriosis in a
19 mammal, the method comprising assaying a biological sample (such as blood, menstrual fluid,
20 or tissue) from said mammal for the presence of at least one fusion transcript described herein
21 by hybridizing said sample with at least one hybridization probe having a nucleic acid sequence
22 complementary to at least a portion of the mitochondrial fusion transcript, wherein the portion
23 includes a fusion junction in the mitochondrial fusion transcript.

24 **[0121]** In another aspect, there is provided methods comprising assaying a biological sample
25 from the mammal by hybridizing the sample with at least two primers. As discussed above, at
26 least one of the primers may have a sequence that allows hybridization to a portion of the fusion
27 transcript including the fusion junction. In other aspects, the primers may have sequences that
28 allow hybridization to flanking regions of the fusion junction.

29 **[0122]** In another aspect, the invention provides a method as above, wherein the assay
30 comprises:

1 [0123] a) conducting a hybridization reaction using at least one of the above-noted
2 probes to allow the at least one probe to hybridize to a complementary mitochondrial fusion
3 transcript;

4 [0124] b) quantifying the amount of the at least one mitochondrial fusion transcript in the
5 sample by quantifying the amount of the transcript hybridized to the at least one probe; and,

6 [0125] c) comparing the amount of the mitochondrial fusion transcript in the sample to at
7 least one known reference value, wherein:

8 [0126] - if the reference value comprises an amount of mitochondrial fusion transcript
9 not associated with endometriosis, a higher amount of it in the sample indicates the presence of
10 endometriosis; and

11 [0127] - if the reference value comprises an amount of mtDNA associated with
12 endometriosis, a lower amount of aberrant mtDNA in the sample indicates the absence of
13 endometriosis.

14 [0128] 3.3) Detection of Translated Proteins

15 [0129] Translation products, proteins, of the fusion transcripts described herein may be
16 detected using commonly known methods, such as immunological assays utilizing antibodies or
17 other such specific binding components. In particular, such components specifically bind to the
18 translated fusion site or junction of mtDNA

19 [0130] 4.0) Kits

20 [0131] The present description encompasses diagnostic or screening kits for the *in vitro*
21 detection, diagnosis, and/or monitoring of endometriosis of a subject. Such kits preferably
22 include one or more probes or primers as described herein, optionally in combination with
23 reagents, instructions, tools, and/or containers etc., as may be needed for conducting an assay.

24 [0132] The kits can include reagents required to conduct a diagnostic assay, such as buffers,
25 salts, detection reagents, anticoagulating agents, and the like. Other components, such as
26 buffers and solutions for the isolation and/or treatment of a biological sample, may also be

1 included in the kit. One or more of the components of the kit may be lyophilised and the kit may
2 further comprise reagents suitable for the reconstitution of the lyophilised components.

3 **[0133]** Where appropriate, the kits described herein may also contain sampling means,
4 reaction vessels, mixing vessels, and/or other components to facilitate the collection and/or
5 preparation of the test sample. The kit may also optionally include instructions for use, which
6 may be provided in paper form or in computer-readable form, such as a disc, CD, DVD or the
7 like.

8 **[0134]** In one aspect, the description provides a kit for conducting an *in vitro* assay for detecting
9 and/or diagnosing endometriosis comprising a hybridization probe described herein and at least
10 one reagent for conducting the assay.

11 **[0135]** In one aspect, a kit described herein comprises at least one hybridization probe
12 complementary to at least a portion of an aberrant mtDNA described herein or at least a portion
13 of a mitochondrial RNA fusion transcript described herein. As discussed above, in one aspect,
14 the portion of the sequence to which the probe hybridizes comprises a junction point, or fusion
15 junction, in the mtDNA or the fusion transcript. In one aspect, the kit may comprise one or more
16 probes that are adapted to hybridize to one or more control sequences.

17 **[0136]** In another aspect, a kit described herein comprises a pair of primers, such as forward
18 and reverse primers, for amplifying at least a portion of an aberrant mtDNA described herein or
19 at least a portion of a mitochondrial RNA fusion transcript described herein. In one aspect, at
20 least one of the primers has a nucleotide sequence that is adapted to hybridize to a junction
21 point, or fusion junction, in the mtDNA or the fusion transcript. In another aspect, at least one of
22 the primers has a nucleotide sequence that is adapted to hybridize to a sequence of the mtDNA
23 or the fusion transcript that is adjacent to the junction point, or fusion junction, in the mtDNA or
24 the fusion transcript. In one aspect, the kit may comprise one or more primers or primer pairs
25 that are adapted to hybridize to one or more control sequences.

26 **[0137] 5.0) Exemplary mtDNA Mutations, Fusion Transcripts, Translation Products,**
27 **Probes, and Primers**

28 **[0138]** Described below are mtDNA mutations (or aberrant mtDNA) and fusion transcripts that
29 have been found to be useful for the presently claimed methods. Putative translation products

are also provided and are believed to also be useful for the same reason. Also provided below are probe and primer sequences that are useful for detecting the subject mtDNA and fusion transcripts.

[0139] 5.1) Exemplary mtDNA Mutations

[0140] Table 1 lists the aberrant mtDNA molecules (i.e. mtDNA molecules having a deletion) that were studied. The listed sequences are based on modifications of the wild type mitochondrial genome (SEQ ID NO: 1) and have been assigned a fusion or "FUS" designation. Where provided, "AltMet" refers to alternate translation start site. The sequences listed in Table 1 are sections of the mtDNA genome that are rejoined, or re-circularized after removal of the subject deletion.

[0141] Table 1: mtDNA aberrations

Deletion ID	SEQ ID NO:	Deletion Name (Location of Deletion, with reference to SEQ ID NO: 1)	Spliced Genes	Location on mtDNA	Junction site (splice location on SEQ ID)	Location of flanking repeats (with reference to SEQ ID NO: 1)
1	2	FUS 8469:13447 (AltMet) ("5.0 kb Deletion" or 4977 kb "common deletion") (8483-13446)	(ATP8) to (ND5)	8366-14148	8366-8469/13447-14148 (nucleotides 81-82 of SEQ ID NO: 2)	[8470-8482] and 13447-13459
4	3	FUS 7992:15730 ("7.7 kb Deletion") (7993-15722)	(CO2) to (Cytb)	7586-15887	7586-7992/15730-15887 (nucleotides 407-408 of SEQ ID NO: 3)	7986-7992 and [15723-15729]
14	4	FUS 9191:12909 (9196-12908)	(ATP6) to (ND5)	8527-14148	8527-9191/12909-14148 (nucleotides 665-666 of SEQ ID NO: 4)	[9192-9195] and 12909-12911

14a	5	FUS 9188:12906 ("3.7 kb Deletion") (9196-12905)	(ATP6) to (ND5)	8527-14148	8527-9188/12906- 14148 (nucleotides 662-663 of SEQ ID NO: 5)	[9189-9195] and 12906-12912
16	6	FUS 10367:12829 (10368-12825)	(ND3) to (ND5)	10059- 14148	10059-10367/12829- 14148 (nucleotides 309-310 of SEQ ID NO: 6)	10365-10367 and [12826-12828]
120	7	FUS 6260:12814 ("6.5 kb Deletion") (6271-12813)	(CO1) to (ND5)	5904-14148	5904-6260/12814- 14148 (nucleotides 357-358 of SEQ ID NO: 7)	[6261-6270] and 12814-12823
122	8	FUS 7973:9023 ("1.0 kb Deletion") (7984-9022)	(CO2) to (ATP6)	7586-9207	7586-7973/9023- 9207 (nucleotides 387-388 of SEQ ID NO: 8)	[7974-7983] and 9023-9032
193	9	FUS 9086:10313 ("1.2 kb Deletion") (9087-10303)	(ATP6) to (ND3)	8527-10404	8527-9086/10313- 10404 (nucleotides 560-561 of SEQ ID NO: 9)	9078-9086 and [10304-10312]
400	10	FUS 9079:14988 (9086-14987)	(ATP6) to (CYTB)	8527-15887	8527-9079/14988- 15887 (nucleotides 553-554 of SEQ ID NO: 10)	[9080-9085] and 14988-14903
516	11	FUS 7260:15540 (7261-15531)	(CO1) to (CYTB)	5904-15887	5904-7260/15540- 15887 (nucleotides 1357- 1358 of SEQ ID NO: 11)	7253-7260 and [15532-15539]
586	12	FUS 8431:10841 ("2.4 kb Deletion") (8440-10840)	(ATP8) to (ND4)	8366-12137	8366-8431/10841- 12137 (nucleotides 66-67 of SEQ ID NO: 12)	[8432-8439] and 10841-10848
8590	74	FUS 8984:13833 ("4.8 kb Deletion") (8994-13832)	(ATPase6) to (ND5)	8527-14101	8527-8984/13833- 14101 (nucleotides 457-458 of SEQ ID NO: 74)	[8985-8993] and 13833-13841

2767	75	FUS 5362:14049 ("8.7 kb Deletion") (5377-14048)	(ND2) to (ND5)	4470-14101	4470-5362/14049- 14101 (nucleotides 893-894 of SEQ ID NO: 75)	[5363-5376] and 14049-14062
------	----	--	-------------------	------------	--	--

1

2 **[0142]** In Table 1, "Deletion ID" is a reference that identifies the mtDNA deletion from among
3 those screened. "SEQ ID NO" indicates the nucleotide sequence identifier ascribed herein to
4 the subject mtDNA deletion. "Deletion Name" identifies the "FUS" designation, wherein A:B
5 represents the junction point between the last mitochondrial nucleotide of the first spliced gene
6 and the first mitochondrial nucleotide of the second spliced gene. The "Location of Deletion"
7 identifies the portion of the respective sequence that is deleted from the parent mtDNA
8 molecule. The following column, "Spliced Genes", identifies the spliced genes resulting from the
9 deletion. In this regard, ATP8 represents ATPase8, ATP6 represents ATPase6, CO2
10 represents COII, and CO1 represents COI. The "mtDNA Location" identifies the segment of the
11 mtDNA sequence corresponding to the wild type mtDNA genome (i.e. SEQ ID NO: 1). The
12 "Junction Site" identifies the location of the junction point of the mutated mtDNA following
13 removal of the deletion (based on the wild type mtDNA genome, SEQ ID NO: 1). Thus, by way
14 of example, for Deletion ID No. 4 (SEQ ID NO: 3), having a "Junction Site" 7586-7992/15730-
15 15887, the deleted mtDNA segment includes nucleotides 7993 to 15729. In such case, the
16 aberrant mtDNA, once re-circularized, comprises a junction at nucleotides 7992 and 15730.
17 The portion in brackets in this column identifies the location of the splice in the respective SEQ
18 ID NO. The final column identifies the repeat sequences flanking the deletion. The repeats
19 shown in square brackets are deleted along with the deletion shown in the third column.

20 **[0143]** As shown in Table 1, one of the flanking repeat sequences is removed along with, and
21 therefore forms part of, the deleted sequence. It is, however, possible that the other of the
22 repeat sequences may be included with the deletion instead. This deletion mechanism is
23 illustrated in Fig. 20, which shows a parent mtDNA molecule 10, wherein 12 and 20 represent
24 the opposed ends of the mtDNA molecule and 16 represents the deletion, or deleted sequence
25 (such as recited in the third column of Table 1). The repeat sequences are represented at 14
26 and 18. During the deletion event, one of the repeats 14 or 18 is deleted along with, and
27 therefore forms part of, the deletion 16. As such, the remaining parent mtDNA, once re-
28 circularized, will comprise segments 12-18-20 or segments 12-14-10, as illustrated in Fig. 20.

Although, as noted above, one entire repeat is described as being included with the deleted sequence, there is a possibility that the only a portion of one or both of the repeats may be included with the deletion.

[0144] Mutant mtDNA sequences according to the present description may comprise any modification that results in the generation of a fusion transcript. Non-limiting examples of such modifications include insertions, translocations, deletions, duplications, recombinations, rearrangements, or combinations thereof.

[0145] The step of detecting the presently described mtDNA mutations can be selected from any technique known to those skilled in the art. For example, analyzing mtDNA can comprise selection of targets by branching DNA, sequencing the mtDNA, amplifying mtDNA by PCR, Southern, Northern, Western, South-Western blot hybridizations, denaturing HPLC, hybridization to microarrays, biochips or gene chips, molecular marker analysis, biosensors, melting temperature profiling or a combination of any of the above.

[0146] Variants or fragments of the mtDNA sequences identified herein are also contemplated. The present description encompasses the use of variants or fragments of these sequences for diagnosing and/or monitoring endometriosis.

[0147] 5.2) Exemplary Fusion Transcripts

[0148] Exemplary fusion transcripts for use in the methods described herein are provided in Table 2. These fusion transcripts were detected and found to be useful in detecting, diagnosing and/or monitoring endometriosis as indicated in the Examples.

[0149] Table 2: Fusion transcripts of the present invention.

Deletion ID	mtDNA Deletion SEQ ID NO	Transcript SEQ ID NO	Fusion Transcript Name	Flanking Genes	Deletion Junction
1	2	13	FUS 8469:13447 (AltMet)	(ATP8) to (ND5)	8469:13447
4	3	14	FUS 7992:15730	(CO2) to (Cytb)	7992:15730
14	4	15	FUS 9191:12909	(ATP6) to (ND5)	9191:12909
14a	5	16	FUS 9188:12906	(ATP6) to (ND5)	9188:12906
16	6	17	FUS 10367:12829	(ND3) to (ND5)	10367:12829
120	7	18	FUS 6260:12814	(CO1) to (ND5)	6260:12814
122	8	19	FUS 7973:9023	(CO2) to (ATP6)	7973:9023
193	9	20	FUS 9086:10313	(ATP6) to (ND3)	9086:10313

400	10	21	FUS 9079:14988	(ATP6) to (CYTB)	9079:14988
516	11	22	FUS 7260:15540	(CO1) to (CYTB)	7260:15540
586	12	23	FUS 8431:10841	(ATP8) to (ND4)	8431:10841
8590	74	76	FUS 8984:13833	(ATPase6) to (ND5)	8984:13833
2767	75	77	FUS 5362:14049	(ND2) to (ND5)	5362:14049

[0150] In Table 2: "Transcript Number" is identification number assigned to the fusion transcript and also corresponds to the mtDNA deletion ID number of Table 1. "mtDNA Deletion SEQ ID NO" is the mtDNA deletion sequence identifier from Table 1. "Transcript SEQ ID NO" is the sequence identifier of the subject fusion transcript. "Fusion Transcript Name" identifies the "FUS" designation, wherein A:B represents the junction point between the last mitochondrial nucleotide of the first spliced gene and the first mitochondrial nucleotide of the second spliced gene. "Flanking Genes" identifies the spliced genes resulting from the deletion. "Deletion Junction" identifies the location of the junction point of the mtDNA molecule after removal of the deletion.

[0151] Naturally occurring fusion transcripts can be extracted from a biological sample and identified according to any suitable method known in the art, such as those methods described in the examples of the present description.

[0152] Fusion transcripts can also be produced by recombinant techniques known in the art. Typically, this involves transformation (including transfection, transduction, or infection) of a suitable host cell with an expression vector comprising an mtDNA sequence of interest.

[0153] Variants or fragments of the fusion transcripts identified herein are also contemplated.

[0154] 5.3) Exemplary Translation Products of Fusion Transcripts

[0155] Putative amino acid sequences corresponding to transcripts of the mtDNA deletions 1, 4, 14, 16, 120, 122, 193, 400, 516, 586, 8590, and 2767 are provided in Table 3.

[0156] Table 3: Putative amino acid sequences corresponding to the fusion transcripts of the present description

Deletion ID	Deletion Name	mtDNA Deletion SEQ ID NO	Transcript SEQ ID NO	Amino Acid SEQ ID NO
1	FUS 8469:13447 (AltMet)	2	13	24

4	FUS 7992:15730	3	14	25
14	FUS 9191:12909	4	15	26
14a	FUS 9188:12906	5	16	27
16	FUS 10367:12829	6	17	28
120	FUS 6260:12814	7	18	29
122	FUS 7973:9023	8	19	30
193	FUS 9086:10313	9	20	31
400	FUS 9079:14988	10	21	32
516	FUS 7260:15540	11	22	33
586	FUS 8431:10841	12	23	34 and/or 84
8590	FUS 8984:13833	74	76	78
2767	FUS 5362:14049	75	77	79

1

2 **[0157] Examples**

3 **[0158]** The following examples are provided to further illustrates aspects of the present
4 description. The examples are not intended to limit the scope of the description in any way.

5 **[0159] Example 1: Large Scale Fusion Transcript Screening in Endometrial Tissue**

6 **[0160]** 268 probes corresponding to fusion transcripts were screened on endometrial tissue
7 samples for evidence of differential expression in samples obtained from patients with
8 endometriosis relative to control samples. Screening methods and results are described herein
9 below.

10 **[0161] Generation of probe library**

11 **[0162]** The 268 probes were identified using a proprietary nucleotide base pair repeat finding
12 program. The program identified over 16000 potential deletions based on direct and indirect
13 repetitive elements which flank the sequence to be deleted at the 5' and 3' end. The selection
14 of the 268 probes was based on the criteria that a minimum of 8 base pair repeats were
15 required; however, deletions with fewer than 8 base pair repeats are also possible. By way of
16 example, the repeat for deletion 16 is 3bp.

17 **[0163] Tissue samples**

18 **[0164]** Large endometrium samples (>0.49 g) were obtained. The "state" or diagnosis of the
19 tissue as determined by the physician upon surgery as well as the reason(s) for surgery are
20 indicated in Table 4.

[0165] Table 4: Tissue samples.

Sample ID	State/Reason for surgery
4360	Endometriosis (ovarian cyst)
3461	Endometriosis (ovarian endometrioma)
3462	Endometriosis (ovarian cyst, pain)
3463	Control (had endometriosis 7-8 yrs prior to surgery; not present at time of surgery for possible tubal ligation)
3464	Control (uterine fibroid)
3465	Control (possible tubal reversal)
3466	Control (ablation of suspected endometriosis lesions)

[0166] Using the samples listed in Table 4, tissue homogenates were prepared using the QuantiGene™ Sample Processing Kit for “Fresh or Frozen Animal Tissues”. For each sample, 4 portions of frozen endometrial tissue were cut and weighed (approximately 100mg each) before being added to 6mL of homogenizing solution containing 60 µL proteinase K. Samples were homogenized using Qiagen’s Tissue Rupture probe then incubated at 65°C overnight. Homogenates were then clarified by centrifuging twice at 16000 x g for 15 minutes. The supernatant was conserved and utilized as the template for the subsequent branched DNA assay. Alternatively, DNA was extracted from the tissue homogenates or directly from fresh frozen tissue according to the protocol for tissue using Qiagen’s QiaAmp™ DNA Mini Kit. DNA was then quantified on the Nanodrop™ spectrophotometer and normalized for subsequent use in a qPCR reaction.

[0167] Mitochondrial DNA deletions and resulting fusion transcripts can be detected using one of many molecular techniques. Herein, branched-DNA and quantitative PCR technologies were employed for the detection of fusion transcripts and the parent aberrant mtDNA molecules, respectively.

[0168] Branched DNA Platform

[0169] Panomics’ QuantiGene™ 2.0 protocol for “Capturing Target RNA from Fresh, Frozen, or FFPE Tissue Homogenates” was followed for tissue samples. A working probe set comprised of water, lysis solution, blocking reagent and probe was first added to the capture plate. The probes (or “capture probes”) used in the present example comprise oligonucleotides that were designed (with complementary nucleotide sequences) to bind to the junction points of the

mtDNA encoding the respective fusion transcripts listed in Table 2 above. In particular, the probes used in the branched DNA (bDNA) analysis were those listed in Table 5 below.

[0170] Homogenate was then added to the capture plate and incubated overnight at 55°C to enable probe-template hybridization. Following a series of wash and hybridization steps, the chemiluminescent substrate was added. Degradation of alkaline phosphatase conjugated to the probe-template hybrid results in a luminescent signal which is reported as relative luminescence units (RLUs). Each capture plate was read in duplicate and the RLUs were measured on the Promega Glomax™ luminometer. RLU values were analyzed bioinformatically. Duplicate plate reads as well as triplicate values were each averaged, providing they had a CV (coefficient of variation; i.e. the ratio of the standard deviation to the mean) of 15%. Also, it was determined whether or not the RLU values were above background for the given probe. Specifically, the Lower Limit of Quantitation, LOQ (LOQ = the RLU average of the probes' background plus 10 standard deviations of the background average), was calculated and subtracted from the sample RLU. The sample RLU values were then transformed to log2 or log10 values for ease of analysis. Finally, for any given probe, the sample RLU was normalized against the housekeeper (HK) RLU by subtracting or dividing it from the sample RLU. Normalized results testing probes 1, 4, 14, 16, 120, 122, 193, 400, 516 and 586 (wherein the probe numbers correspond with the fusion transcript numbers provided above) on 3 endometriosis-positive and 4 control endometrial samples are shown in FIGs. 2A to 2J and summarized below in Table 5.

[0171] Table 5: Performance of fusion transcripts

Transcript / Probe	Mean Log2LOQProbe - Log2LOQHK23 (Control)	Mean Log2LOQProbe - Log2LOQHK23 (Endo. Pos.)	Mean Difference between Control and Endo. Pos.	Significance, P2	Mean Copy #
1	-14.61	-11.32	-3.29	0.007	46.16
4	-16.54	-10.9	-5.64	0.12	89.99
14	-12.25	-9.2	-3.05	0.01	52.63
16	-13.91	-9.8	-4.12	0.01	64.9
120	-4.15	-0.62	-3.53	0.08	159082.93
122	-3.84	-0.76	-3.08	0.09	159914.28
193	-4.2	-1.45	-2.75	0.11	117468.11
400	-10.9	-7.71	-3.19	0.27	1074.81
586	-9.45	-6.7	-2.76	0.07	2489.43
516	-2.07	0.86	-2.93	0.12	550075.23

[0172] Table 5 indicates the average normalized RLU values (Log2LOQProbe-Log2LOQHK23) for control and endometriosis positive ("Endo. Pos.") tissue samples corresponding to scatter plots shown in FIGs. 2A to 2J. The mean difference between the two tissue groups as well as the significance of the difference is provided. The average number of copies of the given fusion transcript (i.e. probes 1, 4, 14, 16, 400, 586, 120, 122, 193 and 516) is also shown in Table 5.

[0173] qPCR Reactions

[0174] qPCR analysis was performed for transcript numbers 1, 4, 14, 16, 120, 122, 193, 586, 8590, and 2767 (see Table 6 below). Purified DNA extracts were normalized to a concentration of 0.25 ng/μL using nuclease-free ultrapure water. qPCR reactions were set-up at room temperature under dim light using Qiagen's Quantitect™ Sybr Green® PCR kit. 10 μL of template was added to 12.5 μL of the 2X master mix along with 0.025 - 0.0625 μL of each 100 μM forward and reverse primers, depending on the target. Primer sequences were designed for the specific DNA targets and these sequences are shown in Tables 6 (junction primers) and 7 (flanking primers). As used herein, the term "junction primer" will be understood to mean a primer that hybridizes to a region of the target DNA molecule having at least one of the pair of nucleotides forming the junction point after removal of the deletion. Thus, in one aspect, the junction primer may overlap both nucleotides forming the junction point or only one of such nucleotides. As shown in the tables below, more than one set of primers was used in some cases. The reaction was made up to a final volume of 25 μL using PCR-grade H₂O. Reaction mixtures were cycled on either the Chromo 4™ (Biorad) or Opticon 2™ (MJ Research) real-time PCR cyclers.

[0175] Table 6: Junction primer sequences used in qPCR reactions to target large-scale mtDNA deletions

Deletion ID	Direction	Sequence (5'-3')	Positions	Seq. Contribution	Length
1 (8469: 13447)	Fwd	TCTACCCCCTCTAGAGCCC ACTGT (SEQ ID NO: 35)	8277- 8300		24

	Rev	CTAGGCTGCCAATGGTGA GGGAGGT Junction (SEQ ID NO: 36)	13741- 13460 8482- 8470	CTAGGCTGCCAA TGGTGAGGGAGGT	25
4 (set 1) (7992: 15730)	Fwd 7983F	TGCGACTCCTAGCCGCAG ACCTCCTCATTC Junction (SEQ ID NO: 37)	7983- 7992 15730- 15749	TGCGACTCCT AGCCGCAGACCTCCTC ATTC	30
	Rev 16053R	GGTACCCAAATCTGCTTCC CCATGAAAG (SEQ ID NO: 38)	16053- 16026		28
4 (set 2)	Fwd 7983F	TGCGACTCCTAGCCGCAG AC Junction (SEQ ID NO: 39)	7983- 7992 15730- 15739	TGCGACTCCT AGCCGCAGAC	20
	Rev 15935R	(same as set 1)			
4 (set 3)	Fwd	CGCCATCATCCTAGTCCTC ATCGC (SEQ ID NO: 40)	7792- 7815		24
	Rev	GAATGAGGAGGTCTGCGG CTAGGAGTC Junction (SEQ ID NO: 41)	15749- 15730 7992- 7986	GAATGAGGAGGTCTGC GGCT AGGAGTC	27
14 (set 1) (9191: 12909)	Fwd 9178F	GTAAGCCTCTACCTACACT CCAACCTC Junction (SEQ ID NO: 42)	9178- 9191 12909- 12920	GTAAGCCTCTACCT ACACTCCAACCTC	26

	Rev 13122R	GCGGATGAGTAAGAAGATT CCTG (SEQ ID NO: 43)	13122- 13100		23
14 (set 2) (9191: 12909)	Fwd 9178F	(same as set 1)			
	Rev 13024R	GGAGACCTAATTGGGCTG ATTG (SEQ ID NO: 44)	13024- 13012		23
14 (set 3) (9188: 12906)	Fwd	GGCCGTACGCCTAACCGC TA (SEQ ID NO: 45)	8944- 9011		20
	Rev	GTTGTGGGTCTCATGAGTT GGAGTGTAGG Junction (SEQ ID NO: 46)	12934- 12913 9195- 9189	GTTGTGGGTCTCATGA GTTGGA GTGTAGG	29
14a (9188: 12906)	Fwd	CCCTGGCCGTACGCCTAA CC (SEQ ID NO: 47)	8989- 9009		20
	Rev	ATTTGTTGTGGGTCTCATG AGTTGGAGTGTAGG Junction (SEQ ID NO: 48)	12938- 12913 9195- 9189	ATTTGTTGTGGGTCTC ATGAGTTGGA GTGTAGG	33
16 (10367: 12829)	Fwd 10354F	CCCTAAGTCTGGCCAACAC AGCAGC Junction (SEQ ID NO: 49)	10354- 10367 12829- 12839	CCCTAAGTCTGGCC AACACAGCAGC	25

	Rev 13028R	GGGTGGAGACCTAATTGG GCTG (SEQ ID NO: 50)	13028- 13007		22
122 (7973: 9023)	Fwd	CGTCTGAACTATCCTGCCC GC (SEQ ID NO: 53)	7773- 7794		21
	Rev	CAATTAGGTGCATGAGTAG GTGGCCTG Junction (SEQ ID NO: 54)	9049- 9033 7983- 7974	CAATTAGGTGCATGAG T AGGTGGCCTG	27
193 (set 1) (9086: 10313)	Fwd	AAGGCACACCTACACCCCT T (SEQ ID NO: 55)	8918- 8937		20
	Rev	GAGGGATGACATAACTATT AGTGGCAGG Junction (SEQ ID NO: 56)	10334- 10312 9086- 9081	GAGGGATGACATAACT ATTAGT GGCAGG	28
193 (set 2) (9086: 10313)	Fwd	AACCAATAGCCCTGGCCGT A (SEQ ID NO: 57)	8981- 9000		20
	Rev	GAGGGATGACATAACTATT AGTGGCAGGTTA Junction (SEQ ID NO: 58)	10334- 10312 9086- 9078	GAGGGATGACATAACT ATTAGT GGCAGGTTA	31
586 (8431: 10841)	Fwd	CTATAGCACCCCCTCTACC CC (SEQ ID NO: 59)	8264- 8284		21

	Rev	GATGCTAATAATTAGGCTG TGGGTGGT Junction (SEQ ID NO: 60)	10867- 10849 8439- 8432	GATGCTAATAATTAGG CTG TGGGTGGT	27
8590 (8984: 13833)	Fwd	TGCCCTAGCCCACCTTCTTA CC (SEQ ID NO: 80)	8895- 8915		21
	Rev	TAGTTGAGGTCTAGGGCTG TTGGTT Junction (SEQ ID NO: 81)	13853- 13833 8984- 8981	TAGTTGAGGTCTAGGG CTG GGTT	25
2767 (5362: 14049)	Fwd	GGGCCATTATCGAAGAATT CACAAA (SEQ ID NO: 82)	5260- 5284		25
	Rev	GAGGTGATGATGGAGGTG GAGTAG Junction (SEQ ID NO: 83)	14069- 14049 5362- 5360	GAGGTGATGATGGAG GTGGAG TAG	24

- 1
- 2 **[0176]** Table 7: Flanking primer sequences used in qPCR reactions to target large-scale
- 3 mtDNA deletions

Deletion ID	Direction	Sequence (5'-3')	Positions	Length
1 (8469: 13447)	Fwd	ACAGTGAAATGCCCCAACTA (SEQ ID NO: 61)	8358-8377	20
	Rev	GCTCAGGCGTTTGTGTATGA (SEQ ID NO: 62)	13540-13559	20

4 (7992: 15730)	Fwd	CAACGATCCCTCCCTTACCA (SEQ ID NO: 63)	7855-7874	20
	Rev 16053R	AGTACGGATGCTACTTGTCCA (SEQ ID NO: 64)	15796-15816	21
14 (9191: 12909)	Fwd	GAAGCGCCACCCTAGCAATA (SEQ ID NO: 65)	9050-9063	20
	Rev	GGTGAGGCTTGGATTAGCGT (SEQ ID NO: 66)	12950-12969	20
16 (10367: 12829)	Fwd	AATCCACCCCTTACGAGTGC (SEQ ID NO: 67)	10156-10175	20
	Rev	(same as 14)		
120 (6260: 12814)	Fwd	ACAACGTTATCGTCACAGCCC (SEQ ID NO: 68)	6064-6084	21
	Rev	GTGAGGCTTGGATTAGCGTT (SEQ ID NO: 69)	12949-12968	20
120a (6260: 12814)	Fwd	ACAACGTTATCGTCACAGCCCATGC (SEQ ID NO: 51)	6064-6088	25
	Rev	GATTGCTTGAATGGCTGCTGTGTTGGC (SEQ ID NO: 52)	12852-12826	27
122 (7973: 9023)	Fwd	CTGAACCTACGAGTACACCGA (SEQ ID NO: 70)	7900-7920	21
	Rev	GTGTGAAAACGTAGGCTTGGA (SEQ ID NO: 71)	9152-9172	21

193 (9086: 10313)	Fwd	TCGAAACCATCAGCCTACTCA (SEQ ID NO: 72)	8957-8977	21
	Rev	CCAATTCGGTTCAGTCTAATCCT (SEQ ID NO: 73)	10385-10407	23

1

2 **[0177]** *Results and Discussion*

3 **[0178]** As noted above, approximately 268 fusion transcripts were screened in the course of
4 this study, from which 10 endometriosis markers, as discussed herein in more detail, were
5 selected for further study. In particular, as described herein, elevated levels of fusions
6 transcripts associated with deletion ID nos. 1, 4, 14, 16, 120, 122, 193, 400, 516, and 586 (i.e.
7 the transcripts of SEQ ID NOs: 13-15, and 17-23, respectively) in endometrial tissue were found
8 to be associated with endometriosis. The presence of each transcript was determined by
9 assaying for the respective probe having a nucleotide sequence complementary to at least a
10 portion of the transcript having a junction point.

11 **[0179]** The scatterplots and performance of all fusion transcript probes are shown in FIGs. 2A
12 to 2J and in Table 5. FIG. 3 illustrates the locations of the fusion transcripts across the
13 mitochondrial genome, gene locations within the genome and the locations of the 10 mtDNA
14 fusion transcripts of the present invention (i.e., “probes” or “targets”) within the genome are
15 indicated by a line spanning the length of each deletion. Probes corresponding to the
16 aforementioned fusion transcripts were tested against 3 endometriosis positive and 4
17 endometriosis negative endometrium samples (See Table 1). For each sample, the RLU values
18 were normalized against the RLU values obtained for housekeeping gene transcripts HK23
19 (Human Beta-2-microglobulin), HK25 (Human GAPD) and HK18 (Peptidyl-prolyl isomerase B).

20 **[0180]** Based on the results of the present study, it is concluded that fusions transcripts 1, 4, 14,
21 16, 120, 122, 193, 400, 516, and 586 (i.e. the transcripts having the sequences set forth in SEQ
22 ID NOs: 13-15, and 17-23, respectively) can be used in the detection of endometriosis,
23 particularly by assaying endometrial tissue. In particular, in the present investigation, elevated
24 levels of the subject transcripts in endometrial tissue samples have been found to be highly

1 correlated with endometriosis. The detection of the subject fusion transcripts can be achieved
2 using the probes identified above that have nucleotide sequences that are at least substantially
3 complementary to the nucleotide sequences of at least a portion of the respective fusion
4 transcript, wherein such portion includes a junction point, such that the probes hybridize to the
5 respective fusion transcript.

6 **[0181]** Based on these findings, it is also concluded that elevated levels of aberrant mtDNA,
7 having the above-identified deletions 1, 4, 14, 16, 120, 122, 193, 400, 516, and 586 (i.e.
8 deletions having the nucleotide sequences set forth in SEQ ID NOs: 2-4, and 6-12,
9 respectively) can be used in the detection of endometriosis. Such deletions can be identified by
10 identifying the junction point of the parent mtDNA after re-circularization (i.e. the re-circularized
11 large sublimon). The junction points can be identified using probes having nucleotide
12 sequences that are at least substantially complementary to at least a portion of the mtDNA
13 nucleotide sequences including the junction point, such that the probes hybridize to the
14 respective mtDNA. The junction points can also be identified using primers wherein at least one
15 of the primers has a nucleotide sequence that is substantially complementary to the mtDNA
16 nucleotide sequence having the junction point. Alternatively, the primers may comprise pairs
17 having nucleotide sequences that are at least substantially complementary to mtDNA
18 sequences adjacent to the junction point.

19 **[0182]** Similarly, it can be concluded that the deletions can also be identified by identifying the
20 junction point of the deleted sequence after re-circularization (i.e. the re-circularized small
21 sublimon).

22 **[0183]** Translation products from the fusion transcripts (i.e. the fusion proteins having amino
23 acid sequences set forth in SEQ ID NOs: 24-26, 28-34, and 84, respectively) may also be
24 usable for such detection method.

25 **[0184]** Thus, as described herein a method for the detection of endometriosis is provided,
26 wherein the method comprises the use of probes and primers for the identification of the
27 aforementioned fusion transcripts or aberrant mtDNA. These probes and primers have nucleic
28 acid sequences that are complementary to such mitochondrial fusion transcripts and their
29 parent aberrant mtDNA molecules, respectively. In particular, the probes described herein are
30 designed to be at least substantially complementary to fusion transcripts encoding a transcribed

junction point corresponding to the re-joined (or re-circularized) mtDNA. The primers described herein are preferably designed so that one of the primer pairs has a nucleotide sequence that is complementary to a junction point of aberrant re-circularized mtDNA following removal of the deletions described herein. It would also be understood that other primer pairs may be designed wherein one of the primer pairs is at least substantially complementary to the junction point of the re-circularized deletion sequence or where the primer pairs are at least substantially complementary to mtDNA nucleotide sequences adjacent to the junction point.

[0185] Example 2: Detection of mtDNA Deletions in Circulatory Blood Samples

[0186] In this example, mitochondrial DNA, mtDNA, deletions were investigated as potential biomarkers for endometriosis. the study focused primarily on mtDNA deletions obtained from circulatory blood samples. Seven deletions were investigated. Two of these deletions, the “1.2 kb Deletion” and the “3.7 kb Deletion”, discussed further below, were determined to have a high diagnostic accuracy as biomarkers using minimally-invasive blood specimens collected from women of child-bearing potential with symptoms of endometriosis. The 1.2 kb and 3.7 kb deletions were discussed above, wherein the 1.2 kb deletion was identified as deletion “193” and the 3.7 kb deletion was identified as deletion “14” or “14a”. These characteristics of these deletions were summarized earlier in Table 1 but are again provided in Table 8 for convenience. It will be understood that references herein to the “3.7kb deletion” will be understood as references to deletion 14 or deletion 14a.

[0187] Table 8: mtDNA aberrations studied in Example 2

Deletion ID	SEQ ID NO.	Deletion Name	Spliced Genes	mtDNA Location	Junction site (splice location on SEQ ID)
14	4	FUS 9191:12909 (“3.7 kb Deletion”)	(ATP6) to (ND5)	8527-14148	8527-9191/12909-14148 (nucleotides 665-666 of SEQ ID NO: 4)
14a	5	FUS 9188:12906 (“3.7 kb Deletion”)	(ATP6) to (ND5)	8527-14148	8527-9188/12906-14148 (nucleotides 662-663 of SEQ ID NO: 4)

193	8	FUS 9086:10313 ("1.2 kb Deletion")	(ATP6) to (ND3)	8527-10404	8527-9086/10313- 10404 (nucleotides 560-561 of SEQ ID NO: 8)
-----	---	--	-----------------	------------	---

1

2 **[0188]** As discussed previously, the 1.2 kb deletion refers to a deletion of nucleotides 9087-
3 10312 from the wild-type mtDNA genome (SEQ ID NO: 1). Such deletion therefore results in a
4 large sublimon having bases 0-9086 and 10313-16568, which, when re-circularized, has a
5 junction between nucleotides 9086 and 10313. Similarly, the 3.7 kb deletion refers to a deletion
6 of nucleotides 9189-12905, resulting in a large sublimon having bases 0-9188 and 12906-
7 16568, which, when re-circularized, has a junction between nucleotides 9188 and 12906.

8 **[0189]** As discussed above, although re-circularization of the large sublimon has been
9 discussed, such re-circularization of the small sublimon is also possible, with the re-circularized
10 small sublimon having a unique junction point as shown. In the present study, the small
11 sublimons corresponding to the 1.2 kb and 3.7 kb deletions were identified in the course of
12 sequencing samples. Thus, the findings in the present example can be extended to the
13 detection of small sublimons resulting from the deletions described herein.

14 **[0190]** Methods

15 **[0191]** Participants and sample collection

16 **[0192]** This study utilized residual de-identified clinical specimens collected from prospectively
17 enrolled patients as part of the EndOx study at Oxford Endometriosis CaRe Centre, John
18 Radcliffe Hospital, University of Oxford. Briefly, specimens were collected from women
19 scheduled to undergo laparoscopy for suspected endometriosis because of pelvic pain
20 (symptomatic) or tubal ligation (asymptomatic). Study participants were female, aged 18 years
21 or older (until menopause), and were confirmed as not pregnant. All specimens were obtained
22 under a study protocol that received appropriate Ethics Committee approval from the National
23 Research Ethics Service (Oxfordshire REC A, 09/H0604/58). All clinical specimens were
24 anonymized to protect the identity of the source patient. The study was designed, implemented,
25 and reported in accordance with the International Council for Harmonisation, Harmonised
26 Tripartite Guidelines for Good Clinical Practice, with applicable local regulations, and the ethical

principles laid down in the Declaration of Helsinki. All patients gave written informed consent prior to participation.

[0193] Blood specimens and extensive clinical phenotypic data were collected prior to surgery. Study specimens were collected, transported, and stored in accordance with the standardized WERF EPHeCT procedures [26, 41-44].

[0194] Patient populations/study cohorts

[0195] Clinical specimens used in this study were classified as asymptomatic controls, symptomatic controls of surgically confirmed absence of endometriosis, or cases of surgically confirmed endometriosis. Asymptomatic controls were defined as specimens collected from a patient that underwent a scheduled tubal ligation without a clinical suspicion of endometriosis, and surgically confirmed absence of endometriosis. Symptomatic controls were defined as specimens collected from patients having pain or other symptoms (excluding infertility) with a clinical suspicion of endometriosis, but no endometriosis lesions visualized by laparoscopy by experienced gynecological surgeons.

[0196] Endometriosis was scored by the operating surgeon using the revised American Society of Reproductive Medicine (rASRM) classification of endometriosis [45]. Cases were grouped by disease subtype (peritoneal, ovarian, deep endometriosis) and rASRM stage, with stages I through IV representing minimal, mild, moderate, and severe disease, respectively.

[0197] Sample handling, processing, and mtDNA amplification

[0198] DNA extraction

[0199] Total DNA was extracted from 200 μ L of plasma using the QIAamp 96 QIAcubeHT™ extraction kit (Qiagen, Crawley, UK), automated on a QIAcube HT™ system (Qiagen, Crawley, UK). Extracted DNA was eluted in 200 μ L of AE buffer.

[0200] mtDNA deletion real-time qPCR

[0201] Amplification was performed in 20 μ L reactions using a 96-well microplate (Bio-Rad, Hemel Hempstead, UK). Each well contained 5 μ L of un-normalised DNA template, 1 $\times$ SYBR

Green® master mix and 250 nM of the respective primers. The primers used for the reactions are provided in Table 9.

[0202] Table 9: Primer sequences used for amplification of 1.2 kb and 3.7 kb deletions

Deletion ID	Primer Direction	Length	Sequence	SEQ ID NO.
193 ("1.2 kb Deletion")	Fwd	20	AACCAATAGCCCTGGCCGTA	57
	Rev (junction primer)	31	GAGGGATGACATAACTATTAGTGGCAGGTTA	58
14a ("3.7kb Deletion")	Fwd	20	GGCCGTACGCCTAACCGCTA	45
	Rev (junction primer)	29	GTTGTGGGTCTCATGAGTTGGAGTGTAGG	46

[0203] PCR and SYBR Green® I fluorescence were analyzed using a Chromo4™ Real-time PCR Detection System (Bio-Rad, Hemel Hempstead, UK). Cycling conditions for the 1.2 kb deletion were as follows: 3 minutes at 95°C, followed by 5 cycles of 30 seconds at 95°C, 30 seconds at 67°C, and 30 seconds at 72°C; for each subsequent cycle, the annealing temperature was decreased by 0.5°C increments. Amplification conditions were: 45 cycles of 30 seconds at 95°C, 30 seconds at 65°C, and 30 seconds at 72°C. All other deletions were amplified with a standard protocol of 45 cycles of 30 seconds at 95°C, 30 seconds at 58-65°C, and 30 seconds at 72°C. Following amplification, melting curve analysis was performed from 70°C to 90°C, reading every 0.5°C. Each plate of samples and controls was amplified in triplicate on three separate occasions.

[0204] Real-time qPCR normalisation with 18S rRNA

[0205] Target amplicon quantity was normalized using the 18S rRNA nuclear DNA gene. Amplification reactions were performed as 20 µL reactions in a 96-well microplate. Each well contained 5µL of un-normalised DNA template, 1× SYBR Green® master mix, and 200 nM of each primer. Amplification and SYBR Green® I fluorescence was analyzed using a Chromo4 Real-Time PCR Detection System. Amplification conditions were: 3 minutes at 95°C, followed

by 40 cycles of 30 seconds at 95°C, 30 seconds at 64.5°C, and 30 seconds at 72°C. Following amplification, a melting curve analysis was performed from 70°C to 90°C, reading every 0.5°C.

[0206] Quality Control

[0207] The quantification cycle (Cq) was calculated using the CFX manager software regression model (Bio-Rad, Hemel Hempstead, UK). The Cq of each deletion amplicon was normalised to the Cq of the multi-copy nuclear target 18S rRNA gene amplicon. All samples were amplified in triplicate on separate plates and were considered to have passed if at least two of the three replicates were within 1.5 Cq and the melting temperature (T_m) was consistent with the target amplification product when present, (deletion T_m 81°C ± 2°C, 18S rRNA T_m 82°C ± 2°C).

[0208] Two no-template control samples were processed alongside each batch of DNA extractions and verified as negative for amplification of both the deletion target and the 18S rRNA gene. Two no-template control reactions were included on each PCR plate and verified negative for amplification of both the deletion targets and the 18S rRNA gene. Deletion primer specificity was evaluated using rho 0 cellular DNA (to detect mitochondrial pseudogene amplification) as well as DNA from healthy male buccal swabs and DNA extracted from the rho 0 parental cell line (prior to depletion of mitochondria).

[0209] For the initial round of standard PCR reactions, extracted DNA from patients with confirmed endometriosis underwent whole genome amplification using the Repli-G™ mitochondrial DNA kit (Qiagen) to ensure sufficient DNA quantity during this phase.

[0210] Rho 0 cell preparation

[0211] Rho 0 cells were prepared as previously described [46]. Briefly, cells from the human osteosarcoma cell line 143B (ATCC CRL-8303) were treated with ethidium bromide to deplete cytoplasmic mitochondrial DNA. Cells were grown to confluence in high glucose DMEM with pyruvate, L-glutamine, uridine (50 µg/ml) and 5% FBS.

[0212] Statistical analysis

[0213] No formal sample size calculation was performed; the number of clinical specimens used was deemed sufficient to satisfy the study objective.

[0214] For qPCR, targets were amplified from all specimens in triplicate and average Cq values were calculated. The normalised deletion value (ΔCq) was determined by quantifying the deletion amplicon relative to the 18S rRNA reference amplicon. Statistical analyses were performed using Graphpad Prism™ 5.0, (Graphpad software Inc., La Jolla, CA, USA) for receiver operating characteristic (ROC) curves and descriptive statistics. SPSS v17.0 (IBM Corp., Armonk, NY, USA) was used to perform correlations and significance tests. Clinical characteristics were summarized using count and percentages for categorical data, and mean, standard deviation (SD), and range for continuous variables. The means of two groups were compared using the Student's t-test and the Mann–Whitney U test for parametric and non-parametric distributions, respectively. Correlation between the two variables was assessed with the Pearson correlation coefficient (r). With respect to the presence of endometriosis, ROC curves were constructed for all but the 6.5 kb deletion. The area under the curve (AUC) of the ROC and the sensitivity and specificity at selected cut-offs (described below) were calculated with 95% confidence intervals (CIs). A p-value <0.05 was considered statistically significant for all tests.

[0215] Results

[0216] Patient population and clinical specimens

[0217] Demographics and clinical characteristics for patients that provided the clinical specimens used to evaluate the 1.2 kb and 3.7 kb deletions are summarized in Table 10.

[0218] Table 10: Demographic and Clinical Characteristics

Characteristic	1.2 kb deletion cohort N (%)	3.7 kb deletion cohort N (%)
TOTAL	171	181
Patient age ¹	34.2 (6.8)	34.4 (6.9)
Hormone therapy status ²		
Yes	48 (28.1)	55 (30.4)
No	116 (67.8)	119 (65.7)
Undetermined	7 (4.1)	7 (3.9)
Menstrual phase ³		
No menstruation	24 (14.0)	26 (14.4)
Irregular menstruation	12 (7.0)	15 (8.3)
Menstruation	31 (18.1)	31 (17.1)
Follicular phase	44 (25.7)	44 (24.3)
Luteal phase and extended	60 (35.1)	65 (35.9)

CONTROLS	28	32
Patient age ¹	36.6 (6.9)	37.2 (6.8)
Symptomatic	18 (64.3)	19 (58.4)
Asymptomatic	10 (35.7)	13 (40.6)
CASES	143	149
Patient age ¹	33.7 (6.7)	33.8 (6.8)
Endometriosis type		
Peritoneal	49 (34.3)	52 (34.9)
Ovarian	45 (31.5)	47 (31.5)
Deep infiltrating	49 (34.3)	50 (33.6)
Endometriosis stage		
Stage I	63 (44.1)	65 (43.6)
Stage II	21 (14.7)	24 (16.1)
Stage III	29 (20.3)	30 (20.1)
Stage IV	28 (18.6)	28 (18.8)
Unknown	2 (1.4)	2 (1.3)

[0219] Abbreviations: N = number of patients/specimens; SD = standard deviation. (1) Mean (SD) is presented; mean and SD were calculated for patients that provided age at time of specimen collection. (2) Patients' status within 3 months of specimen collection. (3) Patients' menstrual status at time of specimen collection.

[0220] The clinical and demographic characteristics of patients and specimens used for the 1.2 kb and 3.7 kb evaluations were similar and differ as a result of inclusion of only those samples with paired qPCR results for 18S rRNA and each deletion that met the acceptance criteria described previously.

[0221] 1.2 kb deletion cohort

[0222] One hundred seventy-one specimens were used in the 1.2 kb deletion evaluation. The mean (SD) age of patients that provided specimens was 34.2 (6.8) years. The mean ages were similar between the control and case groups with mean (SD) ages of 36.6 (6.9) and 33.7 (6.7) years, respectively and were not statistically significantly different ($p=0.113$). Of the 171 patient specimens used in the evaluation, 116 (67.8%) patients reported no hormone therapy within the three months prior to specimen collection, 48 (28.1%) reported having hormone therapy within three months prior to specimen collection, and 7 (4.1%) were undetermined. Menstrual cycle phase data was calculated using the last menstrual period (LMP) prior to the date of blood collection in relation to a patient's normal cycle length. Twenty-

four (14.0%) patients reported no menstruation - 19 of whom were on hormones, 12 (7.0%) reported irregular menstruation, 31 (18.1%) were in the menstrual phase (between 1 to 5 days from the first day of LMP), 44 (25.7%) were in the follicular phase (5 to 14 days from LMP), and 60 (35.1%) were in the luteal + extended menstrual phase (>15 days from LMP).

[0223] The control group included a total of 28 specimens; 18 (64.3%) specimens collected from symptomatic patients (presenting with symptoms consistent with endometriosis other than infertility and surgical confirmed absence for the disease) and 10 (35.7%) specimens collected from asymptomatic patients scheduled for tubal ligation. The test group included 143 specimens from patients with three disease subtypes (peritoneal, ovarian, and deep infiltrating [DI] endometriosis) that were classified into four stages (rASRM I through IV). Forty-nine (34.3%) specimen were collected from women with peritoneal, 45 (31.5%) were from women with ovarian, and 49 (34.3%) were collected from women with deep endometriosis. Sixty-three (44.1%) specimens were from patients with stage I disease, 21 (14.7%) were stage II, 29 (20.3%) were stage III, and 28 (18.6%) were stage IV. Two (1.4%) specimens had an unknown disease stage.

[0224] 3.7 kb deletion cohort

[0225] One hundred eighty-one specimens were used in the 3.7 kb deletion evaluation. The mean (SD) age of patients that provided specimens was 34.4 (6.9) years. The mean ages were similar between the control and case groups with mean (SD) ages of 37.2 (6.8) and 33.8 (6.8) years, respectively and were not statistically significantly different ($p=0.166$). One hundred nineteen (65.7%) patients reported no hormone therapy within the three months prior to specimen collection, 55 (30.4%) reported having hormone therapy within three months prior to specimen collection, and 7 (3.9%) were undetermined. Twenty-six (14.4%) patients reported no menstruation, 15 (8.3%) reported irregular menstruation, 31 (17.1%) were in the menstrual phase (1 to 5 days), 44 (24.3%) were in the follicular phase (5 to 14 days), and 65 (35.9%) were in the luteal + extended menstrual phase (>15 days).

[0226] The control group included a total of 32 specimens; 19 (58.4%) specimens collected from symptomatic patients and 13 (40.6%) specimens collected from asymptomatic patients. The test group included 149 specimens. Fifty-two (34.9%) specimens were collected from women with peritoneal, 47 (31.5%) were from women with ovarian, and 50 (33.6%) were

collected from women with deep endometriosis. Sixty-five (43.6%) specimens were from women with stage I disease, 24 (16.1%) were stage II, 30 (20.1%) were stage III, and 28 (18.8%) were stage IV. Two (1.3%) specimens had an unknown disease stage.

[0227] mtDNA deletions and preliminary evaluation – standard PCR

[0228] Seven candidate deletions were initially selected for evaluation based upon sequence composition, presence of a flanking repeat location within the major arc of the mitochondrial genome where proportionally more deletions are reported [47] and observation previously in endometrial tissue (data not shown). Deletions were selected within the following genomic regions: CO2 to ATP6 (1.0 kb deletion); ATP6 to ND3 (1.2 kb deletion); ATP8 to ND4 (2.4 kb deletion); ATP6 to ND5 (3.7 kb deletion); ATP8 to ND5 (5.0 kb deletion); CO1 to ND5 (6.5 kb deletion); and CO2 to CytB (7.7 kb deletion). An initial round of standard (qualitative) PCR and visualization after gel electrophoresis was used to pre-qualify each deletion target and determine if each of the candidates: (i) were detectable; (ii) had sufficient copy number for reliable detection; (iii) had the predicted amplicon size; (iv) were specific and did not co-amplify nuclear pseudogenes or generate non-specific amplification products.

[0229] All seven predicted deletions were detectable circulating in blood plasma. However, the 5.0 kb, and the 6.5 kb deletions amplified the rho 0 cell DNA indicating potential co-amplification of nuclear mitochondrial pseudogenes (numts). Additionally, the 6.5 kb deletion had insufficient copy number and was not considered a viable candidate for further QPCR testing. The 7.7 kb and the 2.4 kb deletions were lower in copy number, however still potentially detectable with QPCR so these were subject to further evaluation. The 5.0 kb deletion amplified DNA from the buccal swab of a healthy male indicating a potential lack of disease specificity. The 7.7 kb deletion had a low level of amplification from this specimen as well.

[0230] The remaining six deletions were further evaluated using QPCR to determine whether the targets were i) present in sufficient copy number in the absence of whole genome amplification, ii) of sufficient diagnostic accuracy, iii) detectable in rho 0 cells using more sensitive QPCR, and iv) whether the assays' precision was acceptable. Acceptable precision criteria were a maximum deviation of 1.5 Ct between a minimum of two out of three replicates for each target deletion.

[0231] Preliminary evaluation with clinical samples

[0232] As a preliminary assessment of the remaining six candidates, we evaluated the deletions using a set of 55 clinical specimens; 46 specimens were from patients with confirmed endometriosis and nine specimens were from symptomatic control patients. After initial QPCR testing, the 2.4 kb deletion was determined to have insufficient copy number and the 1.0 kb deletion amplified DNA extracted from rho 0 cells indicating co-amplification of numts and the 7.7 kb deletion amplified only at less stringent annealing temperatures meaning mis-priming events would be more probable. These candidate deletions did not satisfy assay requirements as designed here but may still exist as biomarkers benefiting from further assay optimization to obtain better sequence specificity and assay sensitivity.

[0233] Of the seven deletions initially selected, the 1.2 kb and 3.7 kb deletions were present in sufficient copy number in plasma to facilitate easy and reliable detection. The assays were specific under the tested PCR conditions and accurately discriminated between healthy (asymptomatic) control specimens and specimens from confirmed endometriosis patients (data not shown). In addition, both the 1.2 kb and 3.7 kb deletions were also accurate in discriminating between symptomatic controls and endometrial disease cases (all subtypes and stages combined). The AUC (95% CI) for the 1.2 kb deletion was 0.8116 (0.6178 - 1.005), which was statistically significant ($p=0.0034$). Similarly, the AUC (95% CI) for the 3.7 kb deletion was 0.8478 (0.6663 - 1.029), which was also significant ($p=0.0011$; Table 11).

[0234] Table 11: Preliminary Evaluation of the deletions

Deletion ID (Deletion size)	Standard PCR + visualization					Quantitative real-time PCR (N=55)		
	Detectable in plasma	Sufficient copy number	Target specificity (rho 0)	Target specificity (amplicon size)	Disease specificity	Sufficient copy number	Diagnostic accuracy AUC (95% CI)	p-value
1 (5.0 kb)	Yes	Yes	No	Yes	No	Yes	0.82 (0.64 - 1.00)	
4 (7.7 kb)	Yes	Low	Possible	Yes	Possible	Yes	0.84	
14a (3.7 kb)	Yes	Yes	Yes	Yes	Yes	Yes	0.85 (0.6663 - 1.029)	0.0011

120 (6.5 kb)	Yes	No	No	No	N/A	N/A		
122 (1.0 kb)	Yes	Yes	Yes	Yes	Yes	Yes	0.83 (0.64-1.01)	
193 (1.2 kb)	Yes	Yes	Yes	Yes	Yes	Yes	0.81 (0.6178 - 1.005)	0.0034
586 (2.4 kb)	Yes	Low	Yes	Possible	Yes	No	0.82 (0.66-0.97)	

1

2 **[0235]** Abbreviations: AUC = area under the curve; CI = confidence interval; N = number of
3 specimens in evaluation set; PCR = polymerase chain reaction.

4 **[0236]** Diagnostic accuracy of the 1.2 kb and 3.7 kb deletions

5 **[0237]** To more fully evaluate the 1.2 kb and 3.7 kb deletions as clinically viable biomarkers of
6 endometriosis, we determined the ability of these deletions to discriminate between
7 symptomatic controls and all endometriosis types combined, between three subtypes, and four
8 stages of endometriosis using a larger set of clinical specimens (Table 10). Valid paired results
9 (both target and 18S gene amplification) were obtained for 171 specimens with the 1.2 kb
10 deletion and 181 specimens with the 3.7 kb deletion. These analyses were performed using
11 only symptomatic control and confirmed disease specimens in order to more accurately reflect
12 the clinically relevant patient populations – that is, women presenting with symptoms of
13 endometriosis with surgical confirmation of disease status as an outcome. Importantly, the 1.2
14 kb and 3.7 kb deletions detected no difference between symptomatic and asymptomatic control
15 specimens, $p=0.462$ and $p=0.878$, respectively.

16 **[0238]** Symptomatic controls vs all disease

17 **[0239]** Similar to the preliminary analysis using 55 clinical specimens, both the 1.2 kb and
18 3.7 kb deletions accurately discriminated between symptomatic control and endometrial disease
19 specimens (peritoneal, ovarian, and deep endometriosis specimens combined). The AUC (95%
20 CI) for the 1.2 kb deletion was 0.7879 (0.6791 - 0.8967), which was statistically significant
21 ($p<0.0001$). The AUC (95% CI) for the 3.7 kb deletion was 0.807 (0.7063 - 0.9077), which was

also significant ($p < 0.0001$; FIGs. 4A and 4B). Coordinates of the receiver operating (ROC) curve were examined and a threshold, or cut-off, selected to optimize sensitivity. Applying a threshold of -4.43 for discrimination of symptomatic controls and all subtypes/stages of endometriosis using the 1.2 kb deletion results in sensitivity and specificity values of 81.8% and 72.2%, respectively. At a threshold of 10.51, sensitivity and specificity for the 3.7 kb deletion are 85.1% and 57.9%, respectively (Table 12).

[0240] Table 12: Performance of the 1.2 kb and 3.7 kb deletions

	1.2 kb deletion		3.7 kb deletion	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
All disease	81.8	72.2	85.0	57.9
Peritoneal	91.8	72.2	88.5	73.7
Ovarian	75.6	72.2	80.9	68.4
Deep infiltrating	78.6	66.7	80.0	52.6
Stage I/II	82.1	72.2	87.6	63.2
Stage III/IV	80.7	72.2	84.5	52.6

[0241] Combining the 1.2 kb deletion with the 3.7 kb deletion improved diagnostic accuracy (AUC 0.827 (0.722 – 0.931) between all symptomatic controls and all endometriosis, and AUC 0.882 (0.784 – 0.980) between symptomatic controls and stage I/II disease (data not shown).

[0242] Disease by subtype - 1.2 kb deletion

[0243] An important feature of any diagnostic aid for endometriosis is the ability to accurately detect all disease subtypes. We evaluated the ability of the 1.2 kb deletion to differentiate between symptomatic control specimens and specimens from patients with confirmed peritoneal, ovarian, and deep endometriosis. The distribution of the 1.2 kb deletion for each disease subtype is shown in FIG. 5A. The mean (SD) ΔC_t value was -4.312 (2.075), for symptomatic controls, -7.187 (2.581) for peritoneal disease, -6.291 (2.344), for ovarian disease, and -6.193 (2.143), for deep endometriosis. The difference in normalized 1.2 kb deletion quantity between symptomatic controls was statistically significant for peritoneal ($p < 0.0001$), ovarian ($p = 0.003$), and deep endometriosis ($p = 0.0012$).

[0244] Diagnostic accuracy of the 1.2 kb deletion is shown in FIGs. 5B to 5D, with AUC (95% CI) values of 0.8549 (0.7425-0.9672), $p < 0.0001$ for detection of peritoneal, 0.7457 (0.6118-

0.8796), $p=0.0025$ for detection of ovarian, and 0.7596 (0.6292-0.8901), $p=0.0012$ for detection of deep endometriosis. Taken together, these data indicate that the 1.2 kb deletion was able to accurately distinguish between specimens collected from symptomatic controls and peritoneal, ovarian, and deep endometriosis patients. Applying a threshold of -4.430 for discrimination of symptomatic controls and peritoneal endometriosis using the 1.2 kb deletion results in sensitivity and specificity values of 81.8% and 72.2%, respectively. At a threshold of -4.675, the sensitivity and specificity of the 1.2 kb deletion in discriminating between symptomatic controls and ovarian endometriosis is 75.6% and 72.2%, respectively. At a threshold of -4.350, the sensitivity and specificity of the 1.2 kb deletion in discriminating between symptomatic controls and deep endometriosis is 78.6% and 66.7%, respectively (Table 12).

[0245] Disease by subtype - 3.7 kb deletion

[0246] The distribution of the 3.7 kb deletion for each disease subtype is shown in FIG. 6A. The mean (SD) ΔC_t value was 11.12 (2.239), for symptomatic controls, 7.569 (1.843) for peritoneal, 8.549 (2.089), for ovarian, and 8.617 (2.125) for deep endometriosis. The difference in amplicon quantity between symptomatic controls was statistically significant for peritoneal ($p<0.0001$), ovarian ($p<0.0001$), and deep endometriosis ($p=0.0072$). Diagnostic accuracy of the 3.7 kb deletion in detecting each of the three disease subtypes is shown in FIGs. 6B to 6D, with AUC (95% CI) values of 0.8978 (0.8131-0.9824), $p<0.0001$ for detection of peritoneal, 0.8158 (0.7003-0.9313), $p<0.0001$ for detection of ovarian, and 0.7110 (0.5746-0.8475), $p=0.0071$ for detection of deep endometriosis. Taken together, these data indicate that the 3.7 kb deletion was able to accurately distinguish between specimens collected from symptomatic controls and women with peritoneal, ovarian, and deep endometriosis. Applying a threshold of 8.805 for discrimination of symptomatic controls and peritoneal endometriosis using the 3.7 kb deletion results in sensitivity and specificity values of 88.5% and 73.7%, respectively. At a threshold of 8.910, the sensitivity and specificity of the 3.7 kb deletion in discriminating between symptomatic controls and ovarian endometriosis is 80.9% and 68.4%, respectively. At a threshold of 11.01, the sensitivity and specificity of the 3.7 kb deletion in discriminating between symptomatic controls and deep endometriosis is 80.0% and 52.6%, respectively (Table 12).

[0247] Disease by stage - 1.2 kb deletion

[0248] Another important characteristic of a biomarker for endometriosis is the ability to detect both low and high stages of disease. We next evaluated the ability of the 1.2 kb deletion to differentiate between symptomatic control specimens and specimens from confirmed low (I/II) or high (III/IV) stages of disease. The distribution of the 1.2 kb deletion for stage I/II and stage III/IV disease is shown in FIG. 7A. The mean (SD) ΔC_t value was -4.312 (2.075), for symptomatic controls, -6.692 (2.505) for stage I/II, and -6.348 (2.25), for stage III/IV. The difference between symptomatic controls was statistically significant for stage I/II ($p < 0.0001$), and stage III/IV ($p = 0.001$) disease groups. The difference between stage I/II and III/IV was not statistically significant ($p = 0.406$).

[0249] Diagnostic accuracy of the 1.2 kb deletion is shown in FIGs. 7A to 7C, with AUC (95% CI) values of 0.7989 (0.6868-0.9111), $p < 0.0001$ for detection of stage I/II, and 0.7661 (0.6398-0.8924), $p = 0.0007$ for detection of stage III/IV disease. Thus, the 1.2 kb deletion was able to accurately distinguish between symptomatic controls and all stages of disease. At a threshold of -4.430, the sensitivity and specificity of the 1.2 kb deletion in discriminating between symptomatic controls and stage I/II endometriosis is 82.1% and 72.2%, respectively. At a threshold of -4.490, the sensitivity and specificity of the 1.2 kb deletion in discriminating between symptomatic controls and stage III/IV endometriosis is 80.7% and 72.2%, respectively (Table 12).

[0250] Disease by stage - 3.7 kb deletion

[0251] The distribution of the 3.7 kb deletion for stage I/II and stage III/IV disease is shown in FIG. 8A. The mean (SD) ΔC_t value was 11.12 (2.239), for symptomatic controls, 8.243 (2.156) for stage I/II, and 8.112 (2.14) for stage III/IV. The difference between symptomatic controls was statistically significant for stage I/II ($p < 0.0001$), and stage III/IV ($p = 0.0008$) disease groups. Diagnostic accuracy of the 3.7 kb deletion is shown in FIGs. 8B to 8C, with AUC (95% CI) values of 0.8383 (0.7412-0.9353), $p < 0.0001$ for detection of stage I/II, and 0.7591 (0.6354-0.8837), $p = 0.0007$ for detection of stage III/IV disease. The difference between stage I/II and III/IV was statistically significant for the 3.7 kb deletion ($p = 0.016$). These data indicate that the 3.7 kb deletion was able to accurately distinguish between symptomatic controls and all stages of disease. At a threshold of 10.17, the sensitivity and specificity of the 3.7 kb deletion in discriminating between symptomatic controls and stage I/II endometriosis is 87.6% and 63.2%, respectively. At a threshold of 11.00, the sensitivity and specificity of the 3.7 kb deletion in

discriminating between symptomatic controls and stage III/IV endometriosis is 84.5% and 52.6%, respectively.

[0252] Correlation with patient age, specimen age, hormonal therapy, and menstrual phase – 1.2 kb deletion

[0253] An ideal biomarker test would provide accurate results independent of patient and specimen age, treatment with hormonal therapy, and timing of menstrual phase during specimen collection. The effect of these parameters on disease detection is summarized in Table 13.

[0254] Table 13: Effect of patient and specimen age, hormonal therapy, and menstrual cycle

Parameter	Pearson's correlation (r)		p value		T-test ⁽¹⁾ or ANOVA ⁽²⁾ p value	
	1.2 kb	3.7 kb	1.2 kb	3.7 kb	1.2 kb	3.7 kb
Patient age	0.030	0.1034	0.698	0.166	-	
Specimen age	0.072	0.0628	0.353	0.4009	-	
Hormonal status ⁽¹⁾	-		-		0.120	0.195
Menstrual cycle ⁽²⁾	-		-		0.228	0.036

[0255] (1) T-test was used to determine effect of hormonal status on detection of endometriosis

[0256] (2) ANOVA was used to determine effect of menstrual cycle on detection of endometriosis

[0257] For the 1.2 kb deletion, we determined that there was no correlation with the detection of disease and patient age; the correlation coefficient (r) was $r=0.030$ ($p=0.698$). There was also no correlation with disease detection and specimen age by year of collection ($r=0.072$, $p=0.353$). When stratified by hormonal status (patients that received hormonal therapy or had no hormonal therapy within the 3 months prior to specimen collection) the difference in disease detection was not statistically significant ($p=0.120$). When patients were stratified by menstrual phase (no menstruation, irregular menstruation, menstruation, follicular phase, or luteal phase + extended menstruation) there was no statistically significant difference in disease detection with the 1.2 kb deletion ($p=0.228$).

[0258] Similarly, disease detection based on the 3.7 kb deletion was not significantly correlated to patient age ($r=0.1034$; $p=0.166$) or specimen age ($r=0.0628$; $p=0.4009$) or significantly affected by hormonal therapy ($p=0.195$). Detection of endometriosis with the 3.7 kb deletion was significantly correlated to menstrual phase ($p=0.036$), which was driven specifically by the difference in detection between patients who reported no period and those that were in the follicular phase ($p=0.026$) with a difference of 1.72. Combined, these data indicate that the accuracy of the 1.2 kb and is not significantly affected by these clinically relevant variables and accuracy of the 3.7 kb deletion is slightly affected by menstrual phase.

[0259] Discussion

[0260] Endometriosis is a highly prevalent disease in women of reproductive age that is associated with a large economic burden and results in a substantial reduction in the quality of life of those affected. One of the key contributors to this clinical problem is the lack of diagnostic tools to facilitate early detection and intervention. The current diagnostic gold standard is a thorough laparoscopic inspection ideally followed by histologic confirmation of suspected lesions [5, 15]. Because there are minimal objective data available, the diagnostic value of this process is largely unclear. The use of laparoscopic exams has been considered by some to be potentially inaccurate, and even paired with histologic confirmation accuracy reportedly ranges from 60% to 85% [48-51]. The standard diagnostic process could be further complicated in cases where the disease presents itself atypically, or in early stages of disease that are not easily visualized and overlooked by inexperienced surgeons. Additionally, where medical intervention could be initiated in an effort to avoid or delay surgery, a presumptive diagnosis of endometriosis based upon an accurate biomarker test could provide needed evidence in support of this treatment. Thus, there is a clear need to improve upon the current standard, particularly in a way that could provide more routine results early in the course of disease.

[0261] In the current study, we identified and evaluated two novel mtDNA deletions as potential biomarkers for endometriosis. Assays targeting the 1.2 kb and 3.7 kb deletions met criteria for robust diagnostic tests, utilize a minimally-invasive specimen, and if successfully translated into clinical use, could potentially help reduce the delay in time to diagnosis associated with current diagnostic practices and provide an opportunity for medical intervention prior to a surgical one. After setting a diagnostic threshold (as described above), the sensitivity of the 1.2 kb deletion assay was 81.8% and specificity was 72.2%. The diagnostic performance of the 3.7 kb deletion

1 assay was similar with sensitivity and specificity of 85.1% and 57.9%, respectively. Thus,
2 diagnostic assays based on either of these deletions have the potential to compliment the
3 current standard of care. Of particular importance is the diagnostic accuracy of these deletions
4 for early stage disease as later stage disease is more readily detected in current practice using
5 ultrasound. In a primary care setting a positive test result could support initiating first-line
6 medical treatment for endometriosis such as oral contraceptives or trigger a specialist referral.
7 An estimated 10% of women presenting with dysmenorrhea have secondary dysmenorrhea,
8 with the majority caused by endometriosis [52]. In this population the 1.2 kb and 3.7 kb
9 deletions would quite effectively rule out endometriosis with a negative predictive value (NPV) of
10 97%. In a secondary care setting, a positive test could guide the decision to initiate treatment
11 with second-line medication such as gonadotropin-releasing hormone antagonists or to proceed
12 with laparoscopic surgery. Importantly, in the former setting a diagnostic cut-off could be
13 selected to maximize test sensitivity as the risk associated with an incorrect false positive result
14 is less critical, whereas in the latter setting a different diagnostic cut-off to maximize specificity
15 and minimize the exposure of women without the disease to the risks associated with these
16 interventions could be beneficial.

17 **[0262]** In addition to diagnostic accuracy, the ability of these two biomarkers to detect
18 endometriosis was not correlated to patient age, specimen age, or hormone status, and only the
19 3.7 kb deletion had a slight correlation to the patients' phase of menstrual cycle at the time of
20 specimen collection. Further study is needed to confirm whether this correlation with menstrual
21 phase exists in a larger cohort of patients, or whether it is an artifact of the number of
22 specimens used in this study. We demonstrated that both the 1.2 kb and 3.7 kb deletions
23 accurately detect all subtypes and stages of disease. In contrast to the current diagnostic
24 standard that involves visualization during surgery and possible excision of lesions for
25 histological confirmation, assays based on mtDNA deletions require only a blood specimen and
26 could potentially provide objective results before or instead of a surgical intervention. Thus, if
27 successfully translated into clinical use, mtDNA-based assays have the potential to reduce the
28 delays in diagnosis [16] and provide actionable results earlier in the course of disease than
29 currently possible.

30 **[0263]** From a practical standpoint, use of a blood-based biomarker assay has several
31 advantages to effectively augment the current standard of care. The specimen is easy and
32 inexpensive to collect via venipuncture, and there is a low likelihood to have co-morbidities

1 associated with collection. Blood specimens can be readily collected in an outpatient physicians'
2 office or clinic, eliminating the need for dedicated surgical space and equipment. As a result of
3 the high copy number of mtDNA, standard DNA extraction methods are used without the need
4 for enrichment techniques and ample DNA is recovered from a standard blood specimen so a
5 low test failure rate can be anticipated. The assays use PCR-based technology that is cost-
6 effective and widely used in clinical labs, and while the assays are quantitative, the output is
7 easily interpreted - that is, a test result is either above or below a defined diagnostic cut-off
8 which corresponds to either a positive or negative outcome. Finally, a lack of, or minimal
9 correlation with menstrual stage ensures that sampling requirements are simplified, and timing
10 of menstruation need not be considered when scheduling venepuncture.

11 **[0264]** A key element in successful disease management is understanding disease
12 epidemiology. Due in part to a relatively complex diagnostic process and symptoms that overlap
13 with other gynecological disorders, the epidemiology of endometriosis is not well-characterized
14 and varies across patient populations and geographic locations [1, 2, 4, 6]. With the advent of
15 molecular assays such as those described here, additional data could become more readily
16 available and help fill in some of the gaps in our understanding of endometriosis epidemiology.
17 Importantly, this study utilized publicly available standardized processes for specimen collection
18 and processing, which will allow for more direct comparison of test results across different
19 studies and patient populations [41-44, 53].

20 **[0265]** Importantly, the location of the mtDNA deletions may also help shed light on the
21 pathophysiological process of endometriosis. Both the 1.2 kb and 3.7 kb deletions affect all or
22 part of the genes encoding for Complexes I and V (ATP synthase) of the respiratory chain and
23 several tRNAs. Although these deletions likely result in abnormal mitochondrial ATP synthase
24 and Complex I proteins, the heteroplasmic nature of mtDNA likely allows some degree of
25 functional compensation within the population. Interestingly, the two best candidates out of the
26 seven tested in this study are deletions within regions that overlap each other in the
27 mitochondrial genome. Given that the 1.2 kb deletion region (ATP6 to ND3) resides within the
28 larger 3.7 kb deletion (ATP6 to ND5), perhaps it is not surprising that the diagnostic accuracy of
29 the two deletions is similar.

1 [0266] Based on the data presented here, the 1.2 kb and 3.7 kb mtDNA deletions are
2 associated with endometriosis; however, additional study is necessary to understand what
3 mechanistic role this mitochondrial genomic region plays in the development of endometriosis.

4 [0267] Limitations of this study include the use of patient reported hormone and menstrual
5 status, which can be less accurate than taking study-specific data measurements. This data,
6 while encouraging, requires replication and validation in a larger, independent data set. This
7 study is currently underway.

8 [0268] Summary

9 [0269] The following is a summary of the study discussed above:

10 [0270] • Endometriosis is a significant health burden that affects up to 10% of women
11 worldwide. Currently, diagnosis is based on surgical visualization followed by histological
12 confirmation.

13 [0271] • Diagnosis is often complicated due to variable clinical presentation and
14 symptoms that overlap with other gynecological disorders. As a result, definitive diagnosis can
15 be delayed up to a decade, which can result in higher morbidity and decreased quality of life for
16 those affected.

17 [0272] • Thus, there is a clear need for rapid, reliable diagnostic aids that can provide
18 actionable results early in the course of disease.

19 [0273] • Study specimens were collected from women scheduled to undergo laparoscopy
20 for pelvic pain (symptomatic) or tubal ligation (asymptomatic). Study participants were female,
21 aged 18 years or older (until menopause), and were confirmed as not pregnant.

22 [0274] • Seven candidate mtDNA deletions were identified and evaluated to determine
23 whether each was detectable in plasma, had sufficient copy number for reliable detection, had
24 the predicted amplicon size, were specific and did not co amplify nuclear pseudogenes or
25 generate non-specific amplification products.

26 [0275] • Six candidate deletions were further evaluated by QPCR and clinical specimens
27 to determine whether each met the criteria for a robust diagnostic assay and evaluated

accuracy in discriminating between endometriosis and control specimens. Two deletions were selected as potential biomarker candidates (1.2 kb and 3.7 kb deletions).

[0276] • The 1.2 kb and 3.7 kb deletions accurately detected endometriosis, including all subtypes and disease stages, and detection was not correlated to patient or specimen age or hormone therapy. The 3.7 kb deletion was significantly correlated to menstrual phase, which was limited only to two phases.

[0277] • Biomarkers derived from the mitochondrial genome, including the 1.2 kb and 3.7 kb deletions described here, offer a promising and largely unexplored avenue in the pursuit of diagnostic markers for endometriosis that can be effectively translated to clinical application.

[0278] • Based on a minimally invasive specimen, assays based on these markers could positively impact the diagnostic landscape for endometriosis by reducing the delay in diagnosis and providing rapid, actionable, and objective test results.

[0279] Conclusion

[0280] Biomarkers derived from the mitochondrial genome, in particular the 1.2 kb and 3.7 kb deletions described here, offer a promising and largely unexplored avenue in the pursuit of diagnostic markers for endometriosis that can be effectively translated to clinical application. Based on a minimally invasive specimen, assays based on these markers have been found to accurately diagnose endometriosis in blood samples from patients. Thus, the present description provides a rapid, accurate, and efficient means of diagnosing endometriosis thereby resulting in a reduction in the delay in obtaining a diagnosis and administering the necessary treatment protocol. The present description allows for the subject diagnosis to be performed on one or more of the mtDNA deletion (including either the large or small sublimon) and any fusion transcripts resulting therefrom. The same conclusion may also be extended to any translation products resulting from the fusion transcripts.

[0281] Example 3: Identification of 8.7 kb mtDNA deletion for detecting endometriosis

[0282] In this study, we identified the 8.7 kb deletion (Deletion ID No. 2767) using a combination of next generation sequencing (NGS) and proprietary data mining software in a set of 10 cases and 10 controls obtained from Fidelis Research (Sofia, Bulgaria). The methodology used for this identification is described in more detail in Example 4. We detected this biomarker

directly in endometriosis tissue lesions using both qPCR and NGS. We selected the 8.7 kb deletion for evaluation based upon sequence composition, the presence of a flanking repeat location within the major arc of the mitochondrial genome where proportionally more deletions are reported [47] and observation in endometrial tissue. The data from this study is provided in Table 14 and illustrated in Figures 9 to 11.

[0283] Table 14: Identification of 8.7kb mtDNA deletion

	Endometriosis Positive	Symptomatic Controls	Normal/Healthy Controls
Number of values	14	10	12
Minimum	3.18	4.985	6.55
25% Percentile	3.678	6.04	7.753
Median	5.158	7.028	8.82
75% Percentile	6.808	8.156	10.63
Maximum	7.86	10.44	11.78
Range	4.68	5.455	5.23
Mean	5.431	7.216	9.052
Std. Deviation	1.648	1.526	1.686
Std. Error of Mean	0.4406	0.4824	0.4868
Lower 95% CI of mean	4.479	6.125	7.98
Upper 95% CI of mean	6.383	8.307	10.12
Sum	76.04	72.16	108.6

[0284] The 8.7 kb deletion removes all or part of the genes between NADH dehydrogenase subunits 2-5. An initial round of standard (qualitative) PCR and visualization after gel electrophoresis was used to pre-qualify the deletion target and determine if the deletion: (i) was detectable; (ii) had sufficient copy number for reliable detection; (iii) had the predicted amplicon size; (iv) was specific and did not co-amplify nuclear pseudogenes or generate non-specific amplification products.

[0285] We successfully detected the deletion in circulating plasma and performed further evaluation by qPCR to determine whether the target was detectable in rho 0 cells using more sensitive qPCR. We also evaluated if the assay had sufficient diagnostic accuracy and

1 acceptable precision (defined as a maximum deviation of 1.5 Ct between at least two of three
2 replicates).

3 **[0286]** Further investigation of this deletion is described in Example 4.

4 **[0287]** **Example 4: 8.7 kb mtDNA deletion for detecting endometriosis in plasma of**
5 **symptomatic women**

6 **[0288]** In this example, the 8.7 kb mtDNA deletion (FUS 5362:14049) was investigated as a
7 potential biomarker for diagnosing endometriosis, including i) an initial assessment of diagnostic
8 accuracy followed by ii) an evaluation of disease specificity by comparing the biomarker's
9 frequency in plasma from women with: endometriosis and symptomatic controls, and
10 endometrial cancer, ovarian cancer, and breast cancer.

11 **[0289]** Methods

12 **[0290]** Diagnostic accuracy - participants and sample collection

13 **[0291]** This was a case control study in which residual plasma samples prospectively collected
14 from women aged 18 years and over (until menopause) who were not pregnant and were
15 scheduled to undergo laparoscopy for suspected endometriosis because of pelvic pain
16 (symptomatic controls and endometriosis cases) or tubal ligation (asymptomatic controls) were
17 used. This study was conducted as part of the EndOx study at Oxford Endometriosis CaRe
18 Centre, John Radcliffe Hospital, University of Oxford, UK.

19 **[0292]** The collection, anonymization and processing of samples and data were as previously
20 reported [26, 41-44, 57]. Study conduct, relevant authority approvals (Oxfordshire REC A,
21 09/H0604/58) and consent procedures were also as previously reported [26,41-44; 57].

22 **[0293]** Diagnostic accuracy - participant populations/ cohorts

23 **[0294]** Collected samples were classified as either control or case samples. The control group
24 comprised a) asymptomatic controls, which were specimens collected from participants who
25 underwent scheduled tubal ligation without clinical suspicion of endometriosis, and who had
26 surgically confirmed absence of endometriosis; and b) symptomatic controls, which were
27 collected from participants with pain or other symptoms (excluding infertility) with a clinical

suspicion of endometriosis, but no endometriosis lesions visualized by laparoscopy by experienced gynecological surgeons.

[0295] The case group comprised specimens for which the presence of endometriosis was diagnosed during laparoscopy and classified by the operating surgeon using the revised American Society of Reproductive Medicine (rASRM) stages (I: minimal; II: mild; III: moderate; IV: severe disease) [45]. Specimens were also grouped by disease subtype: peritoneal, ovarian, and deep infiltrating (DI) endometriosis.

[0296] Disease specificity - participants and sample collection

[0297] Endometriosis cases and controls from the diagnostic accuracy assessment were utilized for the assessment of disease specificity and compared to residual plasma samples obtained from OBIO (El Segundo, USA) and Ontario Tumour Bank (Toronto, Canada).

[0298] Sample handling, processing, and mtDNA amplification

[0299] Blood collection and processing

[0300] Whole blood was collected in 10 ml K2EDTA Vacutainers® (BD Medical p/n BD366643) and centrifuged within 1 hour of collection at 2500 x g for 10 minutes at 4°C. The plasma layer was removed, aliquotted and stored at -80°C until DNA extraction.

[0301] DNA extraction

[0302] Total deoxyribonucleic (DNA) was extracted from blood plasma (200 µL) using the QIAamp™ 96 QIAcube™ HT extraction kit (Qiagen, Crawley, UK), automated on a QIAcube™ HT system (Qiagen, Crawley, UK), and eluted extracted DNA with buffer AE (200 µL).

[0303] mtDNA deletion qPCR and qPCR normalization with 18s rRNA

[0304] For both real time polymerase chain reaction (qPCR) procedures, we performed amplification in 20 µL reactions using a 96 well microplate (Bio-Rad, Hemel Hempstead, UK), with each well containing un normalized DNA template (5 µL), SYBR® Green master mix and 250 nM of each primer for the 8.7 kb deletion and the 18S ribosomal ribonucleic acid (rRNA). The primers used are provided in Table 15.

[0305] Table 15: Primer sequences used for amplification of 8.7 kb deletion

Deletion ID	Primer Direction	Length	Sequence	SEQ ID NO
2767 ("8.7 kb Deletion")	Fwd	25	GGGCCATTATCGAAGAATTCACAAA	82
	Rev (junction primer)	24	GAGGTGATGATGGAGGTGGAGTAG	83

[0306] We used a CFX96 Touch Real time PCR Detection System (Bio-Rad, Hemel Hempstead, UK) for quantitative polymerase chain reaction (QPCR) with SYBR Green I fluorescence.

[0307] Cycling conditions for the 8.7 kb deletion and 18S rRNA were: 45 cycles of 30 seconds at 95°C, 30 seconds at 66°C, and 30 seconds at 72°C. After amplification, we performed melting curve analysis from 70°C to 90°C, with a reading every 0.5°C. Each plate of samples and controls was amplified in triplicate on three separate occasions.

[0308] Quality Control

[0309] Quality control was performed as previously described [56]. In brief, we calculated the quantification cycle (Cq) and normalized the Cq of the deletion amplicon to the Cq of the multi-copy nuclear target 18s rRNA gene amplicon. We amplified all samples in triplicate on separate plates. Two no-template control samples were processed alongside each batch of DNA extractions and verified as negative for amplification of both the deletion target and the 18s rRNA gene.

[0310] Rho 0 cell preparation

[0311] Rho 0 cells were prepared as previously described [46; Creed 2019]. In brief, cells from the human osteocarcinoma cell line 143B (ATCC CRL 8303) were treated with ethidium bromide to deplete cytoplasmic mtDNA. Cells were grown to confluence in high glucose Dulbecco's Modified Eagle's Medium with pyruvate, L glutamine, uridine (50 µg/mL) and 5% fetal bovine serum.

[0312] Statistical analysis

[0313] No formal sample size calculation was performed; the number of clinical specimens used was deemed sufficient to satisfy the study objective. For qPCR, targets were amplified from all specimens in triplicate and average Cq values were calculated. We determined the normalized deletion value (ΔCq) by quantifying the deletion amplicon relative to the 18s rRNA reference amplicon. Statistical analyses were performed using Graphpad Prism™ 5.0, (Graphpad Software Inc., La Jolla, CA, USA) for ROCs, descriptive statistics, correlations and significance tests. We summarized clinical characteristics using count and percentages for categorical data, and mean, standard deviation (SD), and range for continuous variables. The means of two groups were compared using the Student's t-test and the Mann–Whitney U test for parametric and non-parametric distributions, respectively. Correlation between the two variables was assessed with the Spearman correlation (r) or the Mann Whitney U-Test or Kruskal-Wallis test. With respect to the presence of endometriosis, ROC curves were constructed. The area under the curve (AUC) of the ROC and the sensitivity and specificity at a selected cut-off was calculated with 95% confidence intervals (CIs). A p value <0.05 was considered statistically significant for all tests.

[0314] Results

[0315] Study population and clinical specimens

[0316] Demographics and clinical characteristics for participants who provided specimens are summarized in Table 16.

[0317] Table 16: Study Population Demographic and Clinical Characteristics

Characteristic	Total N (%)	Controls N (%)	Cases N (%)
N	182	32	150
Patient age ¹	34.4 (±6.9)	37.16 (±6.901)	33.78 (±6.820)
Hormone therapy status ²			
Yes	55 (30.2)	11 (34.5)	44 (29.3)
No	120 (65.9)	19 (59.4)	101 (67.3)
Undetermined	7 (3.8)	2 (6.3)	5 (3.3)
Menstrual phase ³			
No menstruation	27 (14.8)	8 (25.0)	19 (12.7)

Irregular menstruation	15 (8.2)	2 (6.3)	13 (8.7)
Menstruation	30 (16.5)	6 (18.8)	24 (16.0)
Follicular phase	46 (25.3)	3 (9.4)	43 (28.7)
Luteal phase and extended	64 (35.2)	13 (40.6)	51 (34.0)
Non-Endometriosis/Endometriosis type			
Symptomatic	18 (9.9)	18 (56.3)	NA
Asymptomatic	14 (7.7)	14 (43.8)	NA
Peritoneal	52 (28.6)	NA	52 (34.7)
Ovarian	48 (26.4)	NA	48 (32.0)
Deep infiltrating	50 (27.5)	NA	50 (33.3)
Endometriosis stage			
Stage I/II	NA	NA	91 (60.7)
Stage III/IV	NA	NA	58 (38.7)
Unknown	NA	NA	1 (0.7)

- 1
- 2 **[0318]** Abbreviations: N = number of participants/specimens; SD = standard deviation. (1)
- 3 Mean (SD) is presented; mean and SD were calculated for participants that provided age at
- 4 time of specimen collection. (2) Participants' status within 3 months of specimen collection. (3)
- 5 Participants' menstrual status at time of specimen collection.
- 6 **[0319]** Overall, the mean (SD) ages of the control and case groups were statistically
- 7 significantly different: 37.2 (6.9) and 33.8 (6.8) years, $p=0.0124$. The majority of participants
- 8 (121; 66.5%) were not undergoing hormone therapy within the three months before specimen
- 9 collection. Most participants who reported no menstruation were on hormones (20/26; 76.9%).
- 10 **[0320]** Of the 182 specimens collected, 32 were from the control group, with 18 (9.49%) from
- 11 symptomatic participants and 14 (7.7%) from asymptomatic participants. The remaining 150
- 12 specimens were from the case group in which 52 (28.6%) participants had peritoneal, 48
- 13 (26.4%) had ovarian and 50 (27.5%) had DI endometriosis and were classified as 91 (60.7%)
- 14 with rASRM stage I/II disease and 58 (31.9%) with stage III/IV. Of the 182 specimens 178
- 15 (97.8%) produced valid assay results, with 2 peritoneal and 1 ovarian endometriosis samples,

and 1 symptomatic control sample invalid for statistical analysis due to out of range 18S rRNA Cq.

[0321] 8.7 kb mtDNA deletion preliminary evaluation – standard PCR

[0322] As discussed above in Example 3, we previously identified the 8.7 kb deletion using a combination of next generation sequencing (NGS) and proprietary data mining software in a set of 10 cases and 10 controls. As discussed above, we successfully detected the deletion in circulating plasma and performed further evaluation by qPCR to determine whether the target was detectable in rho 0 cells using more sensitive qPCR.

[0323] Diagnostic accuracy of the 8.7 kb deletion

[0324] Having successfully detected the 8.7 kb deletion in circulating plasma and endometriosis lesions, we investigated whether the 8.7 kb deletion could discriminate between symptomatic controls versus all endometriosis; between the three subtypes; and between the revised American Society for Reproductive Medicine (r-ASRM) classification stages, in plasma from a larger set of clinical specimens. We performed the analyses using primarily the symptomatic controls and specimens from participants with confirmed disease to more accurately reflect the clinically relevant patient populations, that is, all presenting with symptoms of endometriosis. We also measured the frequency of the deletion in the asymptomatic control samples and did not detect a difference in the 8.7 kb deletion between symptomatic and asymptomatic ($p=0.681$) control specimens.

[0325] Symptomatic controls versus all disease

[0326] We were able to discriminate well between symptomatic control and all endometriosis specimens using the 8.7 kb assay. The AUC (95% CI) of 0.8007 (0.7035 - 0.8979) was statistically significant ($p<0.0001$). We examined ROC coordinates and chose a threshold to optimize sensitivity, with a threshold of 6.650 discriminating between symptomatic controls and all subtypes/stages of endometriosis and gave acceptable sensitivity and specificity values (Table 17).

[0327] Table 17: Performance of the 8.7 kb deletion at 65% specificity, cut-off 6.65

	AUC [95% CI]	Sensitivity (%) [95% CI]
--	--------------	--------------------------

All disease	0.8007 [0.7035-0.8979]	80.95 [73.85-86.48]
Peritoneal	0.8882 [0.8043-0.9722]	94.00 [83.78-98.36]
Ovarian	0.7766 [0.6572-0.8960]	76.60 [62.78-86.40]
Deep infiltrating	0.7359 [0.6057-0.8661]	72.00 [58.33-82.53]
Stage I/II	0.8361 [0.7426-0.9295]	86.52 [77.00-92.12]
Stage III/IV	0.7465 [0.6232-0.8697]	72.41 [59.80-82.25]

1

2 **[0328]** Detection of disease by subtype

3 **[0329]** It is important that we can accurately detect all disease subtypes of endometriosis. In our
4 study, the 8.7 kb deletion assay differentiated between specimens from symptomatic controls
5 and those from patients with peritoneal, ovarian, and DI endometriosis (Figures 13A to 13D),
6 with mean (SD) Δ Ct values of 6.724 (1.192) for asymptomatic controls, 6.908 (1.26) for
7 symptomatic controls, 4.086 (2.134) for peritoneal disease, 5.283 (1.801) for ovarian disease,
8 and 5.617 (1.767) for DI endometriosis. Furthermore, the difference in normalized 8.7 kb
9 deletion quantity between symptomatic controls was statistically significant for peritoneal
10 ($p < 0.0001$), ovarian ($p = 0.0002$), and DI endometriosis ($p = 0.0023$).

11 **[0330]** Diagnostic accuracy of the 8.7 kb deletion assay was also evaluated for each sub-type
12 (Figures 13A to 13D). We accurately distinguished specimens from symptomatic controls and
13 disease subtypes: AUC (95% CI) was 0.8882 (0.8043-0.9722; $p < 0.0001$) for detection of
14 peritoneal disease, 0.7766 (0.6572-0.8960; $p = 0.0008$) for ovarian, and 0.7359 (0.6057-0.8661;
15 $p = 0.0039$) for DI endometriosis. In addition, the threshold value of 6.65 gave acceptable
16 sensitivity and specificity values for distinguishing between symptomatic controls versus
17 peritoneal endometriosis, ovarian, and DI disease (Table 17).

18 **[0331]** Detection of disease by r-ASRM stage

19 **[0332]** Endometriosis cases were classified into two stage groups, r-ASRM Stage I/II and Stage
20 III/IV to determine if both low and high stage disease would be accurately identified using the
21 8.7 kb deletion. The 8.7 kb deletion assay differentiated between specimens from symptomatic
22 controls and those from patients with low stage (Stage I/II) and high stage (Stage III/IV)
23 (Figures 14A-14C), with mean (SD) Δ Ct values of 6.908 (1.26) for symptomatic controls, 4.614
24 (2.063) for low stage and 5.565 (1.794) for high stage disease.

[0333] Diagnostic accuracy assessed by receiver operating curve was highest for Stage I/II: AUC 0.8361 (0.7426-0.9295; $p < 0.0001$) compared to Stage III/IV: AUC 0.7465 (0.6232-0.8697; $p = 0.0021$). At the threshold of 6.65 sensitivity and specificity for all stages was acceptable (Table 17).

[0334] Correlation with patient age, specimen age, hormonal therapy, and menstrual phase

[0335] For an ideal assay, diagnostic accuracy would not be affected by factors such as patient and specimen age, hormonal therapy, and menstrual phase. We found no correlation between the ΔC_t values and patient age ($p = 0.749$) nor specimen age by year of collection ($p = 0.222$) (Table 3). Similarly, no statistically significant differences in ΔC_t values were seen when we stratified participants by hormonal status ($p = 0.838$) or menstrual phase ($p = 0.233$) (Table 18).

[0336] Table 18: Effect of patient and specimen age, hormonal status and menstrual cycle

Parameter	Spearman correlation (r)	p-value	Kruskal-Wallis test or Mann Whitney U-test p-value
Patient age	0.016	0.749	-
Specimen age	0.092	0.222	-
Hormonal status	-	-	0.838
Menstrual cycle	-	-	0.233

[0337] The Mann-Whitney U-Test was used to determine effect of hormonal status on detection of endometriosis. The Kruskal Wallis was used to determine effect of menstrual cycle on detection of endometriosis.

[0338] Evaluating the disease specificity of the 8.7 kb deletion for endometriosis

[0339] To further assess whether other female diseases had elevated levels of the 8.7 kb deletion we obtained plasma samples from women who were subsequently diagnosed with endometrial cancer ($n = 12$), ovarian cancer ($n = 72$), and breast cancer ($n = 51$) and compared the marker frequency to that of the three endometriosis subtypes (peritoneal, ovarian, and deep infiltrating endometriosis) as well as the symptomatic control group (Figures 14A-14C). Significantly less 8.7 kb deletion was detected in all three cancers with endometrial cancer

estimated as having 64-fold less deletion than endometriosis, ovarian cancer 16-fold less, and breast cancer 8-fold less ($p < 0.0001$). The results from this evaluation are provided in Table 19.

[0340] Table 19: Disease specificity of the 8.7 kb deletion for endometriosis

	Endometrial Cancer	Ovarian Cancer	Breast Cancer	Symptomatic Controls	Peritoneal	Ovarian	DIE
N	12	72	51	17	53	48	50
Mean	12.71	8.921	8.218	6.908	4.265	5.304	5.617
Std. Deviation	2.515	4.239	3.953	1.26	2.202	1.788	1.767
Std. Error of Mean	0.7261	0.4996	0.5536	0.3056	0.3024	0.258	0.2499

[0341] The data from Table 19 is illustrated in Figure 15, which shows normalized 8.7 kb deletion distribution for specimens from endometrial cancer, ovarian cancer, breast cancer, symptomatic controls, and participants with peritoneal, ovarian or deep infiltrating endometriosis.

[0342] Discussion

[0343] In the present study, we demonstrated the utility of measuring the levels of the 8.7 kb deletion biomarker in plasma samples as a potential assay for detecting endometriosis. Our assay met robustness criteria for diagnostic tests, utilized a minimally invasive specimen from blood, and accurately detected all subtypes and stages of disease, with best performance seen in the peritoneal sub-type and low stages of endometriosis, both encountered at high frequency in a primary care setting. When translated into clinical use, this assay could potentially shorten the time to diagnosis and enable medical intervention before surgery. Specimens are easy and inexpensive to collect via venipuncture, with a low possibility of co morbidities associated with this type of collection.

[0344] With good diagnostic accuracy, especially for low stage and peritoneal disease, the 8.7 kb deletion assay has the potential to augment the current standard of care, particularly in the diagnosis of peritoneal disease, which, unlike ovarian and deep infiltrating endometriosis, cannot be reliably detected by imaging modalities. The relative simplicity of the 8.7 kb deletion assay means it could be viable for use in both primary and secondary care settings. Blood

1 samples are routinely collected in primary care without the need for dedicated surgical space or
2 equipment. The high copy number of mtDNA means standard DNA extraction methods can be
3 used without enrichment techniques. Furthermore, a high failure rate for the test is unlikely
4 given the ample amount of DNA recovered from a standard blood specimen. Real time PCR
5 based technology is widely used in clinical laboratories producing easily interpreted and
6 quantitative results.

7 **[0345]** In our study we showed an absence of correlation between the deletion and patient age,
8 specimen age, hormone status, or phase of menstrual cycle. A lack of correlation with
9 menstrual stage simplifies sampling requirements, with no need to consider menstruation phase
10 when scheduling sample collection.

11 **[0346]** The current complex diagnostic process, combined with symptoms that overlap with
12 other gynecological disorders, mean the epidemiology of endometriosis has not been fully
13 characterized, yet this is a key element in the successful management of any condition [1, 2, 4,
14 6]. The advent of molecular assays and novel biomarkers will provide additional, more readily
15 available data to help improve our understanding of endometriosis epidemiology. Importantly,
16 our study used publicly available and standardized methods for collecting and processing
17 specimens, which allows a more direct comparison of test results across studies and patient
18 populations [41-44, 53].

19 **[0347]** Conclusion

20 **[0348]** The assay described above, using the mitochondrial derived 8.7 kb deletion biomarker,
21 is a minimally invasive, blood sample based method for diagnosing endometriosis that may be
22 used in both primary and secondary clinical settings. The relatively simple and more patient
23 friendly approach provided by this assay would shorten the time to diagnosis and thereby
24 improve the management of the debilitating condition described herein and thereby improve
25 patients' quality of life.

26 **[0349]** Example 5: Identification of 4.8 kb mtDNA deletion for detecting endometriosis

27 **[0350]** A study similar to that described above was conducted for identifying a correlation
28 between the frequency of the 4.8 kb mtDNA deletion (Deletion ID 8590) and endometriosis.

The method of Example 4 was followed for this analysis. The primers used for this study are shown in Table 20.

[0351] Table 20: Primer sequences used for amplification of 4.8 kb deletion

Deletion ID	Primer Direction	Length	Sequence	SEQ ID NO
8590 ("4.8 kb Deletion")	Fwd	21	TGCCCTAGCCCACTTCTTACC	80
	Rev (junction primer)	25	TAGTTGAGGTCTAGGGCTGTTGGTT	81

Data illustrating the utility of the 4.8 kb deletion in detecting endometriosis is provided in Table 21 and illustrated in Figures 12-14.

[0353] Table 21: Identification of 4.8 kb mtDNA deletion

	Endo Positive	Symptomatic Controls	Normal Healthy Controls
Number of values	14	10	12
Minimum	1.005	2.93	4.575
25% Percentile	1.67	4.771	5.85
Median	3.033	5.945	6.89
75% Percentile	5.489	6.7	8.224
Maximum	9.47	11.74	8.93
Range	8.465	8.81	4.355
Mean	3.833	6.161	6.899
Std. Deviation	2.537	2.547	1.436
Std. Error of Mean	0.6781	0.8053	0.4147
Lower 95% CI of mean	2.368	4.339	5.986
Upper 95% CI of mean	5.298	7.982	7.811
Sum	53.67	61.61	82.79

Although the above description includes reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art. Any examples provided

herein are included solely for the purpose of illustration and are not intended to be limiting in any way. Any drawings provided herein are solely for the purpose of illustrating various aspects of the description and are not intended to be drawn to scale or to be limiting in any way. The scope of the claims appended hereto should not be limited by the preferred embodiments set forth in the above description but should be given the broadest interpretation consistent with the present specification as a whole. The disclosures of all documents recited herein are incorporated herein by reference in their entirety.

References

- [0355] 1. Eskenazi B, Warner ML. Epidemiology of endometriosis. *Obstet. Gynecol. Clin. North Am.* 24(2), 235–258 (1997).
- [0356] 2. Kjerulff KH, Erickson BA, Langenberg PW. Chronic gynecological conditions reported by US women: findings from the National Health Interview Survey, 1984 to 1992. *Am J Public Health.* 86(2), 195–199 (1996).
- [0357] 3. Houston DE, Noller KL, Melton LJ, Selwyn BJ, Hardy RJ. Incidence of pelvic endometriosis in Rochester, Minnesota, 1970–1979. *Am. J. Epidemiol.* 125(6), 959–969 (1987).
- [0358] 4. Wheeler JM. Epidemiology of endometriosis-associated infertility. *J Reprod Med.* 34(1), 41–46 (1989).
- [0359] 5. Giudice LC, Kao LC. Endometriosis. *Lancet.* 364(9447), 1789–1799 (2004).
- [0360] 6. Vessey MP, Villard-Mackintosh L, Painter R. Epidemiology of endometriosis in women attending family planning clinics. *BMJ.* 306(6871), 182–184 (1993).
- [0361] 7. Zondervan KT, Becker CM, Koga K, Missmer SA, Taylor RN, Viganò P. Endometriosis. *Nature Reviews Disease Primers.* 4(1), 9 (2018).
- [0362] 8. Barlow DH, Kennedy S. Endometriosis: new genetic approaches and therapy. *Annu. Rev. Med.* 56, 345–356 (2005).

- 1 **[0363]** 9. Giudice LC. Clinical practice. Endometriosis. N. Engl. J. Med. 362(25), 2389–
2 2398 (2010).
- 3 **[0364]** 10. Fassbender A, Vodolazkaia A, Saunders P, et al. Biomarkers of endometriosis.
4 Fertil. Steril. 99(4), 1135–1145 (2013). **Review on recent work in biomarkers of endometriosis
- 5 **[0365]** 11. SK, Kantawala KP, Menias CO. Beyond the boundaries-endometriosis: typical
6 and atypical locations. Curr Probl Diagn Radiol. 40(6), 219–232 (2011).
- 7 **[0366]** 12. Berkley KJ, Rapkin AJ, Papka RE. The pains of endometriosis. Science.
8 308(5728), 1587–1589 (2005).
- 9 **[0367]** 13. Nnoaham KE, Hummelshoj L, Webster P, et al. Impact of endometriosis on
10 quality of life and work productivity: a multicenter study across ten countries. Fertil. Steril. 96(2),
11 366-373.e8 (2011).
- 12 **[0368]** 14. Simoens S, Dunselman G, Dirksen C, et al. The burden of endometriosis: costs
13 and quality of life of women with endometriosis and treated in referral centres. Hum. Reprod.
14 27(5), 1292–1299 (2012).
- 15 **[0369]** 15. Dunselman G a. J, Vermeulen N, Becker C, et al. ESHRE guideline:
16 management of women with endometriosis. Hum. Reprod. 29(3), 400–412 (2014).
- 17 **[0370]** 16. Hadfield R, Mardon H, Barlow D, Kennedy S. Delay in the diagnosis of
18 endometriosis: a survey of women from the USA and the UK. Hum. Reprod. 11(4), 878–880
19 (1996).
- 20 **[0371]** 17. Ahn SH, Singh V, Tayade C. Biomarkers in endometriosis: challenges and
21 opportunities. Fertil. Steril. 107(3), 523–532 (2017). **Review on recent work in biomarkers of
22 endometriosis
- 23 **[0372]** 18. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the
24 long and uncertain path to clinical utility. Nat. Biotechnol. 24(8), 971–983 (2006).
- 25 **[0373]** 19. Hu C, van der Heijden R, Wang M, van der Greef J, Hankemeier T, Xu G.
26 Analytical strategies in lipidomics and applications in disease biomarker discovery. J.
27 Chromatogr. B Analyt. Technol. Biomed. Life Sci. 877(26), 2836–2846 (2009).

- 1 **[0374]** 20. Liotta LA, Petricoin EF. Serum peptidome for cancer detection: spinning biologic
2 trash into diagnostic gold. *J. Clin. Invest.* 116(1), 26–30 (2006).
- 3 **[0375]** 21. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to
4 identify ovarian cancer. *Lancet.* 359(9306), 572–577 (2002).
- 5 **[0376]** 22. Petricoin EF, Belluco C, Araujo RP, Liotta LA. The blood peptidome: a higher
6 dimension of information content for cancer biomarker discovery. *Nat. Rev. Cancer.* 6(12), 961–
7 967 (2006).
- 8 **[0377]** 23. Zhang Z, Bast RC, Yu Y, et al. Three biomarkers identified from serum proteomic
9 analysis for the detection of early stage ovarian cancer. *Cancer Res.* 64(16), 5882–5890 (2004).
- 10 **[0378]** 24. Anderson NL, Anderson NG. The human plasma proteome: history, character,
11 and diagnostic prospects. *Mol. Cell Proteomics.* 1(11), 845–867 (2002).
- 12 **[0379]** 25. May KE, Conduit-Hulbert SA, Villar J, Kirtley S, Kennedy SH, Becker CM.
13 Peripheral biomarkers of endometriosis: a systematic review. *Hum. Reprod. Update.* 16(6),
14 651–674 (2010).
- 15 **[0380]** 26. Rahmioglu N, Fassbender A, Vitonis AF, et al. Harmonization of Clinical and
16 Laboratory Data to Improve Biomarker Discovery in Endometriosis: WERF EPHeC. In:
17 Biomarkers for Endometriosis. Springer, Cham, 199–226. ** Summary of standardized methods
18 in endometriosis research
- 19 **[0381]** 27. Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. *Nat.*
20 *Rev. Genet.* 6(5), 389–402 (2005).
- 21 **[0382]** 28. Polyak K, Li Y, Zhu H, et al. Somatic mutations of the mitochondrial genome in
22 human colorectal tumours. *Nat. Genet.* 20(3), 291–293 (1998).
- 23 **[0383]** 29. Petros JA, Baumann AK, Ruiz-Pesini E, et al. mtDNA mutations increase
24 tumorigenicity in prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* 102(3), 719–724 (2005).
- 25 **[0384]** 30. Creed J, Klotz L, Harbottle A, et al. A single mitochondrial DNA deletion
26 accurately detects significant prostate cancer in men in the PSA “grey zone.” *World J Urol.*
27 36(3), 341–348 (2018).

- 1 **[0385]** 31. Parrella P, Xiao Y, Fliss M, et al. Detection of mitochondrial DNA mutations in
2 primary breast cancer and fine-needle aspirates. *Cancer Res.* 61(20), 7623–7626 (2001).
- 3 **[0386]** 32. Liang BC. Evidence for association of mitochondrial DNA sequence amplification
4 and nuclear localization in human low-grade gliomas. *Mutat. Res.* 354(1), 27–33 (1996).
- 5 **[0387]** 33. Guo X-G, Guo Q-N. Mutations in the mitochondrial DNA D-Loop region occur
6 frequently in human osteosarcoma. *Cancer Lett.* 239(1), 151–155 (2006).
- 7 **[0388]** 34. Alonso A, Martin P, Albarran C, et al. Detection of somatic mutations in the
8 mitochondrial DNA control region of colorectal and gastric tumors by heteroduplex and single-
9 strand conformation analysis. *Electrophoresis.* 18(5), 682–685 (1997).
- 10 **[0389]** 35. Liu VW, Shi HH, Cheung AN, et al. High incidence of somatic mitochondrial DNA
11 mutations in human ovarian carcinomas. *Cancer Res.* 61(16), 5998–6001 (2001).
- 12 **[0390]** 36. Dai JG, Xiao YB, Min JX, Zhang GQ, Yao K, Zhou RJ. Mitochondrial DNA 4977
13 BP deletion mutations in lung carcinoma. *Indian J Cancer.* 43(1), 20–25 (2006).
- 14 **[0391]** 37. Liu VWS, Yang HJ, Wang Y, et al. High frequency of mitochondrial genome
15 instability in human endometrial carcinomas. *Br. J. Cancer.* 89(4), 697–701 (2003).
- 16 **[0392]** 38. Wang Y, Xue WC, Liu VWS, Ngan HYS. Detection of mosaic pattern of
17 mitochondrial DNA alterations in different populations of cells from the same endometrial tumor.
18 *Mitochondrion.* 7(1–2), 171–175 (2007).
- 19 **[0393]** 39. Semczuk A, Lorenc A, Putowski L, et al. Clinicoprognostical features of
20 endometrial cancer patients with somatic mtDNA mutations. *Oncol. Rep.* 16(5), 1041–1045
21 (2006).
- 22 **[0394]** 40. Pejovic T, Ladner D, Intengan M, et al. Somatic D-loop mitochondrial DNA
23 mutations are frequent in uterine serous carcinoma. *Eur. J. Cancer.* 40(16), 2519–2524 (2004).
- 24 **[0395]** 41. Rahmioglu N, Fassbender A, Vitonis AF, et al. World Endometriosis Research
25 Foundation Endometriosis Phenome and Biobanking Harmonization Project: III. Fluid
26 biospecimen collection, processing, and storage in endometriosis research. *Fertil. Steril.* 102(5),
27 1233–1243 (2014). *Harmonized methods for endometriosis research

- 1 **[0396]** 42. Fassbender A, Rahmioglu N, Vitonis AF, et al. World Endometriosis Research
2 Foundation Endometriosis Phenome and Biobanking Harmonisation Project: IV. Tissue
3 collection, processing, and storage in endometriosis research. *Fertil. Steril.* 102(5), 1244–1253
4 (2014). *Harmonized methods for endometriosis research
- 5 **[0397]** 43. Becker CM, Laufer MR, Stratton P, et al. World Endometriosis Research
6 Foundation Endometriosis Phenome and Biobanking Harmonisation Project: I. Surgical
7 phenotype data collection in endometriosis research. *Fertil. Steril.* 102(5), 1213–1222 (2014).
8 *Harmonized methods for endometriosis research
- 9 **[0398]** 44. Vitonis AF, Vincent K, Rahmioglu N, et al. World Endometriosis Research
10 Foundation Endometriosis Phenome and Biobanking Harmonization Project: II. Clinical and
11 covariate phenotype data collection in endometriosis research. *Fertil. Steril.* 102(5), 1223–1232
12 (2014). *Harmonized methods for endometriosis research
- 13 **[0399]** 45. American Society for Reproductive Medicine. Revised American Society for
14 Reproductive Medicine classification of endometriosis: 1996. *Fertility and Sterility.* 67(5), 817–
15 821.
- 16 **[0400]** 46. King MP, Attardi G. Human cells lacking mtDNA: repopulation with exogenous
17 mitochondria by complementation. *Science.* 246(4929), 500–503 (1989).
- 18 **[0401]** 47. Brandon MC, Lott MT, Nguyen KC, et al. MITOMAP: a human mitochondrial
19 genome database--2004 update. *Nucleic Acids Res.* 33(Database issue), D611-613 (2005).
- 20 **[0402]** 48. Stratton P, Winkel C, Premkumar A, et al. Diagnostic accuracy of laparoscopy,
21 magnetic resonance imaging, and histopathologic examination for the detection of
22 endometriosis. *Fertil. Steril.* 79(5), 1078–1085 (2003).
- 23 **[0403]** 49. Mettler L, Schollmeyer T, Lehmann-Willenbrock E, et al. Accuracy of
24 laparoscopic diagnosis of endometriosis. *JSLs.* 7(1), 15–18 (2003). *Evaluation of current gold
25 standard for diagnosis of endometriosis
- 26 **[0404]** 50. Falcone T, Mascha E. The elusive diagnostic test for endometriosis. *Fertil. Steril.*
27 80(4), 886–888 (2003).

- 1 **[0405]** 51. Wykes CB, Clark TJ, Khan KS. Accuracy of laparoscopy in the diagnosis of
2 endometriosis: a systematic quantitative review. *BJOG*. 111(11), 1204–1212 (2004).
- 3 **[0406]** 52. Harel Z. Dysmenorrhea in adolescents and young adults: etiology and
4 management. *J Pediatr Adolesc Gynecol*. 19(6), 363–371 (2006).
- 5 **[0407]** 53. Casper RF. Introduction: new tools for enhancing collaborative endometriosis
6 research. *Fertil. Steril*. 102(5), 1211–1212 (2014).
- 7 **[0408]** 54. Salahi, E., et al.; Mitochondrial DNA 4977-bp deletion in endometriosis. *Genes &*
8 *Genomics*, 35(5), 563-567 (2013).
- 9 **[0409]** 55. Taylor HS. Diagnosing endometriosis: Is laparoscopy the gold standard?
10 Supplement to *OBG Management*. S1-S2 (2019).
- 11 **[0410]** 56. Ballard K, Lowton K, Wright J. What's the delay? A qualitative study of women's
12 experiences of reaching a diagnosis of endometriosis. *Fertil. Steril* 86(5), 1296-301 (2006).
- 13 **[0411]** 57. Creed J, Maggrah A, Reguly B, Harbottle A. Mitochondrial DNA deletions
14 accurately detect endometriosis in symptomatic females of child-bearing age. *Biomark. Med.*
15 13(4), 291–306 (2019).
- 16 **[0412]** 58. The Practice Committee of the American Society for Reproductive Medicine.
17 Endometriosis and infertility. *Fertil. Steril*. 81, 1441-1446 (2004).
- 18 **[0413]** 59. Nicholls, T.J., et al., Linear mtDNA fragments and unusual mtDNA
19 rearrangements associated with pathological deficiency of MGME1 exonuclease. *Hum Mol*
20 *Genet.*, 23(23):6147-62 (2014).
- 21 **[0414]** 60. Zhang C. et al., Multiple mitochondrial DNA deletions in an elderly human
22 individual. *FEBS*, 297, 1, 2, 34-38 (1992).
- 23 **[0415]** 61. Wei, Y.H., Mitochondrial DNA alterations as ageing-associated molecular events.
24 *Mutat. Res.*, 275(3-6), 145-55 (1992).
- 25 **[0416]** 62. Tanaka, M. et al., Direct sequencing of deleted mitochondrial DNA in myopathic
26 patients. *Biochem Biophys Res Commun.*, 164(1):156-63, (1989).

- 1 **[0417]** 63. Kapsa, R. et al., A novel mtDNA deletion in an infant with Pearson syndrome. J
2 Inherit Metab Dis., 17(5):521-6, (1974).
- 3 **[0418]** 64. Hseih, R.H., et al., Multiple rearrangements of mitochondrial DNA in unfertilized
4 human oocytes. Fertil Steril., 77(5):1012-7, (2002).
- 5 **[0419]** 65. Hjelm, B.E. et al., [http://mitobreak.portugene.com/cgi-](http://mitobreak.portugene.com/cgi-bin/Mitobreak_individual_page.cgi?Name=5371:14058&Breaktype=Del&Species=Homsap&Published=Pub)
6 [bin/Mitobreak_individual_page.cgi?Name=5371:14058&Breaktype=Del&Species=Homsap&Pub](http://mitobreak.portugene.com/cgi-bin/Mitobreak_individual_page.cgi?Name=5371:14058&Breaktype=Del&Species=Homsap&Published=Pub)
7 [lished=Pub](http://mitobreak.portugene.com/cgi-bin/Mitobreak_individual_page.cgi?Name=5371:14058&Breaktype=Del&Species=Homsap&Published=Pub)
- 8

CLAIMS:

1. A method of identifying, in a biological sample from a mammalian subject, an aberrant mitochondrial DNA, mtDNA, molecule having a deletion, wherein the deletion comprises a nucleotide sequence between nucleotides 5362-14049; 8469-13447; 7992-15730; 9191-12909; 9188-12906; 10367-12829; 6260-12814; 7973-9023; 9086-10313; 9079-14988; 7260-15540; 8431-10841; or 8984-13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1, and wherein, once re-circularized, the mtDNA includes a junction point.

2. A method of identifying, in a biological sample from a mammalian subject, an aberrant mitochondrial DNA, mtDNA, molecule having a deletion, wherein once re-circularized, the mtDNA includes a junction point consisting of first and second nucleotides, and wherein, with respect to SEQ ID NO: 1:

- a) the deletion includes nucleotides 5377-14048, the first nucleotide is between nucleotides 5362-5377 and the second nucleotide is between nucleotides 14048-14063;
- b) the deletion includes nucleotides 8483-13446, the first nucleotide is between nucleotides 8469-8483 and the second nucleotide is between nucleotides 13446-13460;
- c) the deletion includes nucleotides 7993-15722, the first nucleotide is between nucleotides 7985-7993 and the second nucleotide is between nucleotides 15722-15730;
- d) the deletion includes nucleotides 9196-12908, the first nucleotide is between nucleotides 9191-9196 and the second nucleotide is between nucleotides 12908-12912;
- e) the deletion includes nucleotides 9196-12905, the first nucleotide is between nucleotides 9188-9196 and the second nucleotide is between nucleotides 12905-12913;
- f) the deletion includes nucleotides 10368-12825, the first nucleotide is between nucleotides 10364-10368 and the second nucleotide is between nucleotides 12825-12829;
- g) the deletion includes nucleotides 6261-12813, the first nucleotide is between nucleotides 6260-6271 and the second nucleotide is between nucleotides 12813-12824;
- h) the deletion includes nucleotides 7984-9022, the first nucleotide is between nucleotides 7973-7984 and the second nucleotide is between nucleotides 9022-9033;
- i) the deletion includes nucleotides 9087-10303, the first nucleotide is between nucleotides 9077-9087 and the second nucleotide is between nucleotides 10303-10313;
- j) the deletion includes nucleotides 9086-14987, the first nucleotide is between nucleotides 9079-9086 and the second nucleotide is between nucleotides 14987-14904;

k) the deletion includes nucleotides 7261-15531, the first nucleotide is between nucleotides 7252-7261 and the second nucleotide is between nucleotides 15531-15540;

l) the deletion includes nucleotides 8440-10840, the first nucleotide is between nucleotides 8431-8440 and the second nucleotide is between nucleotides 10840-10849; or,

m) the deletion includes nucleotides 8994-13832, the first nucleotide is between nucleotides 8984-8994 and the second nucleotide is between nucleotides 13832-13842.

3. The method of claim 2, wherein:

a) the deletion includes nucleotides 5377-14048, the first nucleotide is at position 5362 and the second nucleotide is at position 14049;

b) the deletion includes nucleotides 8483-13446, the first nucleotide at position 8469 and the second nucleotide is a position 13447;

c) the deletion includes nucleotides 7993-15722, the first nucleotide is at position 7992 and the second nucleotide is at position 15730;

d) the deletion includes nucleotides 9196-12908, the first nucleotide is at position 9191 and the second nucleotide is at position 12909;

e) the deletion includes nucleotides 9196-12905, the first nucleotide is at position 9188 and the second nucleotide is at position 12906;

f) the deletion includes nucleotides 10368-12825, the first nucleotide is at position 10367 and the second nucleotide is at position 12829;

g) the deletion includes nucleotides 6261-12813, the first nucleotide is at position 6260 and the second nucleotide is at position 12814;

h) the deletion includes nucleotides 7984-9022, the first nucleotide is at position 7973 and the second nucleotide is at position 9023;

i) the deletion includes nucleotides 9087-10303, the first nucleotide is at position 9086 and the second nucleotide is at position 10313;

j) the deletion includes nucleotides 9086-14987, the first nucleotide is at position 9079 and the second nucleotide is at position 14988;

k) the deletion includes nucleotides 7261-15531, the first nucleotide is at position 7260 and the second nucleotide is at position 15532;

l) the deletion includes nucleotides 8440-10840, the first nucleotide is at position 8431 and the second nucleotide is at position 10841; or,

m) the deletion includes nucleotides 8994-13832, the first nucleotide is at position 8984 and the second nucleotide is at position 13833.

4. A method of identifying, in a biological sample from a mammalian subject, an aberrant mitochondrial DNA, mtDNA, molecule wherein the mtDNA molecule has a deletion in its nucleic acid sequence and, once re-circularized, the mtDNA includes a junction point at nucleotides 5362:14049; 8469:13447; 7992:15730; 9191:12909; 9188:12906; 10367:12829; 6260:12814; 7973:9023; 9086:10313; 9079:14988; 7260:15540; 8431:10841; or 8984:13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1.

5. The method of any one of claims 1 to 4, wherein the aberrant mtDNA comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 75, 2 to 12 and 74.

6. The method of any one of claims 1 to 5, wherein the identification comprises contacting the biological sample with a nucleic acid probe having a nucleotide sequence at least substantially complementary to a portion of the nucleotide sequence of the aberrant mtDNA comprising the junction point.

7. The method of any one of claims 1 to 5, wherein the identification comprises contacting the biological sample with a nucleic acid primer pair, wherein one of the primers has a nucleotide sequence complementary to a portion of the nucleotide sequence of the aberrant mtDNA comprising the junction point.

8. The method of claim 7, wherein the one of the primers of the pair of primers has the nucleotide sequence selected from SEQ ID NO: 83, 36, 37, 39, 41, 42, 44, 46, 48, 49, 54, 56, 58, 60, or 81.

9. The method of any one of claims 1 to 5, wherein the identification comprises contacting the biological sample with a nucleic acid primer pair, wherein each of the primers has a nucleotide sequence complementary to nucleotide sequences of the aberrant mtDNA adjacent to the junction point.

10. The method of claim 9, wherein the primer pairs comprise: SEQ ID NOs: 61 and 62; SEQ ID NOs: 63 and 64; SEQ ID NOs: 65 and 66; SEQ ID NOs: 67 and 66; SEQ ID NOs: 68 and 69; SEQ ID NOs: 70 and 71; or, SEQ ID NOs: 72 and 73.

11. A method of identifying, in a biological sample from a mammalian subject, a fusion transcript encoded by at least a portion of an aberrant mitochondrial DNA, mtDNA, molecule having a deletion, wherein the deletion comprises nucleotides 5362-14049; 8469-13447; 7992-15730; 9191-12909; 9188-12906; 10367-12829; 6260-12814; 7973-9023; 9086-10313; 9079-14988; 7260-15540; 8431-10841; or 8984-13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1.

12. A method of identifying, in a biological sample from a mammalian subject, a fusion transcript encoded by at least a portion of an aberrant mitochondrial DNA, mtDNA, molecule having a deletion resulting in a junction point after the mtDNA is re-circularized, wherein the junction point is between nucleotides 5362:14049; 8469:13447; 7992:15730; 9191:12909; 9188:12906; 10367:12829; 6260:12814; 7973:9023; 9086:10313; 9079:14988; 7260:15540; 8431:10841; or 8984:13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1.

13. A method of identifying, in a biological sample from a mammalian subject, a fusion transcript encoded by at least a portion of an aberrant mitochondrial DNA, mtDNA, molecule as defined in any one of claims 1 to 5.

14. The method of claim 13, wherein the method comprises contacting the biological sample with a nucleic acid probe having a nucleic acid sequence that is complementary to the nucleotide sequence of the fusion transcript having the transcribed junction point.

15. The method of any one of claims 11 to 14, wherein the fusion transcript comprises the nucleotide sequence set forth in any one of SEQ ID NOs: 77, 13 to 23, and 76.

16. The method of any one of claims 11 to 15, wherein the step of identification of the fusion transcript comprises identifying a translation product of the fusion transcript.

17. The method of claim 16, wherein the translation product has the amino acid sequence set forth in any one of SEQ ID NOs: 79, 24 to 34, 84, and 78.

18. A method of identifying, in a biological sample from a mammalian subject, a fusion protein comprising an amino acid sequence set forth in any one of SEQ ID NOs: 79, 24 to 34, 84, and 78.

19. The method of any one of claims 16 to 18, wherein the identification of the protein comprises an immunological assay.

20. The method of claim 19, wherein the identification comprises an antibody specific to the protein.

21. The method of any one of claims 1 to 10, wherein the step of identification comprises identifying a fusion protein encoded by the aberrant mtDNA.

22. The method of any one of claims 1 to 21, wherein the step of identification further comprises detecting, diagnosing, and/or monitoring endometriosis in the mammalian subject.

23. The method of any of claims 1 to 22, further comprising quantifying, in the biological sample, the amount of the aberrant mtDNA.

24. The method of claim 23, further comprising comparing the quantified amount of aberrant mtDNA to a reference value indicative of the presence of endometriosis or the development of endometriosis in the subject.

25. The method of claim 24, wherein, prior to the step of comparing, the quantified amount of aberrant mtDNA is normalized against a nuclear DNA sequence.

26. The method of claim 25, wherein the nuclear DNA sequence encodes rRNA.

27. The method of claim 25, wherein the rRNA is 18S rRNA.

28. The method of any one of claims 1 to 27, wherein the subject is a member of a sub-population suspected of having endometriosis or a member of a subpopulation not suspected of having endometriosis.
29. The method of any one of claims 1 to 28, wherein the biological sample is one or more of: blood; a blood derivative; tissue; and menstrual fluid.
30. The method of claim 29, wherein the blood derivative is plasma and/or serum.
31. A method of identifying, in a biological sample from a mammalian subject, a deleted mitochondrial DNA, mtDNA, molecule, wherein the deletion comprises nucleotides 5362-14049; 8469-13447; 7992-15730; 9191-12909; 9188-12906; 10367-12829; 6260-12814; 7973-9023; 9086-10313; 9079-14988; 7260-15540; 8431-10841; or 8984-13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1.
32. The method of claim 31, wherein the step of identification further comprises detecting, diagnosing, and/or monitoring endometriosis in the mammalian subject.
33. The method of claim 31 or 32, further comprising quantifying, in the biological sample, the amount of the deleted mtDNA.
34. The method of claim 33, further comprising comparing the quantified amount of the deleted mtDNA to a reference value indicative of the presence of endometriosis or the development of endometriosis in the subject.
35. The method of any one of claims 31 to 34, wherein the biological sample is one or more of: blood; a blood derivative; tissue; and menstrual fluid.
36. The method of any one of claims 1 to 35, wherein the method is conducted in vitro.
37. A kit for conducting the method according to any one of claims 1 to 36.

38. The kit of claim 37, wherein the kit comprises a nucleic acid primer pair, wherein one of the primers has a nucleotide sequence complementary to a portion of the nucleotide sequence of the aberrant mtDNA comprising the junction point.

39. The kit of claim 38, wherein the one of the primers of the pair of primers has the nucleotide sequence selected from SEQ ID NO: 83, 36, 37, 39, 41, 42, 44, 46, 48, 49, 54, 56, 58, 60, or 81.

40. The kit of claim 37, wherein the kit comprises a nucleic acid primer pair, wherein each of the primers has a nucleotide sequence complementary to nucleotide sequences of the aberrant mtDNA adjacent to the junction point.

41. The kit of claim 40, wherein the primer pairs comprise: SEQ ID NOs: 61 and 62; SEQ ID NOs: 63 and 64; SEQ ID NOs: 65 and 66; SEQ ID NOs: 67 and 66; SEQ ID NOs: 68 and 69; SEQ ID NOs: 70 and 71; or, SEQ ID NOs: 72 and 73.

42. The kit of claim 37, wherein the kit comprises primers and/or probes complementary to one or more fusion transcripts of the aberrant mtDNA molecules.

43. The kit of claim 37, wherein the kit comprises binding agents that are adapted to bind to proteins encoded by the aberrant mtDNA molecules.

44. A kit for in vitro identification of an aberrant mitochondrial DNA, mtDNA, molecule in a biological sample, wherein the mtDNA has a deletion of a nucleotide sequence between nucleotides 5362-14049; 8469-13447; 7992-15730; 9191-12909; 9188-12906; 10367-12829; 6260-12814; 7973-9023; 9086-10313; 9079-14988; 7260-15540; 8431-10841; or 8984-13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1, and wherein, once re-circularized, the mtDNA includes a junction point.

45. A kit for in vitro identification of an aberrant mitochondrial DNA, mtDNA, molecule in a biological sample, wherein the mtDNA molecule has a deletion, and wherein once re-circularized, the mtDNA includes a junction point consisting of first and second nucleotides, and wherein, with respect to SEQ ID NO: 1:

- a) the deletion includes nucleotides 5377-14048, the first nucleotide is between nucleotides 5362-5377 and the second nucleotide is between nucleotides 14048-14063;
- b) the deletion includes nucleotides 8483-13446, the first nucleotide is between nucleotides 8469-8483 and the second nucleotide is between nucleotides 13446-13460;
- c) the deletion includes nucleotides 7993-15722, the first nucleotide is between nucleotides 7985-7993 and the second nucleotide is between nucleotides 15722-15730;
- d) the deletion includes nucleotides 9196-12908, the first nucleotide is between nucleotides 9191-9196 and the second nucleotide is between nucleotides 12908-12912;
- e) the deletion includes nucleotides 9196-12905, the first nucleotide is between nucleotides 9188-9196 and the second nucleotide is between nucleotides 12905-12913;
- f) the deletion includes nucleotides 10368-12825, the first nucleotide is between nucleotides 10364-10368 and the second nucleotide is between nucleotides 12825-12829;
- g) the deletion includes nucleotides 6261-12813, the first nucleotide is between nucleotides 6260-6271 and the second nucleotide is between nucleotides 12813-12824;
- h) the deletion includes nucleotides 7984-9022, the first nucleotide is between nucleotides 7973-7984 and the second nucleotide is between nucleotides 9022-9033;
- i) the deletion includes nucleotides 9087-10303, the first nucleotide is between nucleotides 9077-9087 and the second nucleotide is between nucleotides 10303-10313;
- j) the deletion includes nucleotides 9086-14987, the first nucleotide is between nucleotides 9079-9086 and the second nucleotide is between nucleotides 14987-14904;
- k) the deletion includes nucleotides 7261-15531, the first nucleotide is between nucleotides 7252-7261 and the second nucleotide is between nucleotides 15531-15540;
- l) the deletion includes nucleotides 8440-10840, the first nucleotide is between nucleotides 8431-8440 and the second nucleotide is between nucleotides 10840-10849; or,
- m) the deletion includes nucleotides 8994-13832, the first nucleotide is between nucleotides 8984-8994 and the second nucleotide is between nucleotides 13832-13842.

46. The kit of claim 45, wherein:

- a) the deletion includes nucleotides 5377-14048, the first nucleotide is at position 5362 and the second nucleotide is at position 14049;
- b) the deletion includes nucleotides 8483-13446, the first nucleotide at position 8469 and the second nucleotide is a position 13447;

- c) the deletion includes nucleotides 7993-15722, the first nucleotide is at position 7992 and the second nucleotide is at position 15730;
- d) the deletion includes nucleotides 9196-12908, the first nucleotide is at position 9191 and the second nucleotide is at position 12909;
- e) the deletion includes nucleotides 9196-12905, the first nucleotide is at position 9188 and the second nucleotide is at position 12906;
- f) the deletion includes nucleotides 10368-12825, the first nucleotide is at position 10367 and the second nucleotide is at position 12829;
- g) the deletion includes nucleotides 6261-12813, the first nucleotide is at position 6260 and the second nucleotide is at position 12814;
- h) the deletion includes nucleotides 7984-9022, the first nucleotide is at position 7973 and the second nucleotide is at position 9023;
- i) the deletion includes nucleotides 9087-10303, the first nucleotide is at position 9086 and the second nucleotide is at position 10313;
- j) the deletion includes nucleotides 9086-14987, the first nucleotide is at position 9079 and the second nucleotide is at position 14988;
- k) the deletion includes nucleotides 7261-15531, the first nucleotide is at position 7260 and the second nucleotide is at position 15532;
- l) the deletion includes nucleotides 8440-10840, the first nucleotide is at position 8431 and the second nucleotide is at position 10841; or,
- m) the deletion includes nucleotides 8994-13832, the first nucleotide is at position 8984 and the second nucleotide is at position 13833.

47. A kit for in vitro identification of an aberrant mitochondrial DNA, mtDNA, molecule in a biological sample, wherein, the mtDNA has a deletion and, once re-circularized, the mtDNA includes a junction point between nucleotides 5362:14049; 8469:13447; 7992:15730; 9191:12909; 9188:12906; 10367:12829; 6260:12814; 7973:9023; 9086:10313; 9079:14988; 7260:15540; 8431:10841; or 8984:13833 of the mtDNA nucleotide sequence of SEQ ID NO: 1.

48. The kit of any one of claims 44 to 47, wherein the kit comprises a nucleic acid primer pair, wherein one of the primers has a nucleotide sequence complementary to a portion of the nucleotide sequence of the aberrant mtDNA comprising the junction point.

49. The kit of claim 48, wherein the one of the primers of the pair of primers has the nucleotide sequence SEQ ID NO: 83, 36, 37, 39, 41, 42, 44, 46, 48, 49, 54, 56, 58, 60, or 81.

50. The kit of any one of claims 44 to 47, wherein the kit comprises a nucleic acid primer pair, wherein each of the primers has a nucleotide sequence complementary to nucleotide sequences of the aberrant mtDNA adjacent to the junction point.

51. The kit of claim 50, wherein the primer pairs comprise: SEQ ID NOs: 61 and 62; SEQ ID NOs: 63 and 64; SEQ ID NOs: 65 and 66; SEQ ID NOs: 67 and 66; SEQ ID NOs: 68 and 69; SEQ ID NOs: 70 and 71; or, SEQ ID NOs: 72 and 73.

52. The kit of any one of claims 44 to 47, wherein the kit comprises primers and/or probes complementary to one or more fusion transcripts of the aberrant mtDNA molecules.

53. The kit of any one of claims 44 to 47, wherein the kit comprises binding agents that are adapted to bind to proteins encoded by the aberrant mtDNA molecules.

54. The kit of claim 54, wherein the binding agents comprise antibodies or antibody fragments.

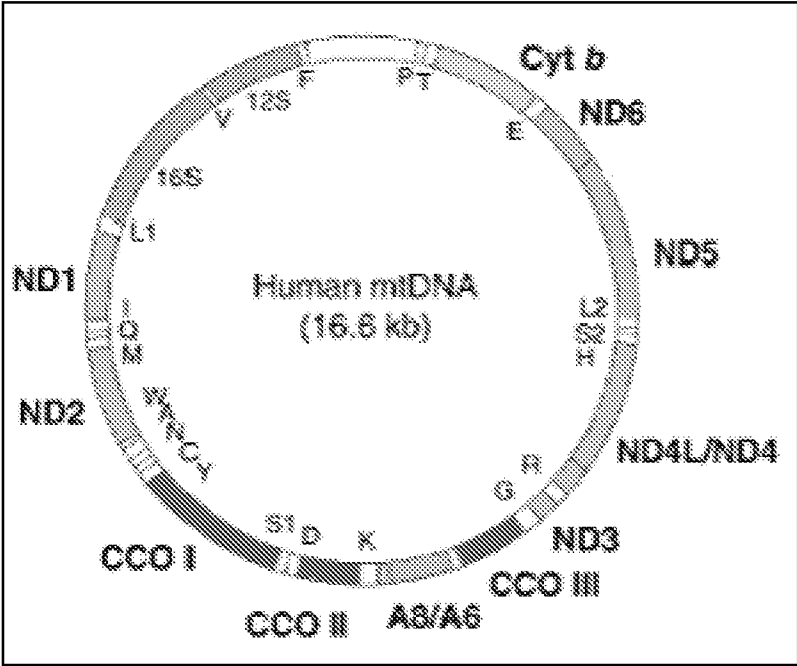


FIG. 1

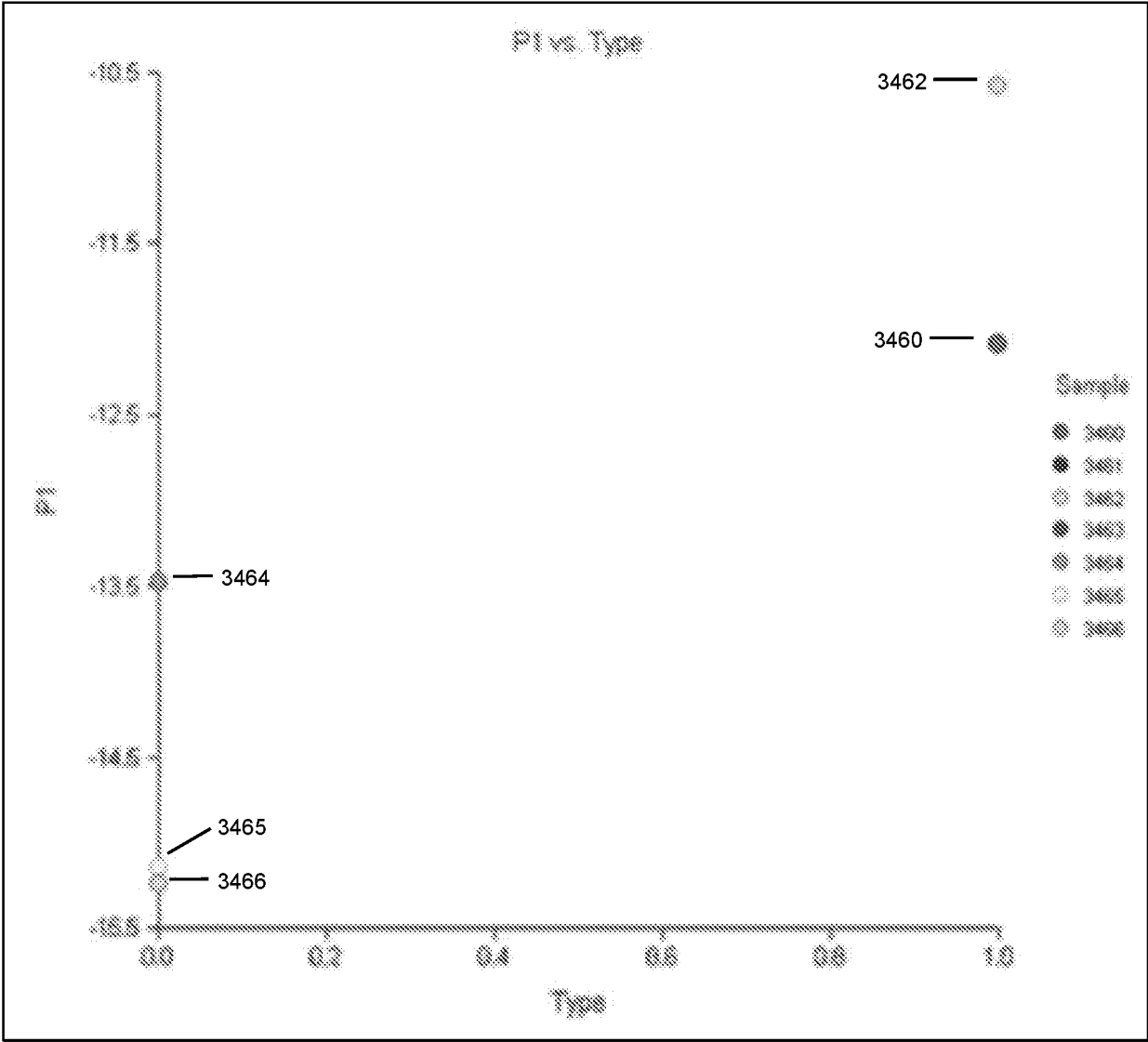


FIG. 2A

3/35

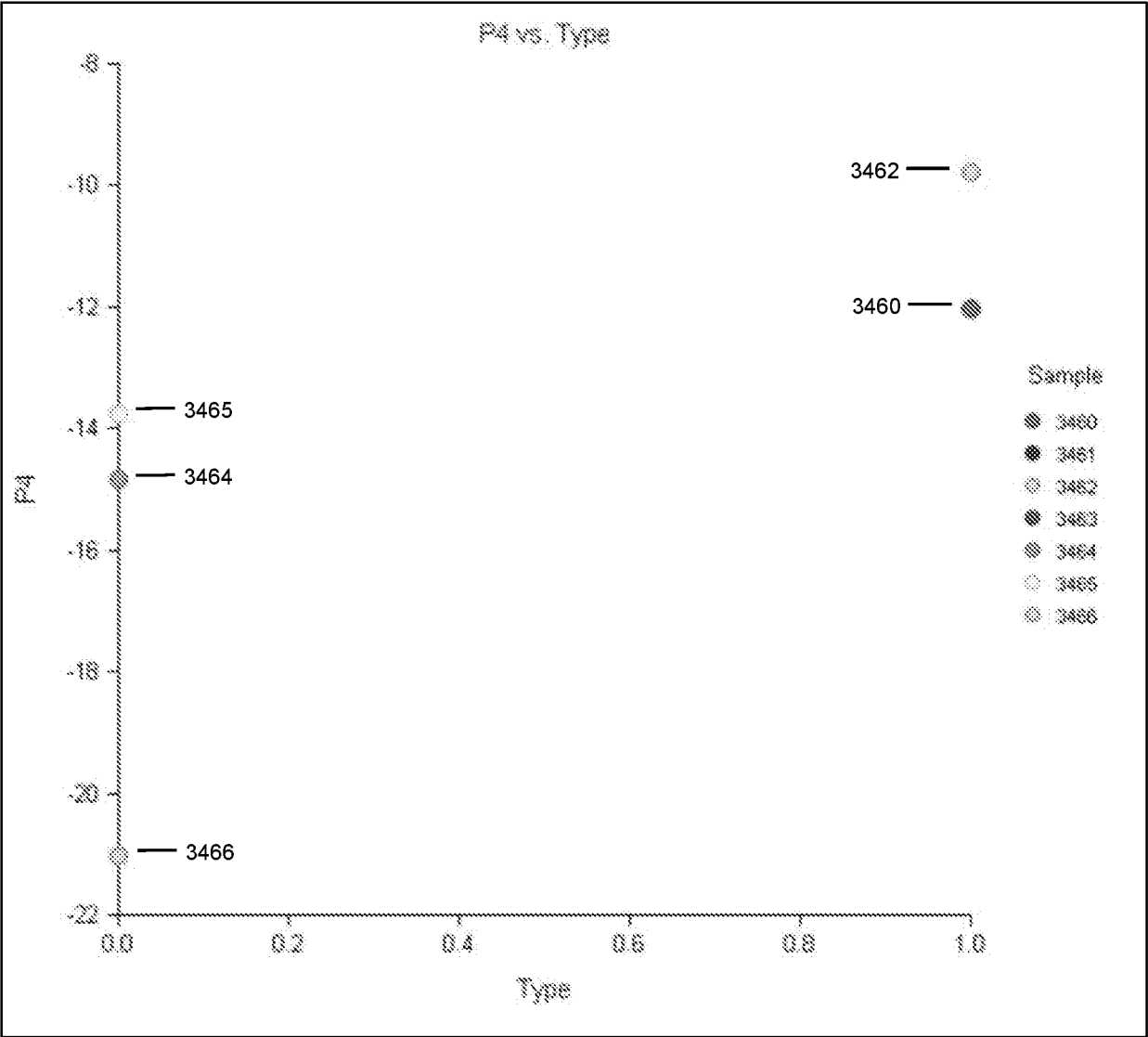


FIG. 2B

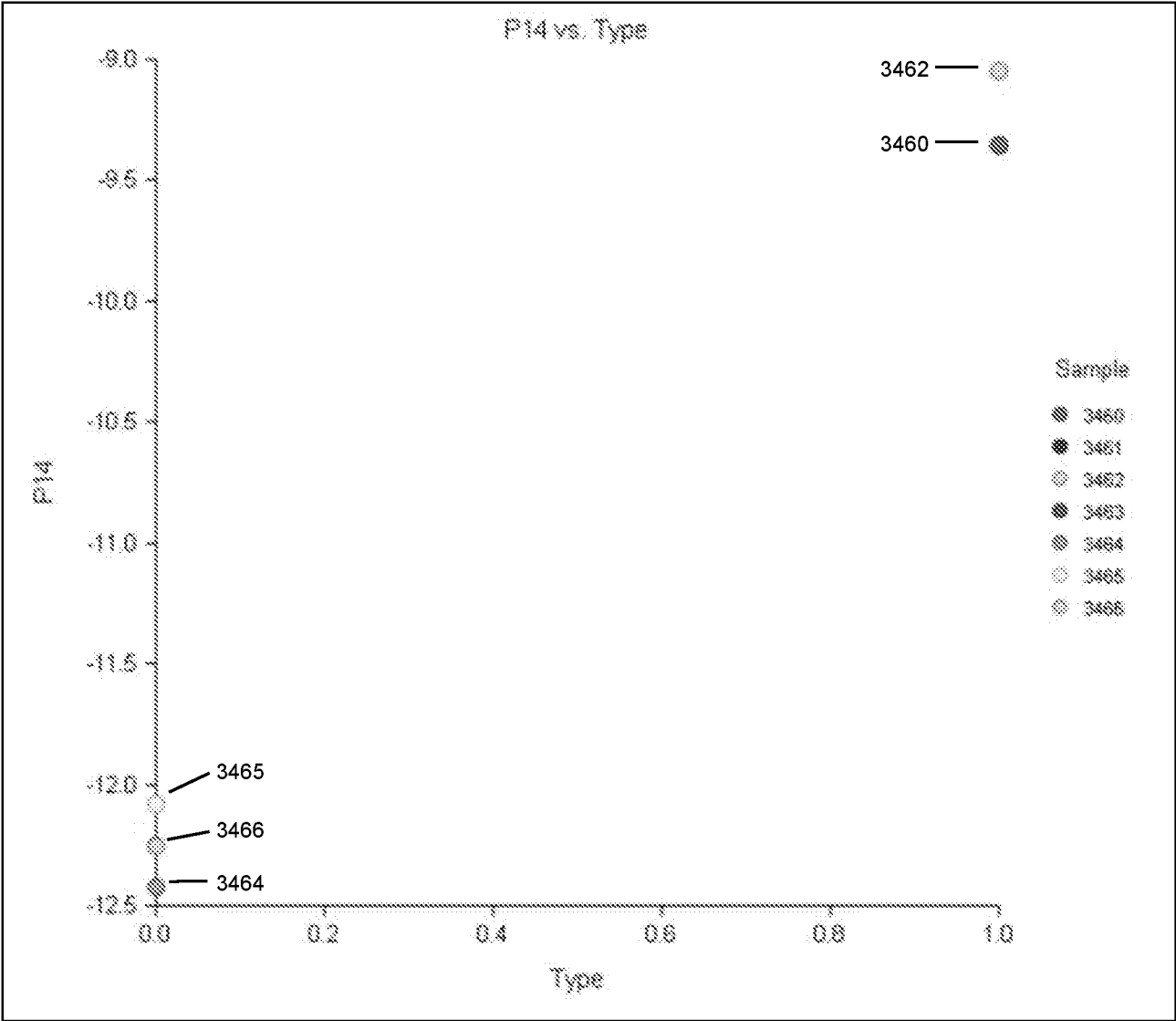


FIG. 2C

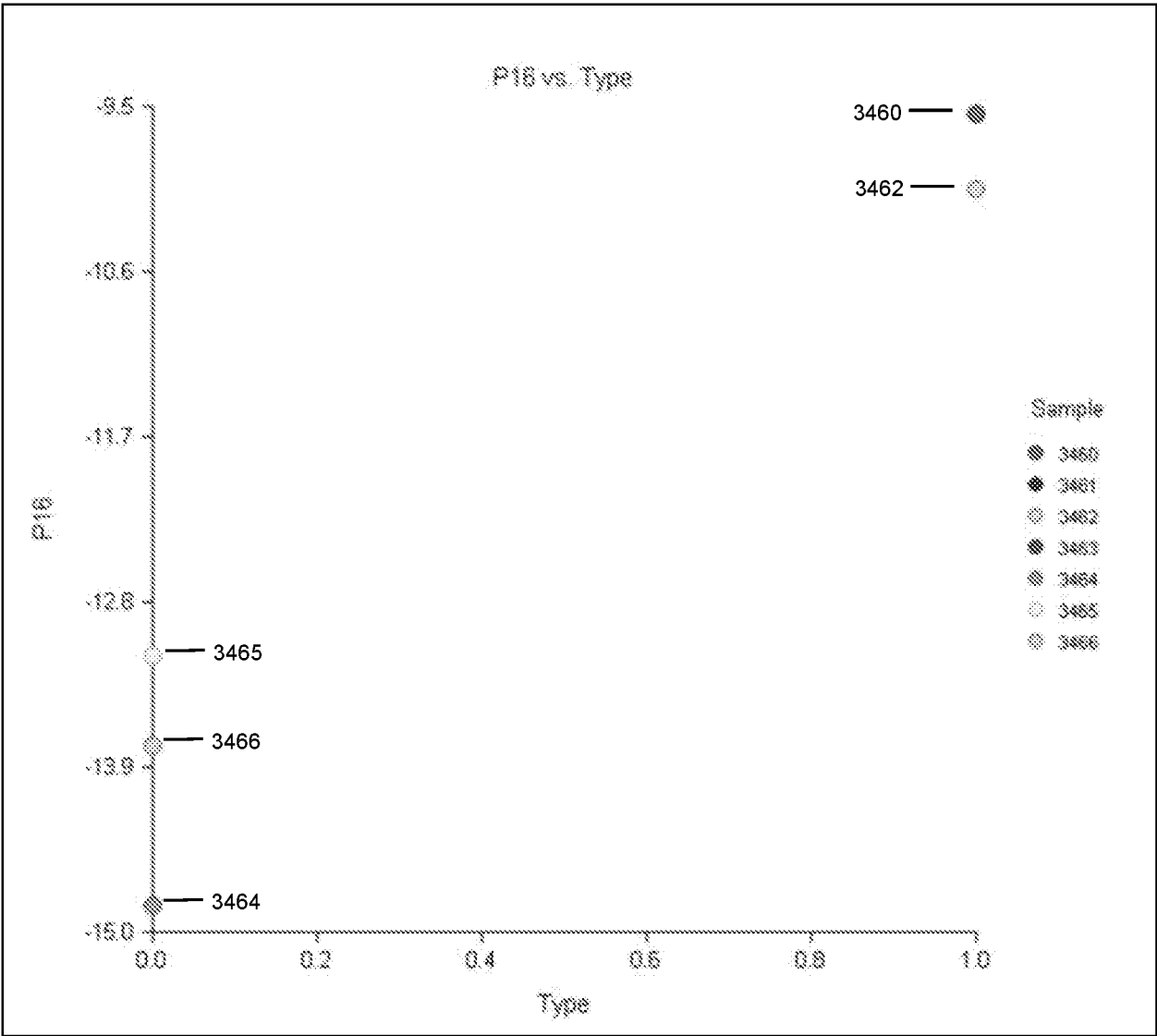


FIG. 2D

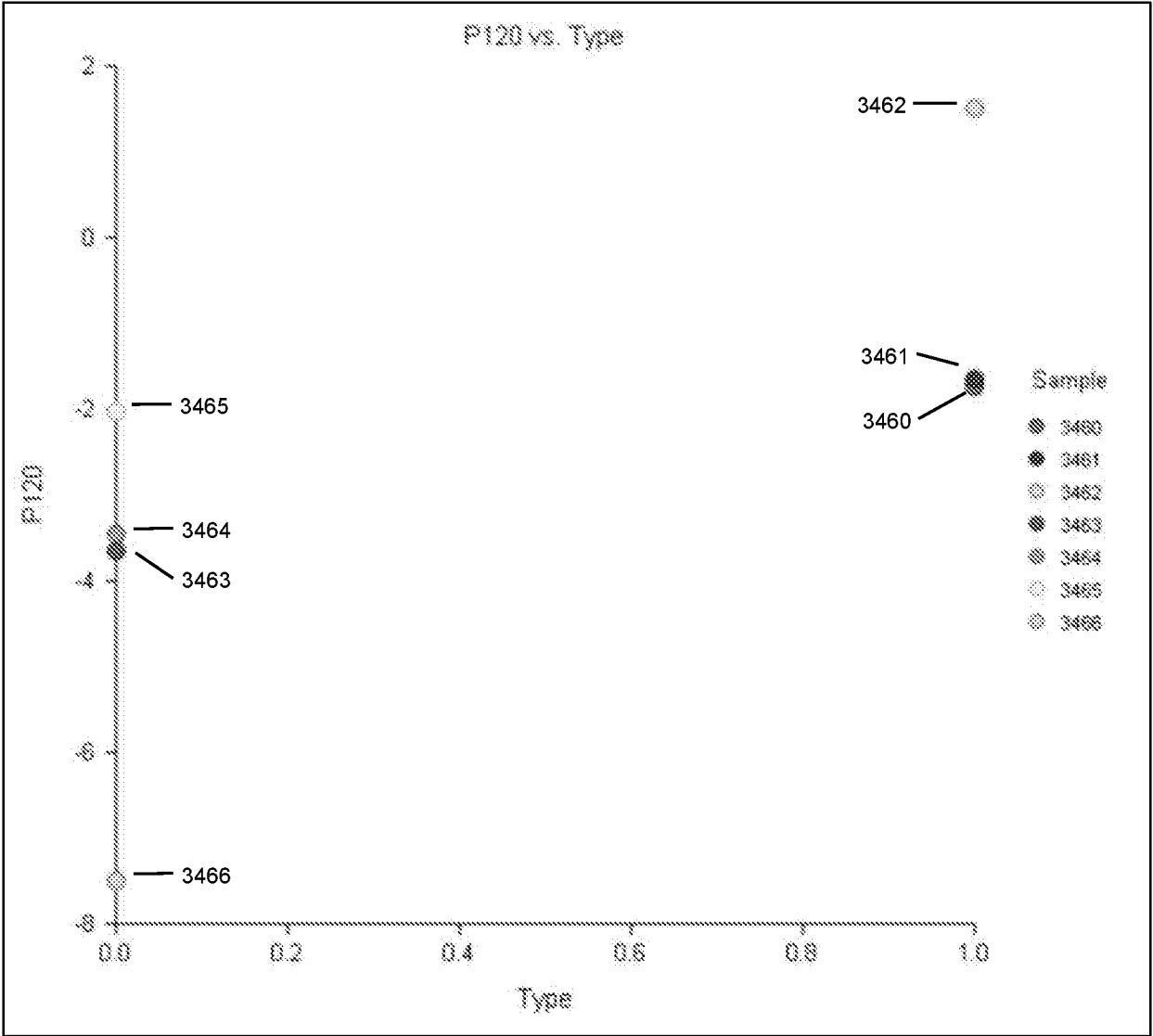


FIG. 2E

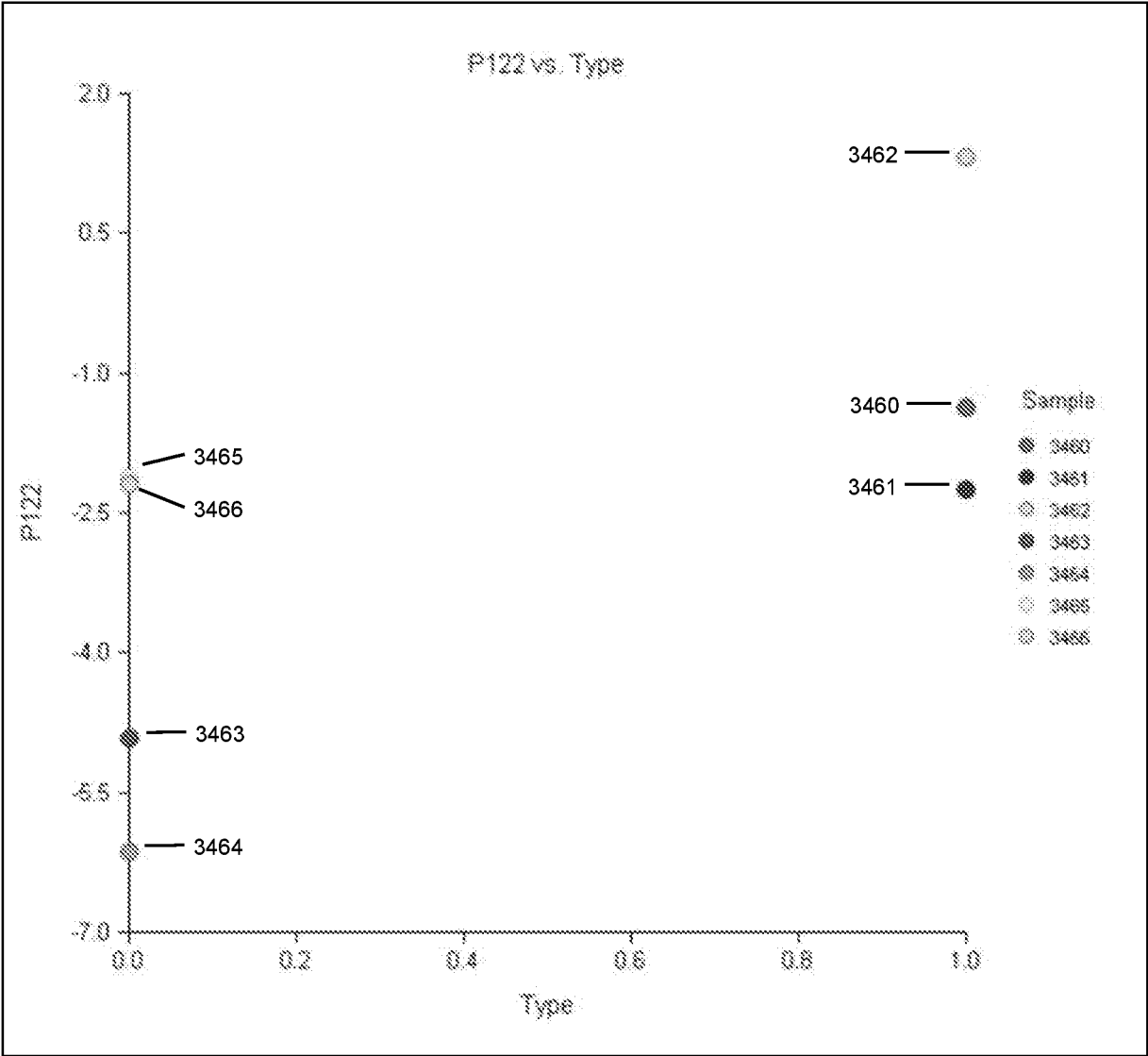


FIG. 2F

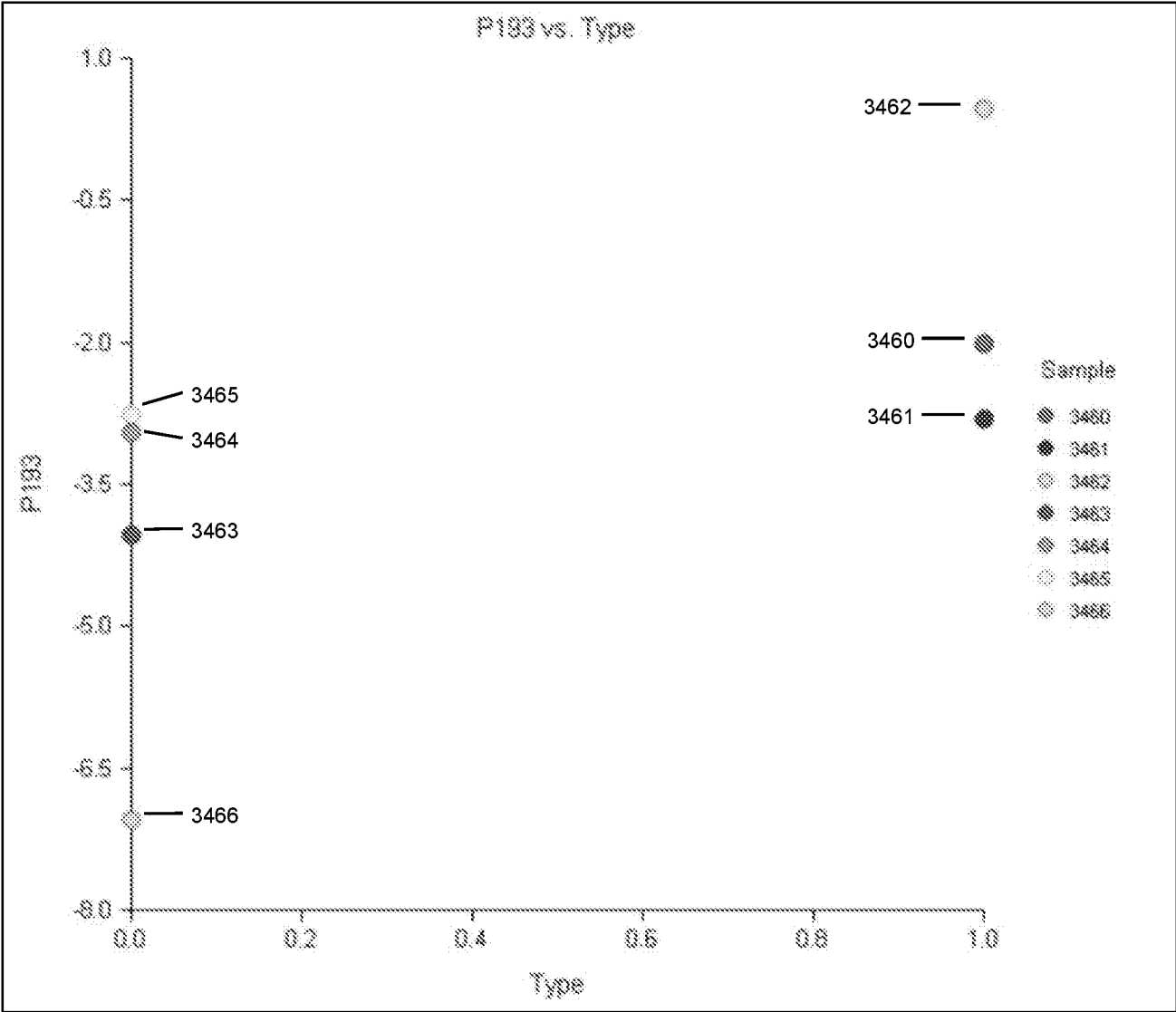


FIG. 2G

9/35

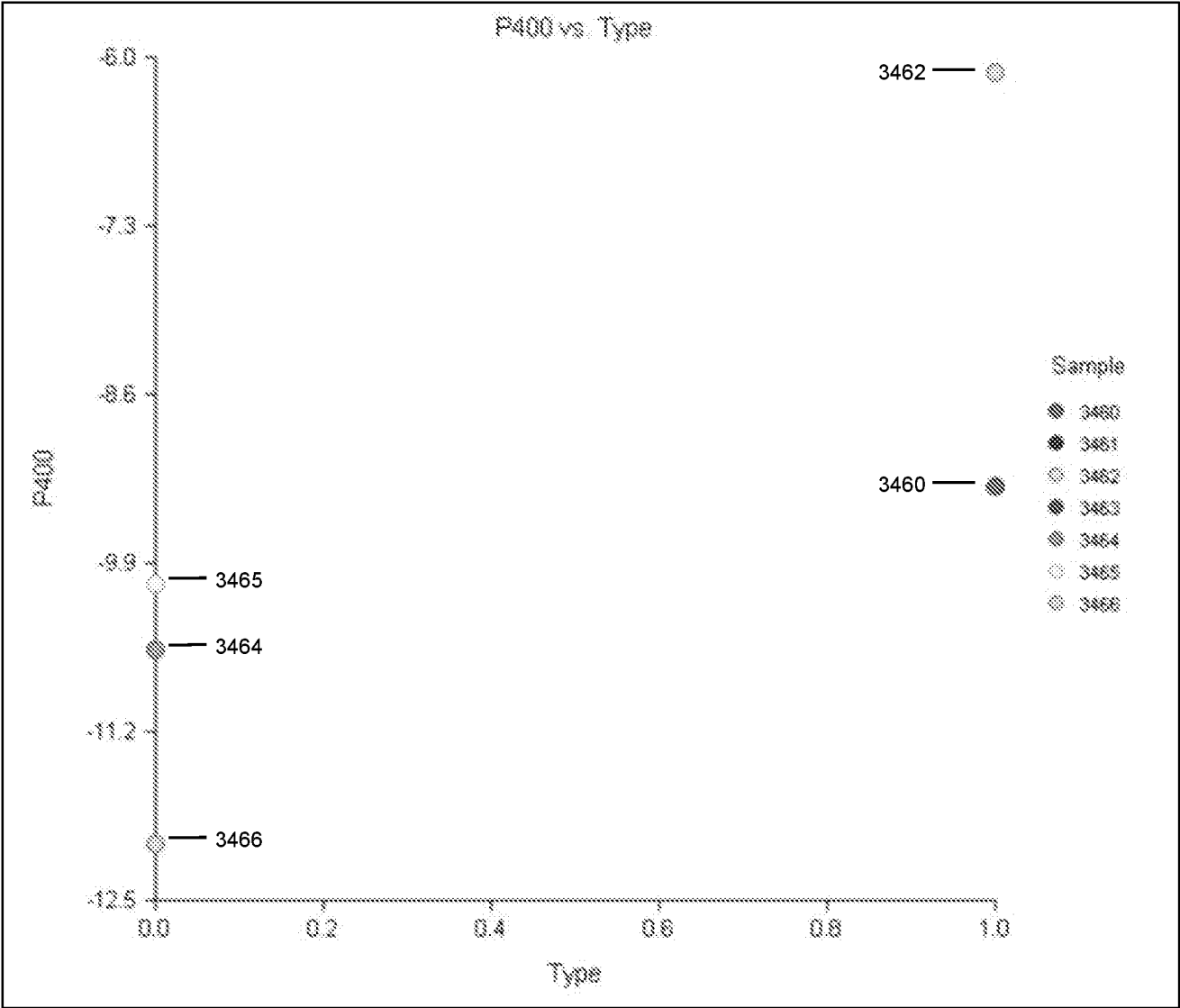


FIG. 2H

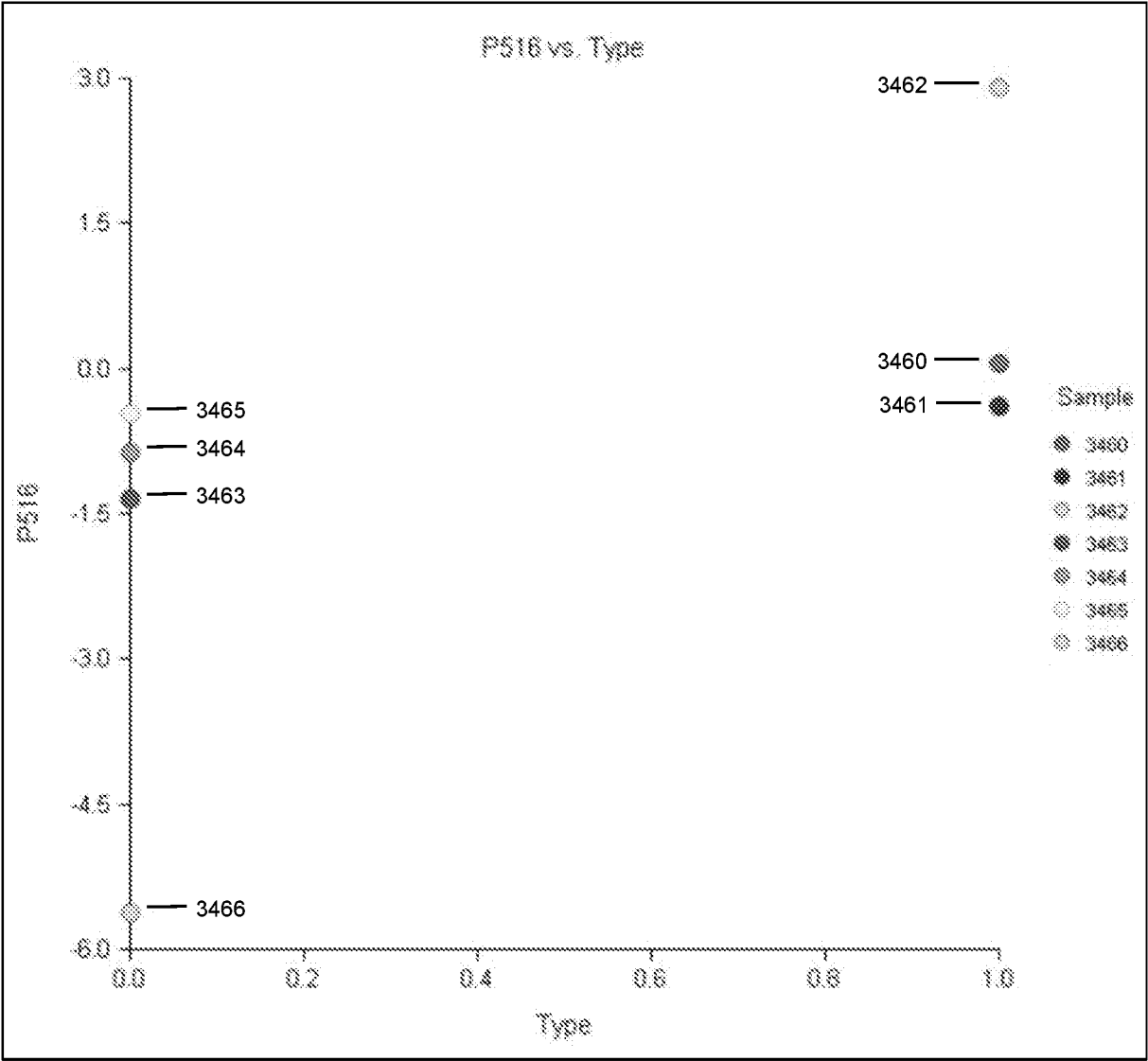


FIG. 2I

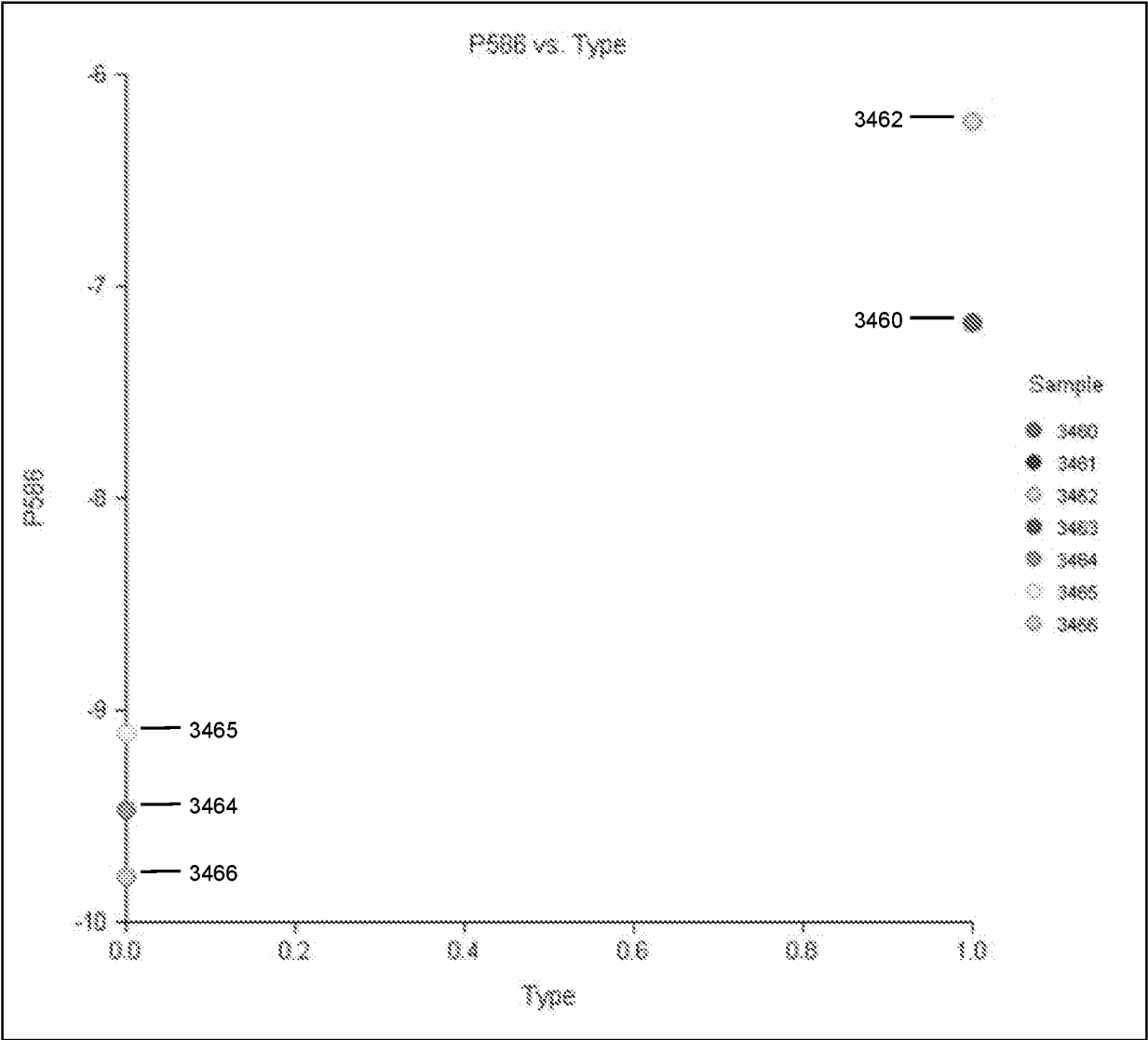


FIG. 2J

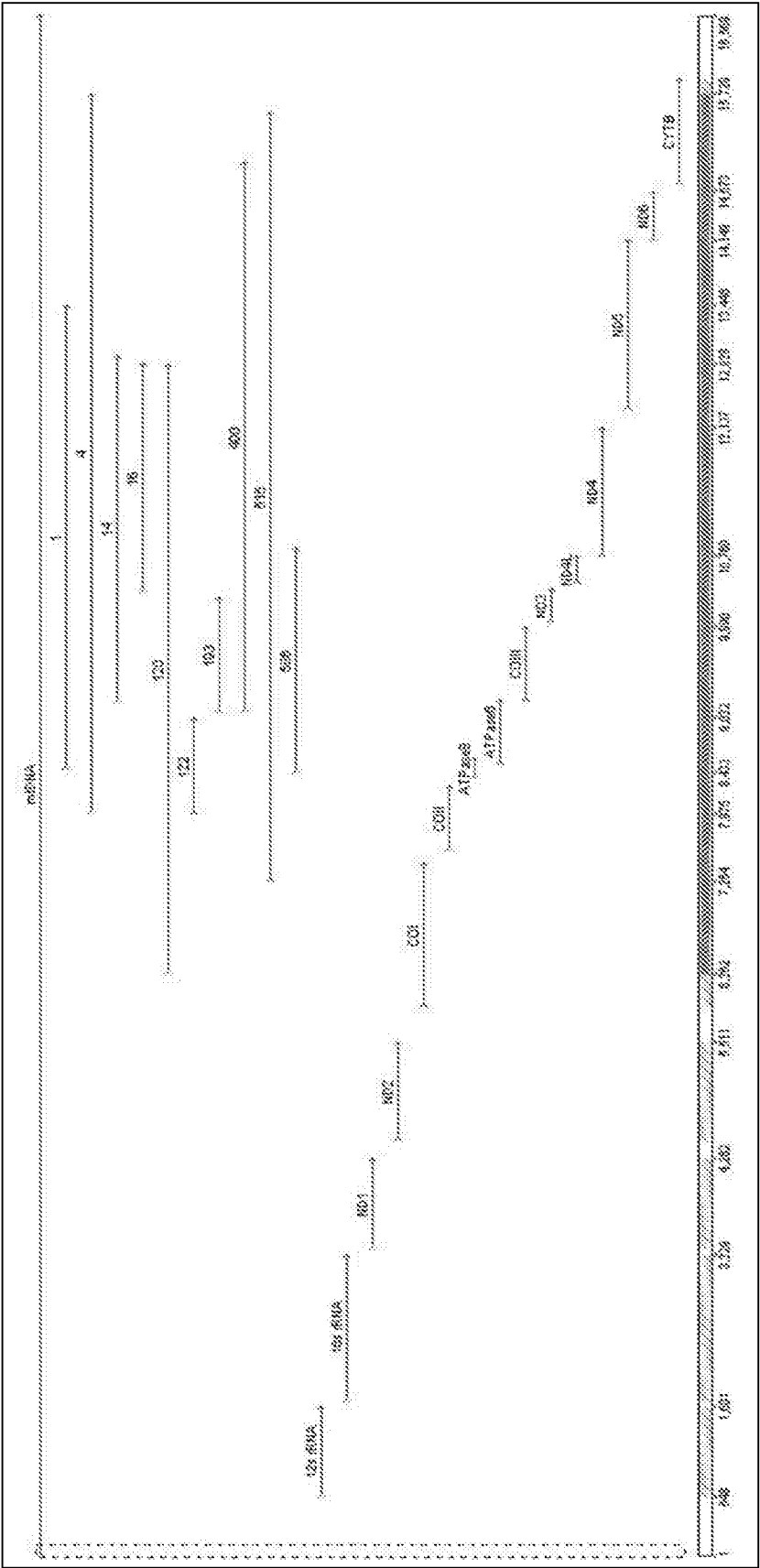


FIG. 3

13/35

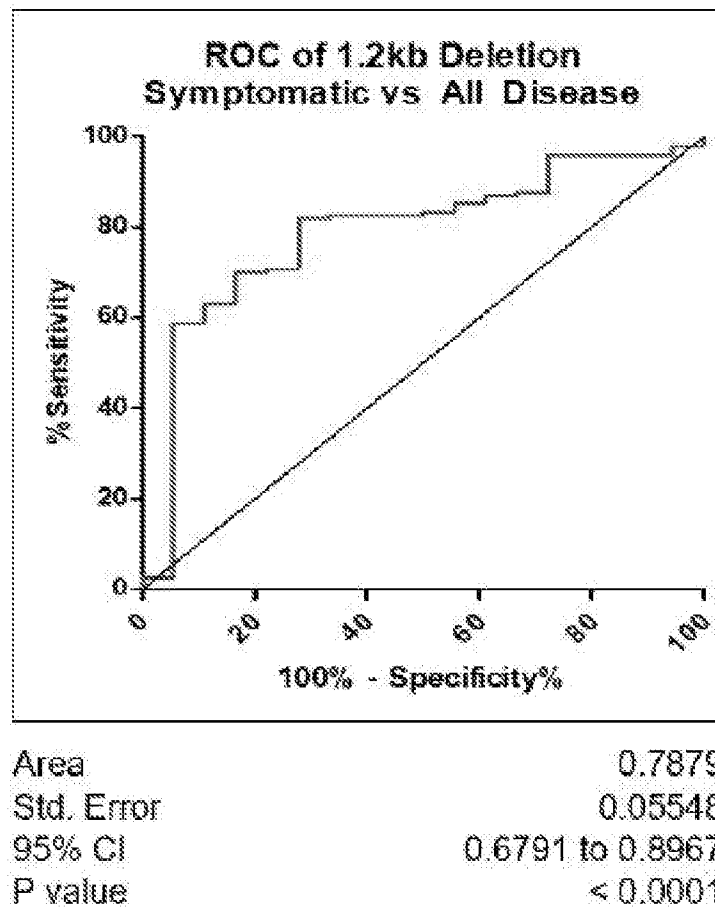


FIG. 4A

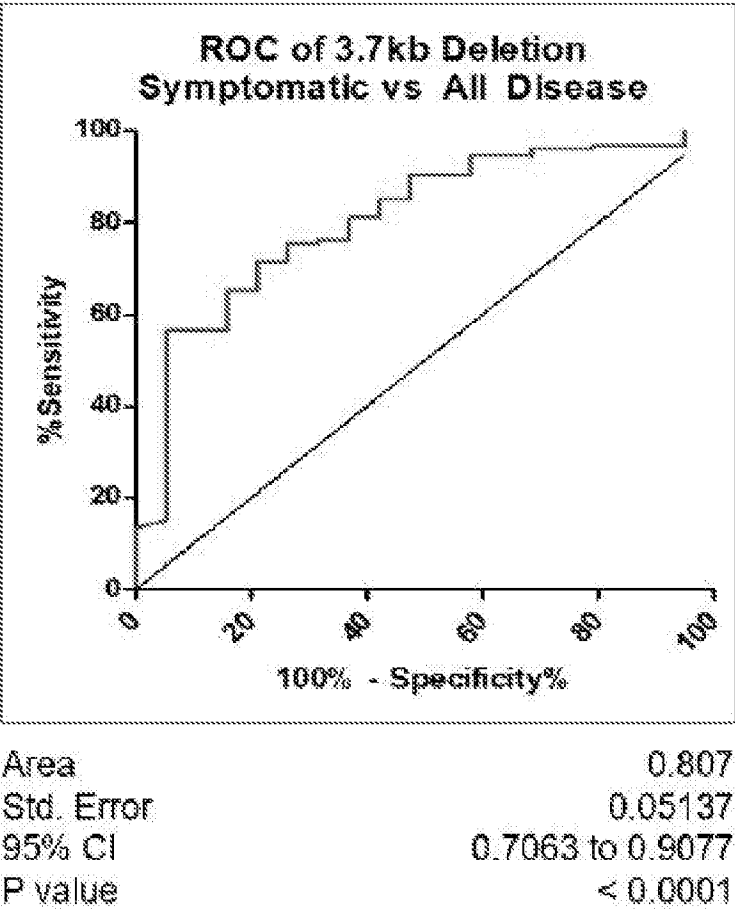
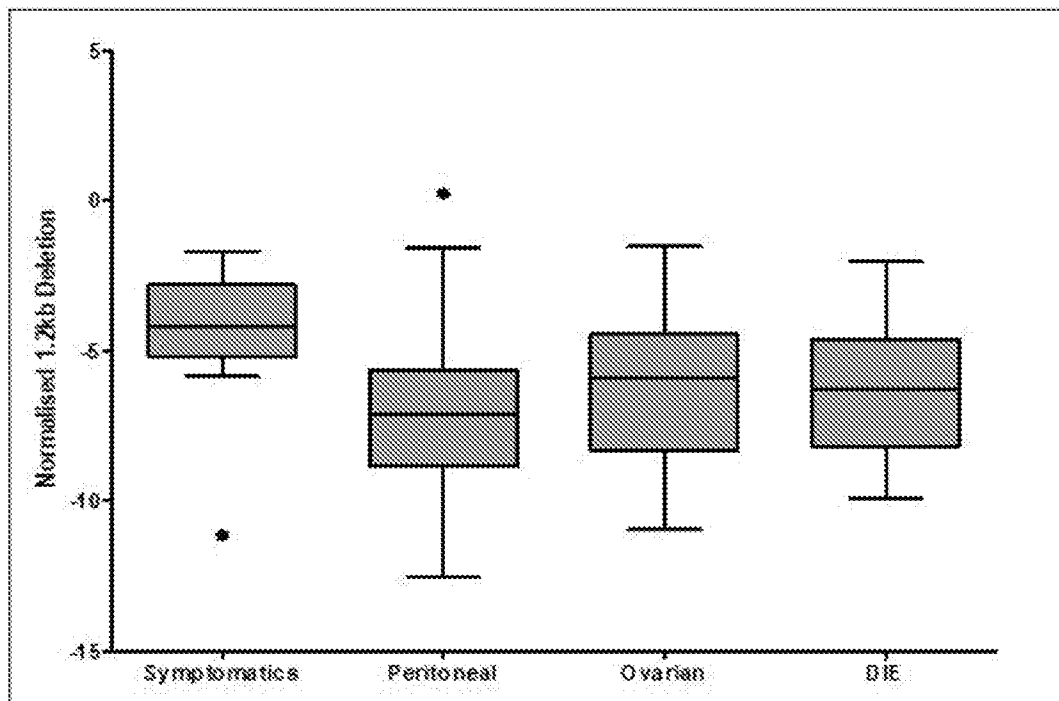


FIG. 4B

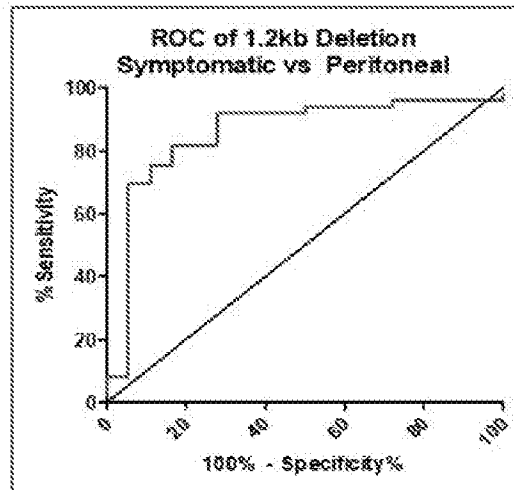
15/35



	Sympto	Peritoneal	Ovarian	DIE
N	18	49	45	49
Mean	-4.312	-7.187	-6.291	-6.193
Std. Dev	2.075	2.581	2.344	2.143
Std. Error	0.4891	0.3687	0.3494	0.3062
Lower 95% CI	-5.344	-7.928	-6.996	-6.809
Upper 95% CI	-3.28	-6.446	-5.587	-5.578

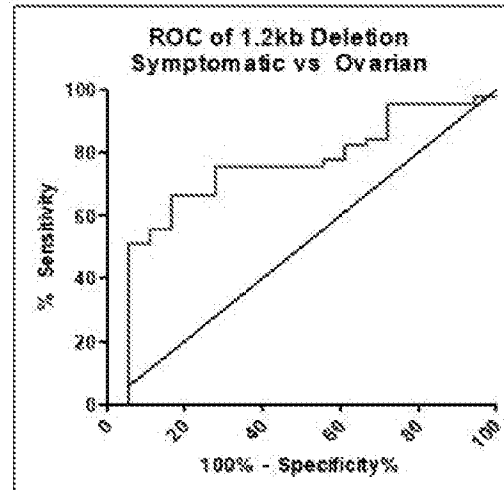
Fig. 5A

16/35



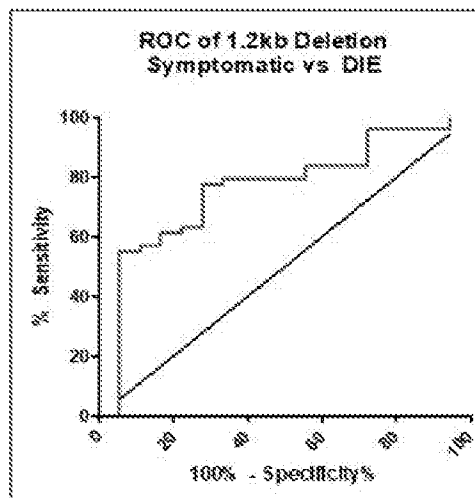
Area 0.8549
Std. Error 0.05731
95% CI 0.7425 to 0.9672
P value < 0.0001

Fig. 5B



Area 0.7457
Std. Error 0.0683
95% CI 0.6118 to 0.8796
P value 0.002477

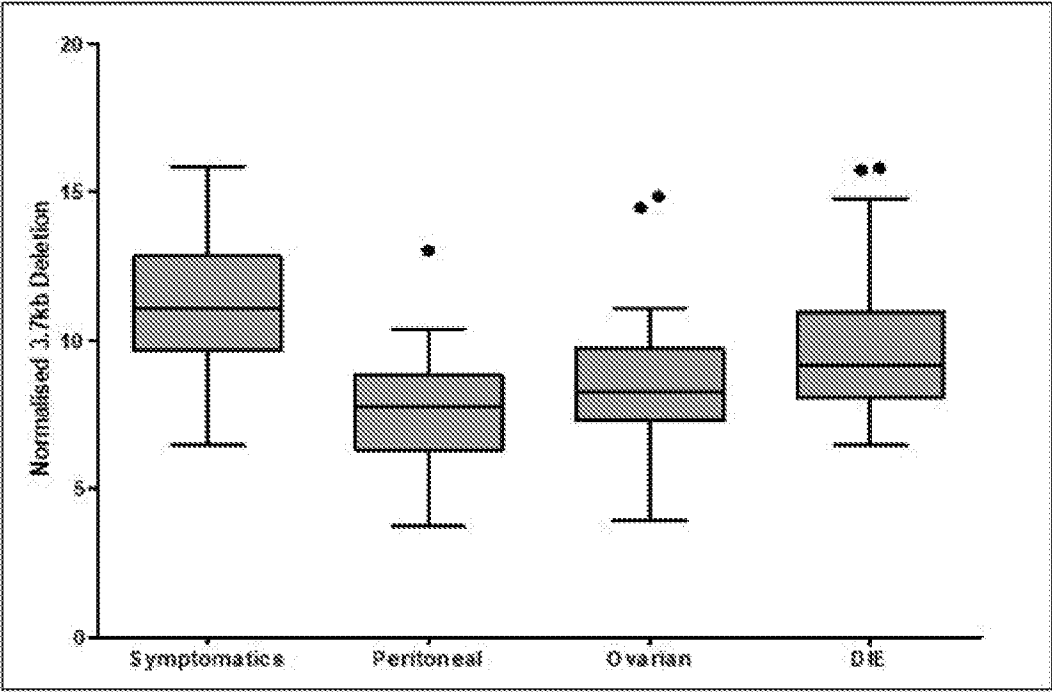
Fig. 5C



Area 0.7596
Std. Error 0.06655
95% CI 0.6292 to 0.8901
P value 0.001207

Fig. 5D

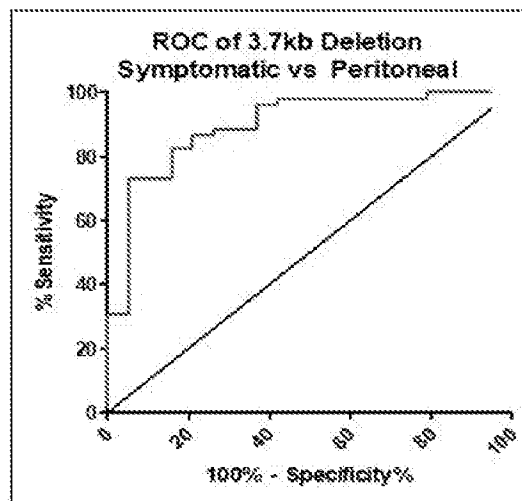
17/35



	Sympto	Peritoneal	Ovarian	DIE
N	19	52	47	50
Mean	11.12	7.569	8.549	9.617
Std. Dev	2.239	1.843	2.089	2.125
Std. Error	0.5136	0.2555	0.3047	0.3005
Lower 95% CI	10.04	7.056	7.936	9.013
Upper 95% CI	12.19	8.082	9.162	10.22

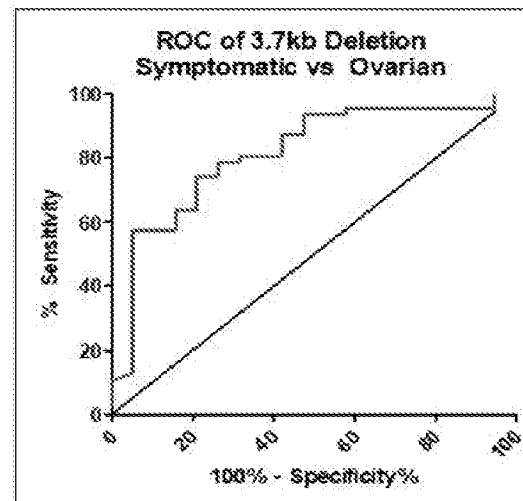
Fig. 6A

18/35



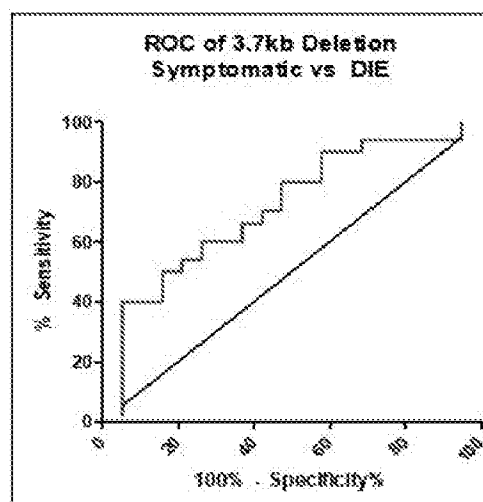
Area 0.8978
 Std. Error 0.04318
 95% CI 0.8131 to 0.9824
 P value < 0.0001

Fig. 6B



Area 0.8158
 Std. Error 0.05891
 95% CI 0.7003 to 0.9313
 P value < 0.0001

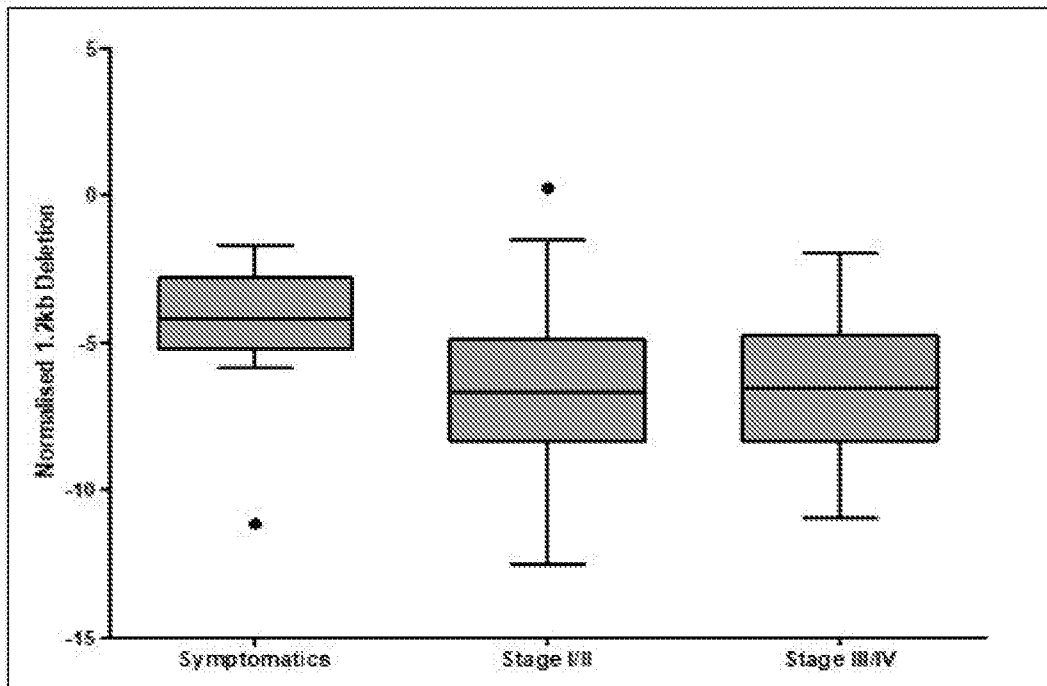
Fig. 6C



Area 0.7111
 Std. Error 0.06962
 95% CI 0.5746 to 0.8475
 P value 0.007097

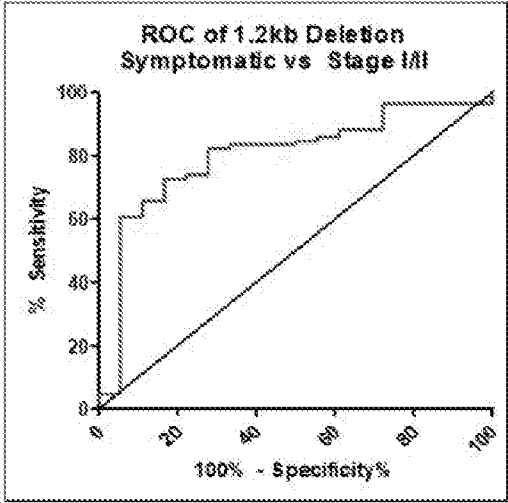
Fig. 6D

19/35



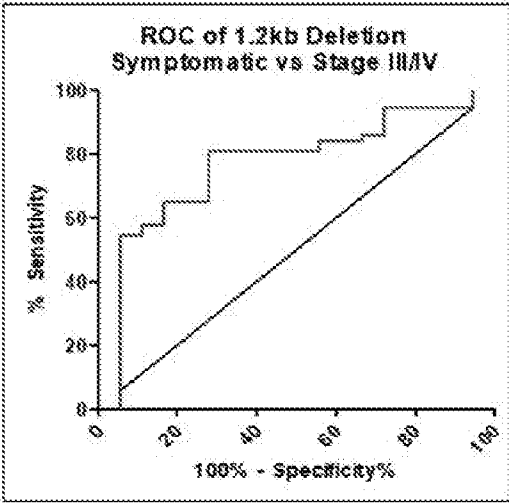
	Sympto	Stage I/II	Stage III/IV
N	18	84	57
Mean	-4.312	-6.692	-6.348
Std. Dev	2.075	2.505	2.25
Std. Error	0.4691	0.2733	0.2981
Lower 95% CI	-5.344	-7.236	-6.945
Upper 95% CI	-3.28	-6.148	-5.751

Fig. 7A



Area 0.7989
Std. Error 0.05719
95% CI 0.6868 to 0.9111
P value < 0.0001

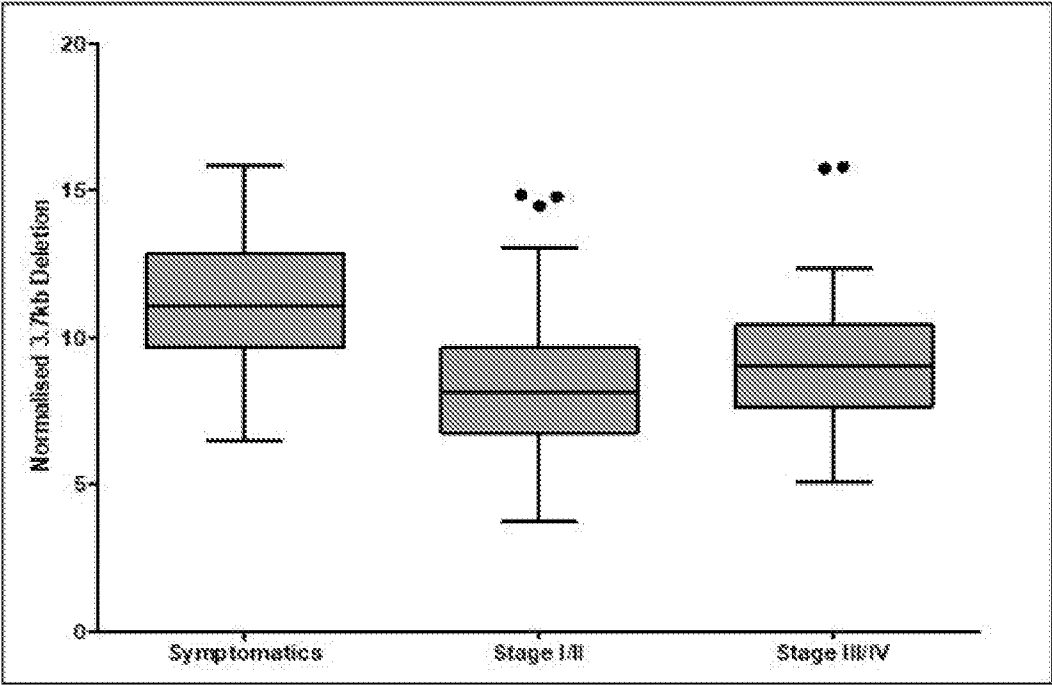
Fig. 7B



Area 0.7661
Std. Error 0.06443
95% CI 0.6398 to 0.8924
P value 0.0007129

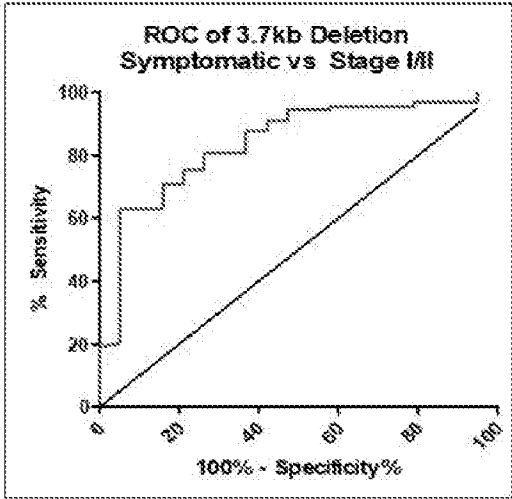
Fig. 7C

21/35



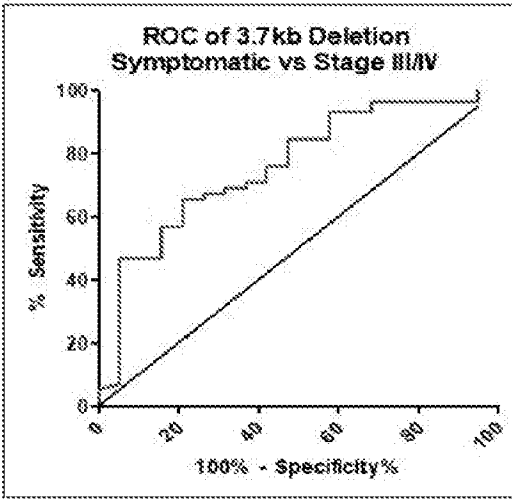
	Sympto	Stage I/II	Stage III/IV
N	19	89	58
Mean	11.12	8.243	9.112
Std. Dev	2.239	2.156	2.14
Std. Error	0.5136	0.2285	0.2811
Lower 95% CI	10.04	7.789	8.549
Upper 95% CI	12.19	8.697	9.675

Fig. 8A



Area 0.8383
Std. Error 0.04949
95% CI 0.7412 to 0.9353
P value < 0.0001

Fig. 8B



Area 0.7591
Std. Error 0.06356
95% CI 0.6345 to 0.8837
P value 0.0007483

Fig. 8C

23/35

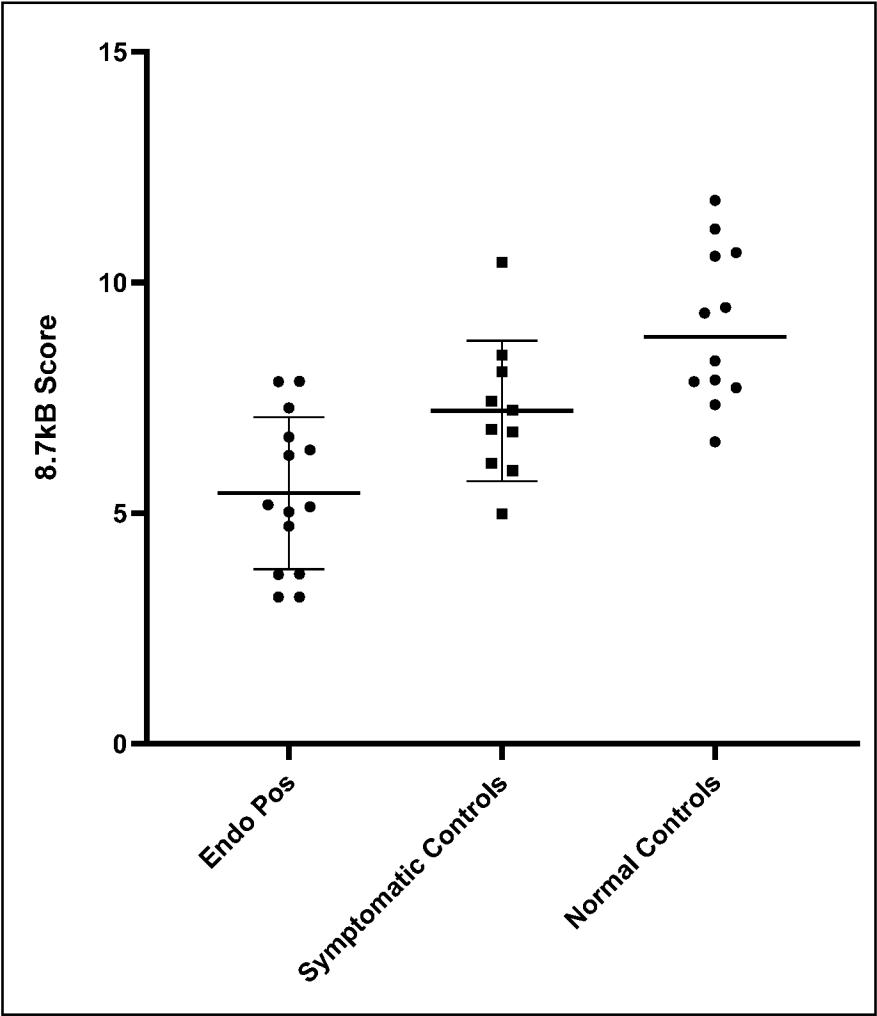


Fig. 9

24/35

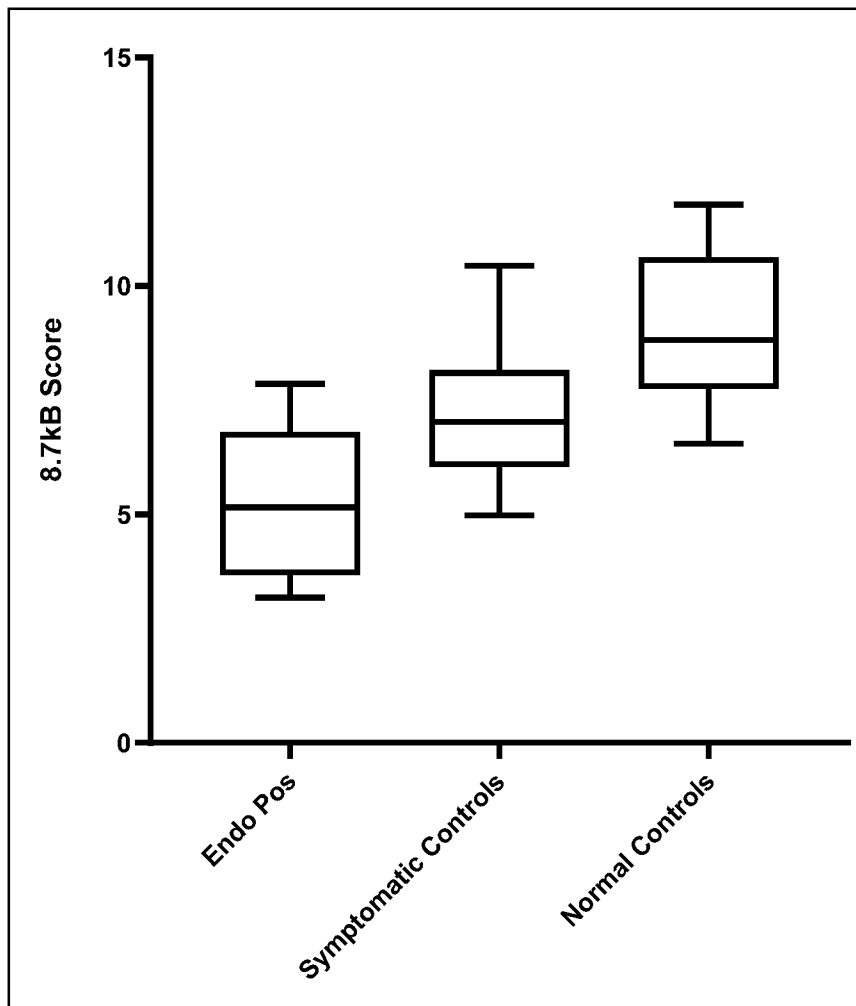
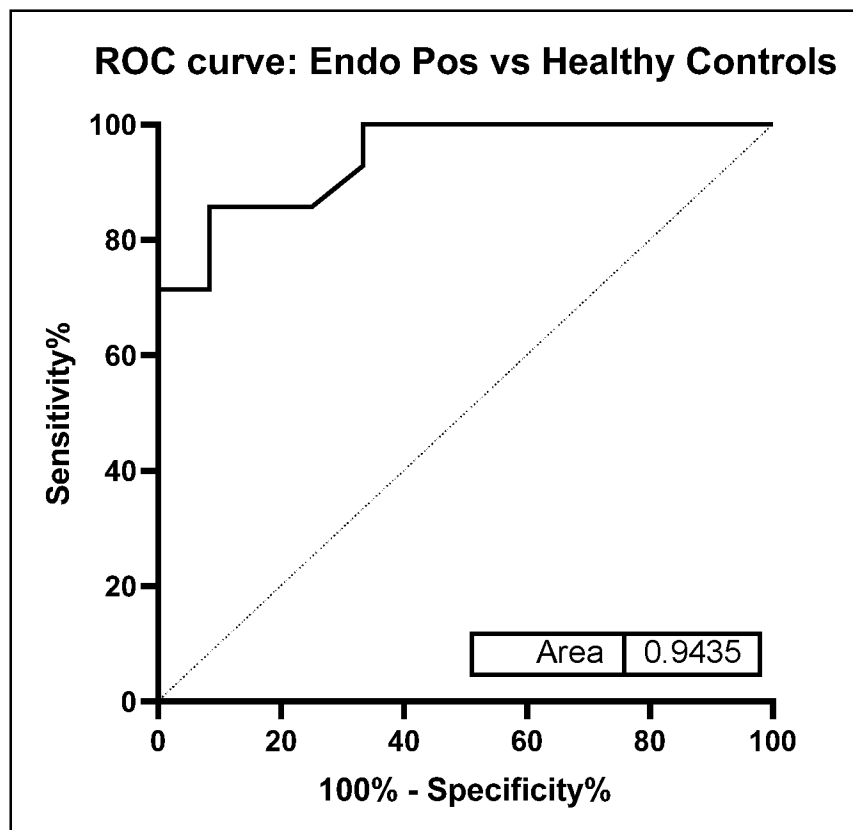


Fig. 10

25/35



Area under the ROC curve	
Area	0.9435
Std. Error	0.04138
95% confidence interval	0.8623 to 1.000
P value	0.0001
Data	
Controls (BioRec Normal Controls)	12
Patients (Endo Pos)	14
Missing Controls	0
Missing Patients	10

Fig. 11

26/35

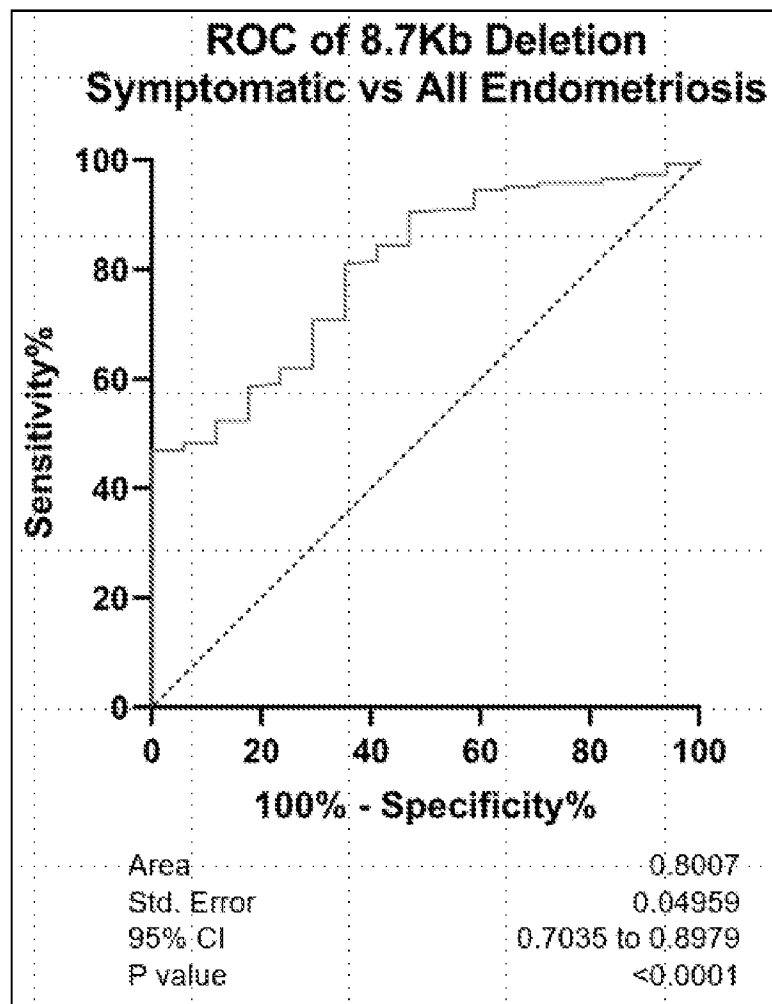


Fig. 12

27/35

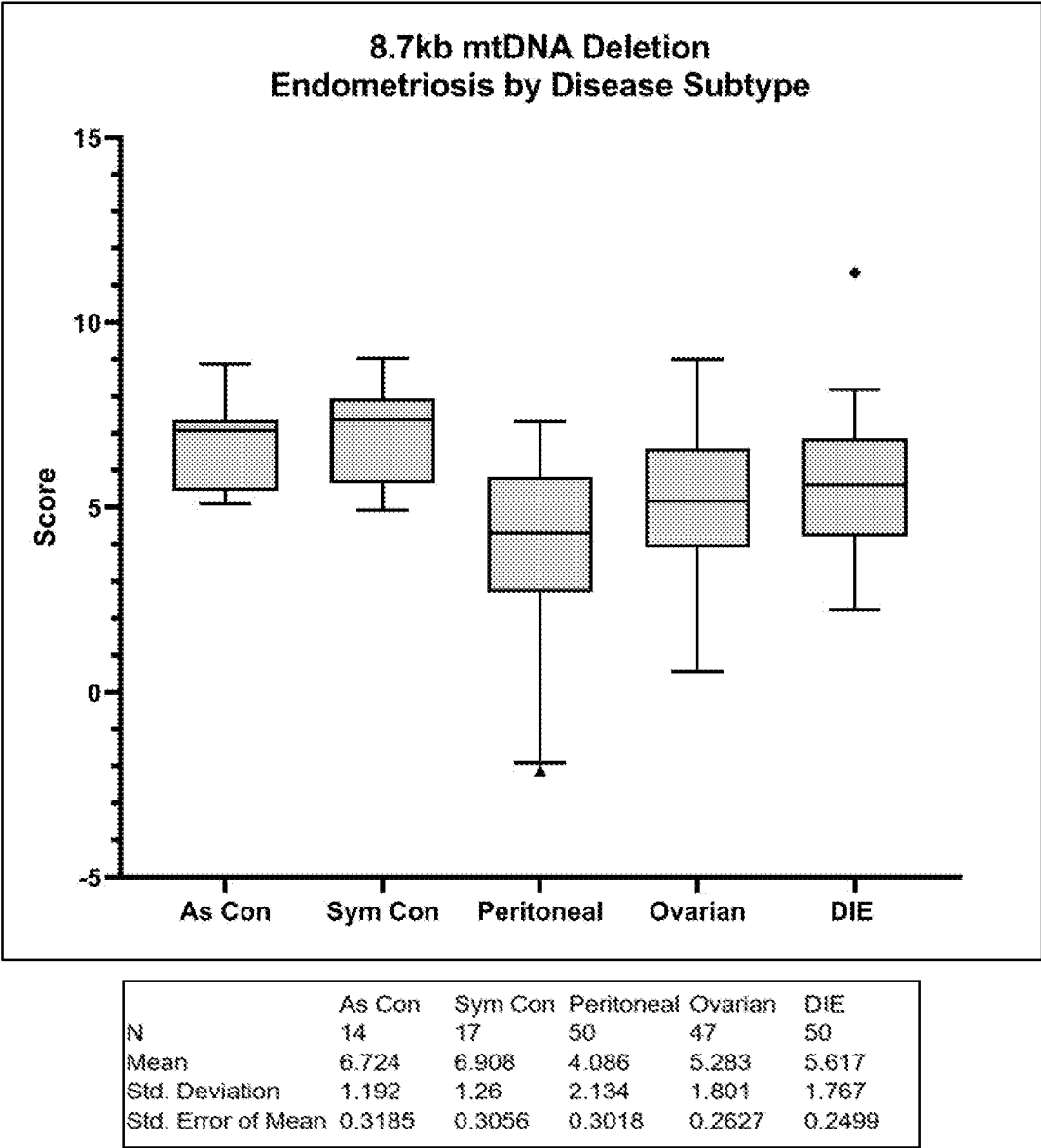
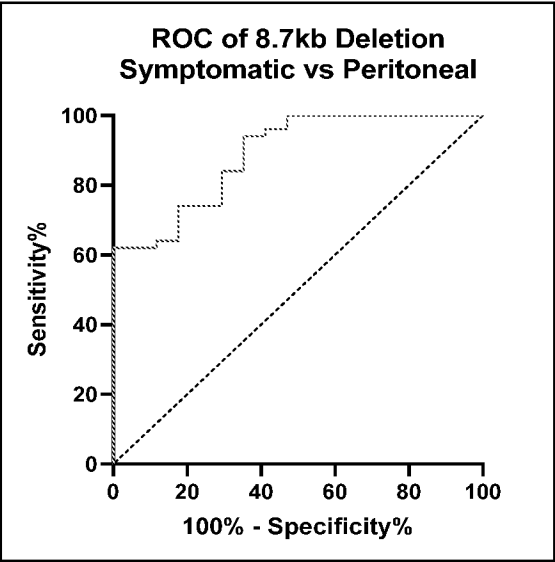
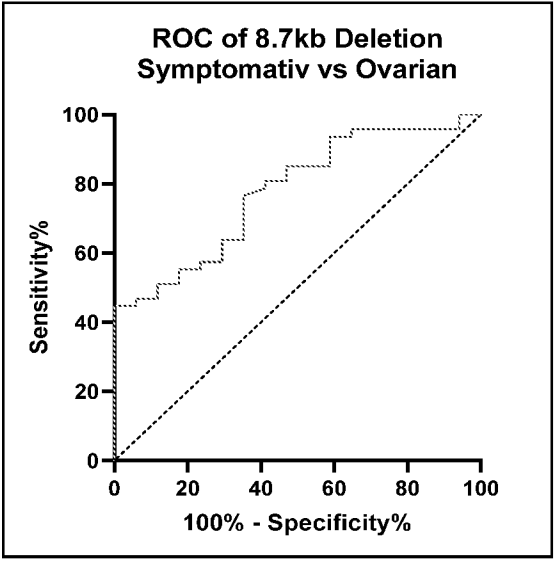


Fig. 13A



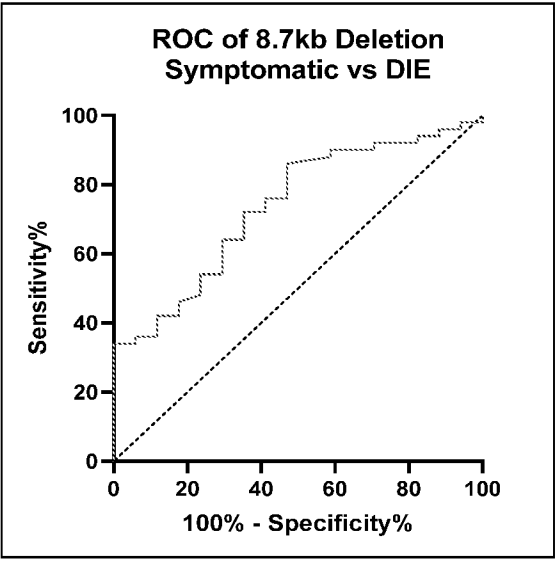
Area	0.8882
Std. Error	0.04284
95% CI	0.8043 to 0.9722
P value	<0.0001

Fig.13B



Area	0.7766
Std. Error	0.0609
95% CI	0.6572 to 0.8960
P value	0.0008

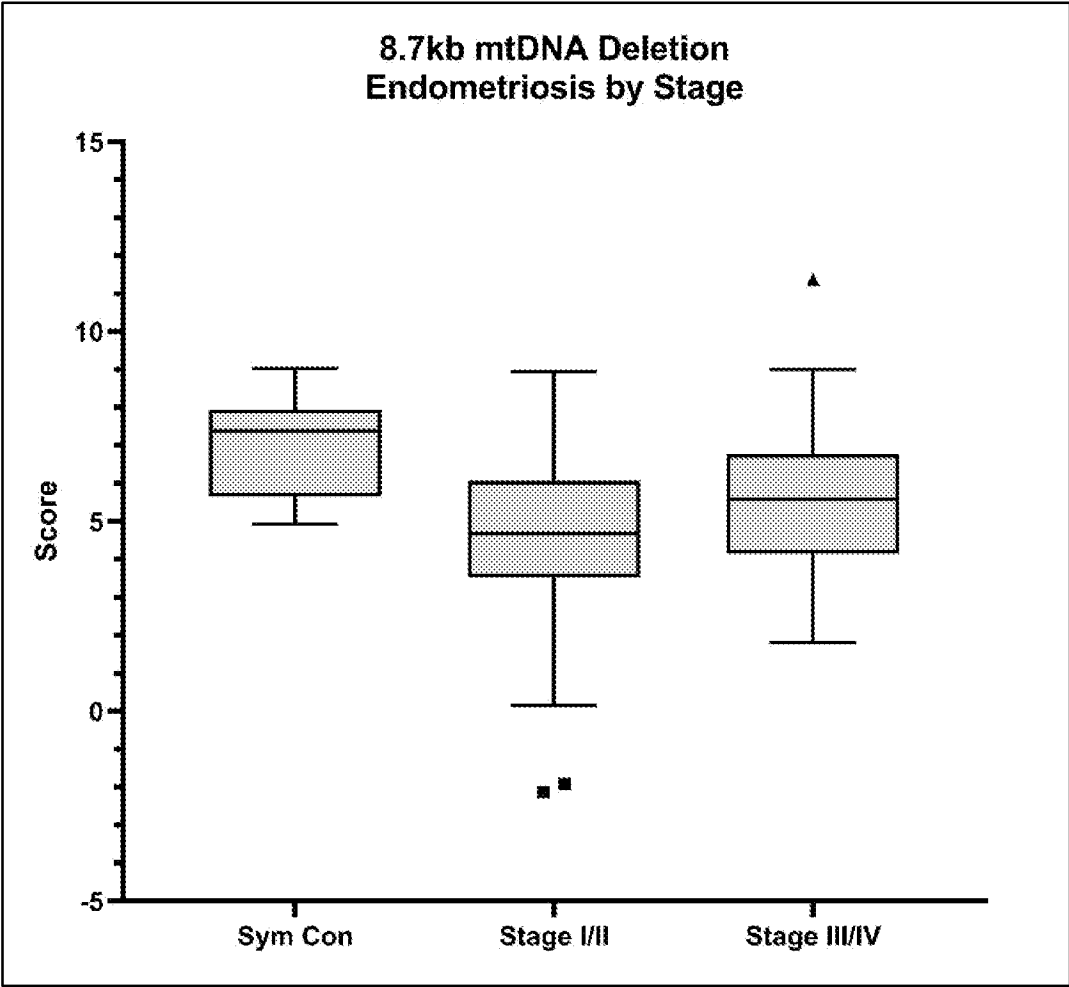
Fig. 13C



Area	0.7359
Std. Error	0.06643
95% CI	0.6057 to 0.8661
P value	0.0039

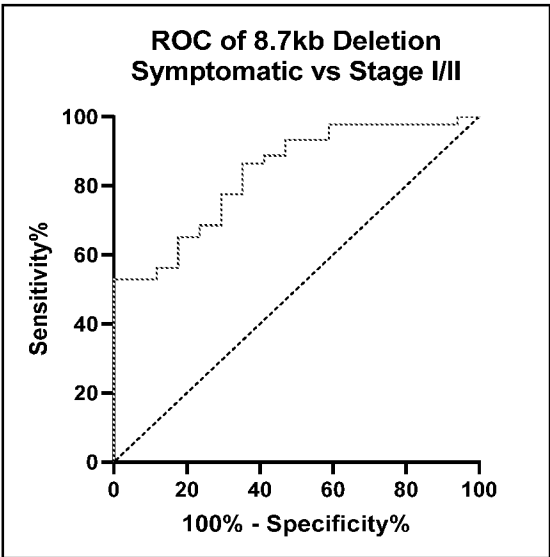
Fig.13D

29/35



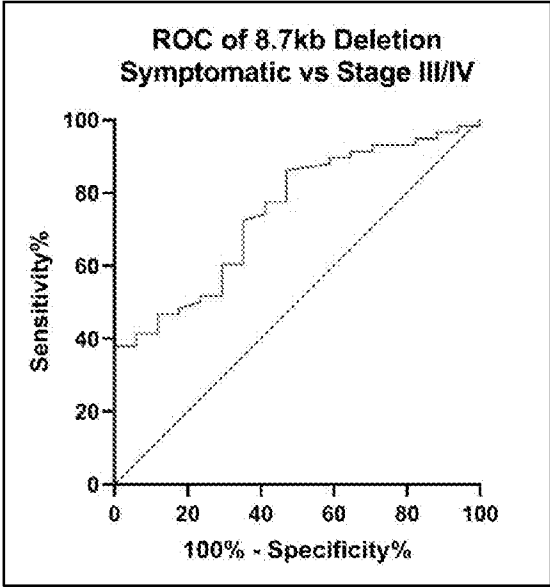
	Sym Con	Stage I/II	Stage III/IV
N	17	89	58
Mean	6.908	4.614	5.565
Std. Deviation	1.26	2.063	1.794
Std. Error of Mean	0.3056	0.2187	0.2356

Fig. 14A



Area	0.8361
Std. Error	0.04768
95% CI	0.7426 to 0.9295
P value	<0.0001

Fig 14B



Area	0.7465
Std. Error	0.06287
95% CI	0.6232 to 0.8697
P value	0.0021

Fig. 14C

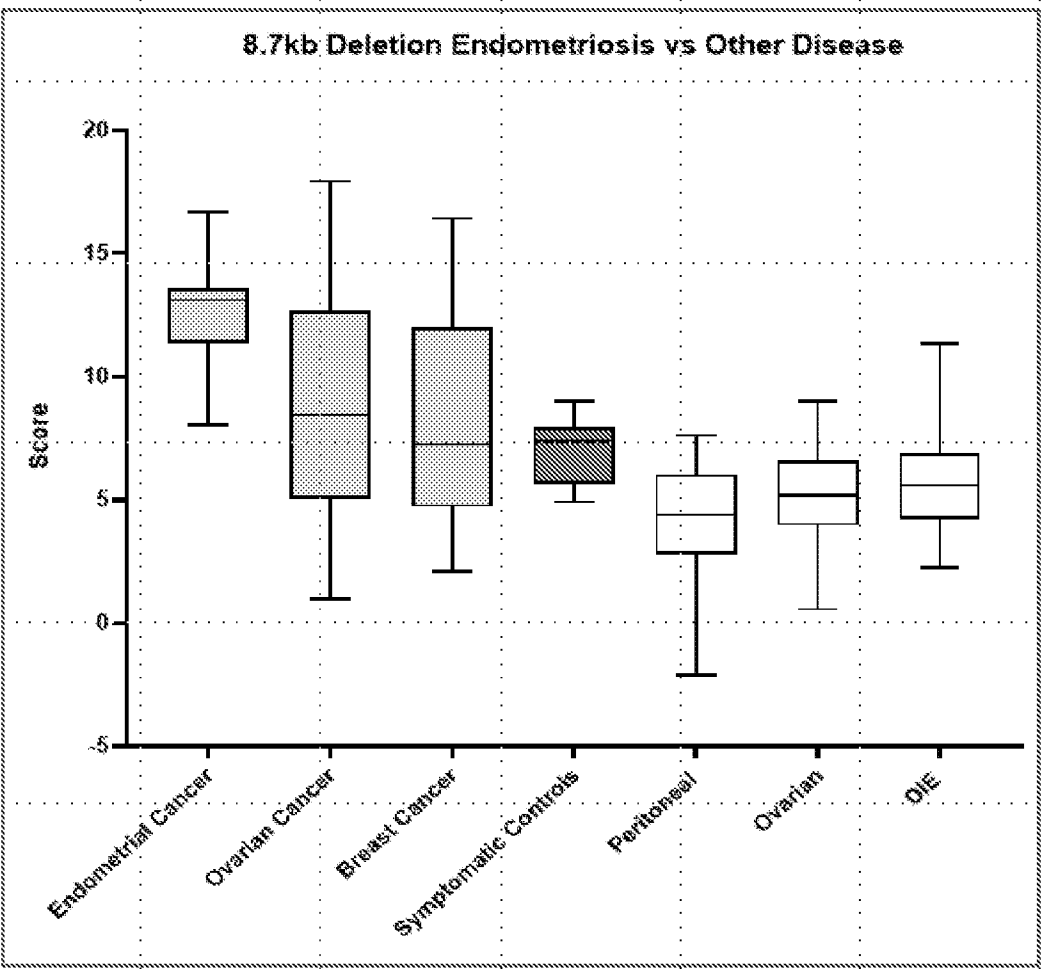


Fig. 15

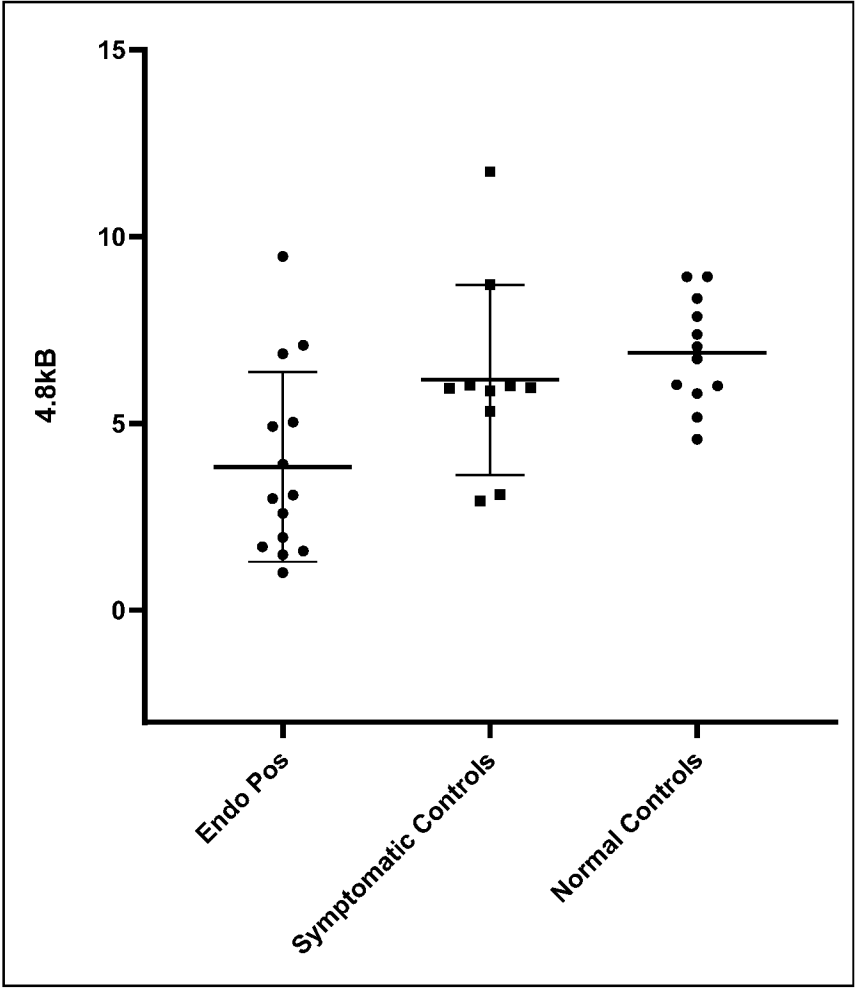


Fig. 16

33/35

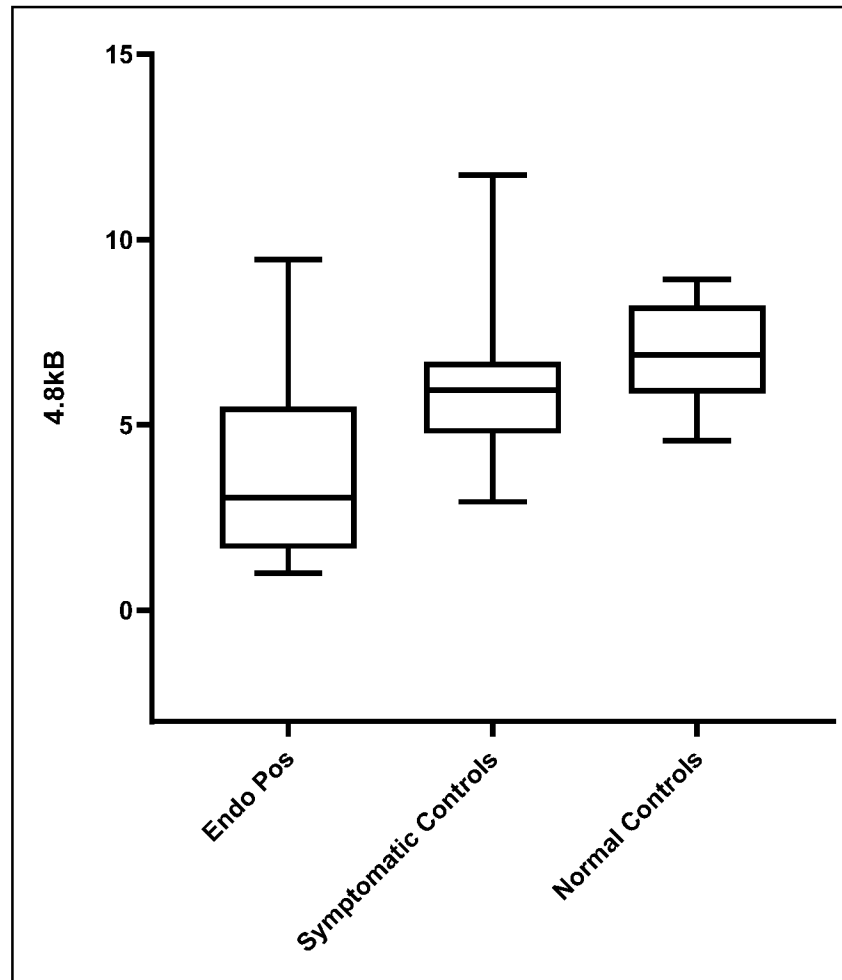


Fig. 17

34/35

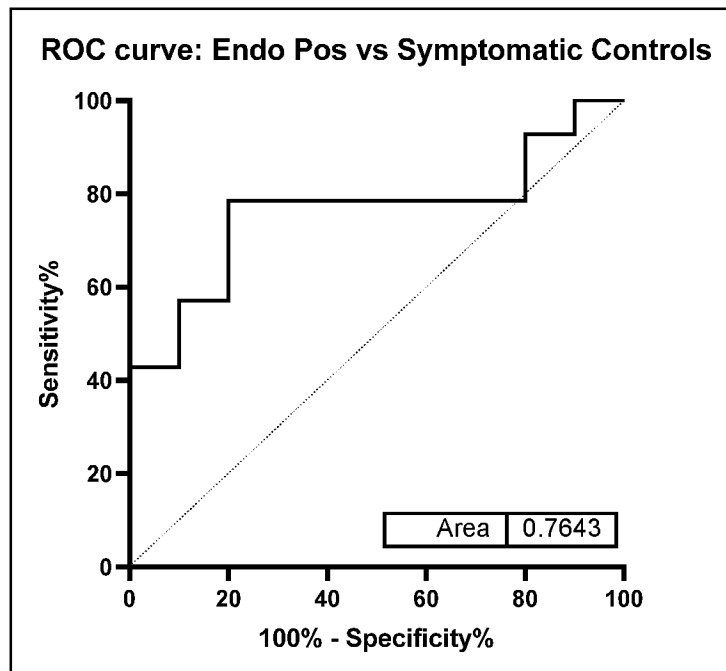


Fig. 18

Area under the ROC curve	
Area	0.7643
Std. Error	0.1007
95% confidence interval	0.5669 to 0.9616
P value	0.0303
Data	
Controls (Symptomatic Controls)	10
Patients (Endo Pos)	14
Missing Controls	0
Missing Patients	10

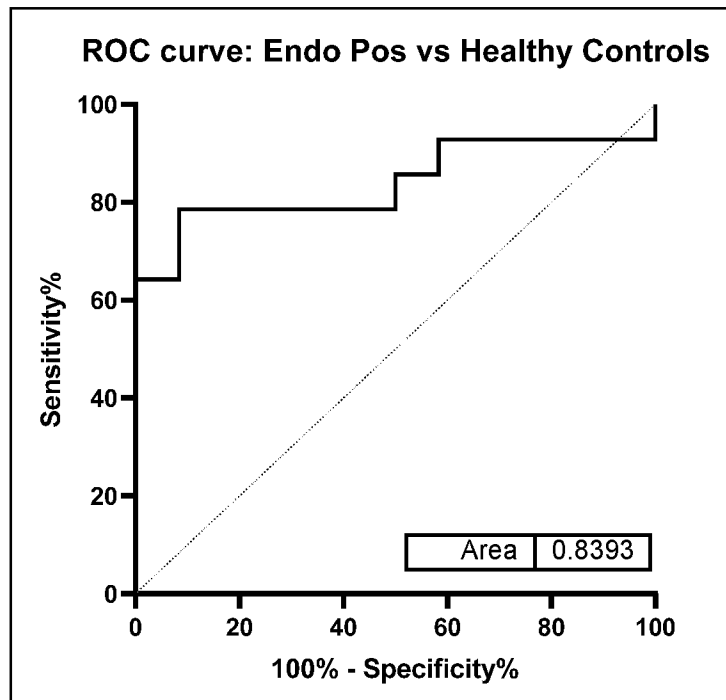


Fig. 19

Area under the ROC curve	
Area	0.8393
Std. Error	0.08494
95% confidence interval	0.6728 to 1.000
P value	0.0034
Data	
Controls (Normal Controls)	12
Patients (Endo Pos)	14
Missing Controls	0
Missing Patients	10

35/35

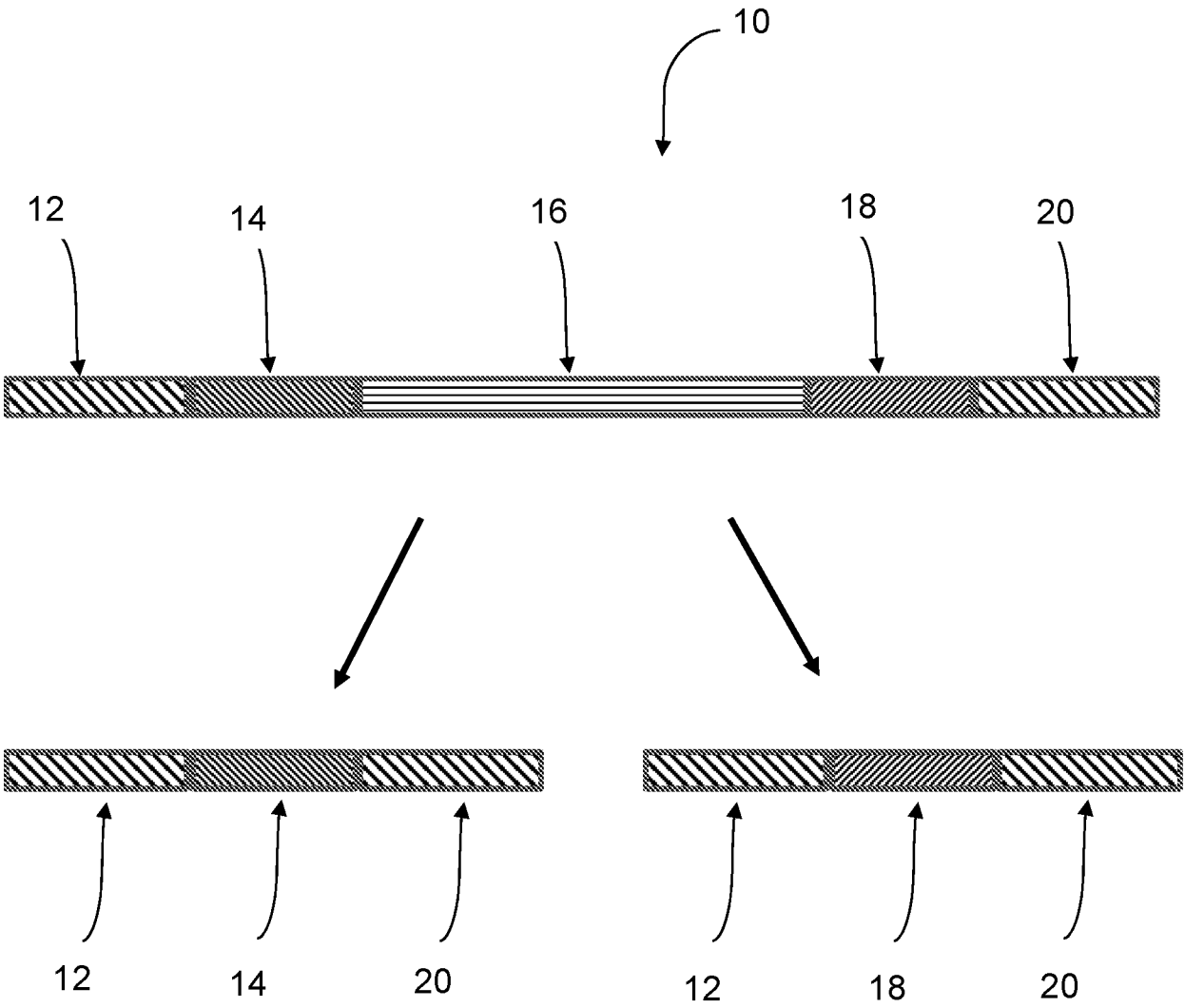


Fig. 20

SEQUENCE LISTING

<110> MDNA Life Sciences Inc.

<120> MITOCHONDRIAL DNA DELETIONS ASSOCIATED WITH ENDOMETRIOSIS

<130> 11871/00286

<150> 62/784,403

<151> 2018-12-22

<150> 62/931,173

<151> 2019-11-05

<160> 84

<170> PatentIn version 3.5

<210> 1

<211> 16569

<212> DNA

<213> Human

<220>

<221> misc_feature

<222> (3107)..(3107)

<223> Gap

<400> 1

gatcacaggt ctatcacctt attaacctt cacgggagct ctccatgcat ttggtatatt	60
cgtctggggg gtatgcacgc gatagcattg cgagacgctg gagccggagc accctatgtc	120
gcagtatctg tctttgattc ctgcctcatc ctattattta tcgcacctac gttcaatatt	180
acaggcgaac atacttacta aagtgtgtta attaatatt gcttgtagga cataataata	240
acaattgaat gtctgcacag ccactttcca cacagacatc ataacaaaaa atttccacca	300
aacccccct ccccgcttc tggccacagc acttaaacc atctctgcca aacccccaaa	360
acaaagaacc ctaacaccag cctaaccaga tttcaaattt tatcttttgg cggtatgcac	420
ttttaacagt cccccccaa ctaacacatt attttcccct cccactcca tactactaat	480
ctcatcaata caacccccgc ccactctacc cagcacacac acaccgctgc taaccccata	540
ccccgaacca accaaacccc aaagacaccc cccacagttt atgtagctta cctcctcaaa	600
gcaatacact gaaaatgttt agacgggctc acatcacccc ataaacaaat aggtttggtc	660
ctagcctttc tattagctct tagtaagatt acacatgcaa gcatccccgt tccagttagt	720

tcaccctcta aatcaccacg atcaaaagga acaagcatca agcacgcagc aatgcagctc	780
aaaacgctta gcctagccac acccccacgg gaaacagcag tgattaacct ttagcaataa	840
acgaaagttt aactaagcta tactaaccac agggttggtc aatttcgtgc cagccaccgc	900
ggtcacacga ttaacccaag tcaatagaag ccggcgtaaa gagtgtttta gatcaccccc	960
tccccaataa agctaaaact cacctgagtt gtaaaaaact ccagttgaca caaaatagac	1020
tacgaaagtg gctttaacat atctgaacac acaatagcta agacccaaac tgggattaga	1080
taccccacta tgcttagccc taaacctcaa cagttaaadc aacaaaactg ctcgccagaa	1140
cactacgagc cacagcttaa aactcaaagg acctggcggg gcttcataat cctctagagg	1200
agcctgttct gtaatcgata aaccccgatc aacctcacca cctcttgctc agcctatata	1260
ccgccatctt cagcaaacc tgatgaaggc tacaagtaa gcgcaagtac ccacgtaaag	1320
acgttaggtc aaggtgtagc catgaggtg gcaagaaatg ggctacattt tctaccccag	1380
aaaactacga tagcccttat gaaacttaag ggtcgaagg ggtatttagc gtaactaag	1440
agtagagtgc ttagttgaac agggccctga agcgcgta caaccgcccgt caccctctc	1500
aagtatactt caaaggacat ttaactaaaa cccctacgca tttatataga ggagacaagt	1560
cgtaacatgg taagtgtact ggaaagtgc cttggacgaa ccagagtgt gcttaacaca	1620
aagcacccaa cttacactta ggagatttca acttaacttg accgctctga gctaaaccta	1680
gccccaaacc cactccacct tactaccaga caaccttagc caaacattt acccaaataa	1740
agtataggcg atagaaattg aaacctggcg caatagatat agtaccgcaa gggaaagatg	1800
aaaaattata accaagcata atatagcaag gactaaccac tataccttct gcataatgaa	1860
ttaactagaa ataactttgc aaggagagcc aaagctaaga ccccgaaac cagacgagct	1920
acctaagaac agctaaaaga gcacaccgct ctatgtagca aaatagtggg aagatttata	1980
ggtagaggcg acaaacctac cgagcctggg gatagctggg tgtccaagat agaacttag	2040
ttcaacttta aatttgccca cagaaccctc taaatcccct tgtaaattta actgttagtc	2100
caaagaggaa cagctctttg gacactagga aaaaaccttg tagagagagt aaaaaattta	2160
acacccatag taggcctaaa agcagccacc aattaagaaa gcgttcaagc tcaacacca	2220
ctacctaata aatcccaaac atataactga actcctcaca cccaattgga ccaatctatc	2280
accctataga agaactaatg ttagtataag taacatgaaa acattctcct ccgcataagc	2340

ctgcgtcaga ttaaaacact gaactgacaa ttaacagccc aatatctaca atcaaccaac	2400
aagtcattat taccctcact gtcaacccaa cacaggcatg ctcataagga aaggttaaaa	2460
aaagtaaaag gaactcggca aatcttacct cgctgttta ccaaaaacat cacctctagc	2520
atcaccagta ttagaggcac cgctgccc gtgacacatg tttaacggcc gcggtaccct	2580
aaccgtgcaa aggtagcata atcacttggt ccttaaataag ggacctgtat gaatggctcc	2640
acgagggttc agctgtctct tacttttaac cagtgaatt gacctgccg tgaagaggcg	2700
ggcataacac agcaagacga gaagacccta tggagcttta atttattaat gcaaacagta	2760
cctaacaaac ccacaggctc taaactacca aacctgcatt aaaaatttcg gttggggcga	2820
cctcggagca gaaccaacc tccgagcagt acatgctaag acttcaccag tcaaagcgaa	2880
ctactatact caattgatcc aataacttga ccaacggaac aagttaccct agggataaca	2940
gcgcaatcct attctagagt ccatatcaac aatagggttt acgacctcga tgttgatca	3000
ggacatccc atggtgcagc cgctattaaa ggttcgtttg ttcaacgatt aaagtcctac	3060
gtgatctgag ttcagaccgg agtaatccag gtcggtttct atctacnttc aaattcctcc	3120
ctgtacgaaa ggacaagaga aataaggcct acttcacaaa gcgccttccc ccgtaaataga	3180
tatcatctca acttagtatt ataccacac ccaccaaga acagggtttg ttaagatggc	3240
agagcccgg aatcgcataa aacttaaaac tttacagtca gaggttcaat tcctcttctt	3300
aacaacatac ccatggccaa cctcctactc ctcatgtac ccattctaata cgcaatggca	3360
ttcctaatagc ttaccgaacg aaaaattcta ggctatatac aactacgcaa aggcccaac	3420
gttgtaggcc cctacgggct actacaacc ttcgtgacg ccataaaact cttaccaaaa	3480
gagcccctaa aaccgccac atctaccatc accctctaca tcaccgccc gacctagct	3540
ctcaccatcg ctcttctact atgaaccccc ctccccatac ccaaccccct ggtcaacctc	3600
aacctaggcc tcctatztat tctagccacc tctagcctag ccgtttactc aatcctctga	3660
tcagggtgag catcaaactc aaactacgcc ctgatcggcg cactgcgagc agtagcccaa	3720
acaatctcat atgaagtcac cctagccatc attctactat caacattact aataagtggc	3780
tcctttaacc tctccaccct tatcacaaca caagaacacc tctgattact cctgccatca	3840
tgacccttgg ccataatatg atttatctcc aactagcag agaccaaccg aacccccttc	3900
gaccttgccg aaggggagtc cgaactagtc tcaggcttca acatcgaata cgccgcaggc	3960

cccttcgccc tatttcttcat agccgaatac acaaacatta ttataataaa caccctcacc	4020
actacaatct tcctaggaac aacatatgac gcaactctccc ctgaactcta cacaacatat	4080
tttgtcacca agaccctact tctaacctcc ctgttcttat gaattcgaac agcatacccc	4140
cgattccgct acgaccaact catacacctc ctatgaaaaa acttcctacc actcaccta	4200
gcattactta tatgatatgt ctccataccc attacaatct ccagcattcc ccctcaaacc	4260
taagaaatat gtctgataaa agagttactt tgatagagta aataatagga gcttaaacc	4320
ccttatttct aggactatga gaatcgaacc catccctgag aatccaaaat tctccgtgcc	4380
acctatcaca ccccatccta aagtaaggtc agctaaataa gctatcgggc ccataccccg	4440
aaaatgttgg ttataccctt cccgtactaa ttaatccctt ggcccaacc gtcacttact	4500
ctaccatctt tgcaggcaca ctcatcacag cgctaagctc gcaactgattt ttacactgag	4560
taggcctaga aataaacatg ctagctttta ttccagttct aaccaaaaaa ataaaccctc	4620
gttccacaga agctgccatc aagtatttcc tcacgcaagc aaccgcatcc ataatccttc	4680
taatagctat cctcttcaac aatatactct ccggacaatg aaccataacc aatactacca	4740
atcaatactc atcattaata atcataatag ctatagcaat aaaactagga atagccccct	4800
ttcacttctg agtcccagag gttacccaag gcacccctct gacatccggc ctgcttcttc	4860
tcacatgaca aaaactagcc cccatctcaa tcatatacca aatctctccc tcactaaacg	4920
taagccttct cctcactctc tcaatcttat ccatcatagc aggcagttga ggtggattaa	4980
accaaacca gctacgcaa atcttagcat actcctcaat taccacata ggatgaataa	5040
tagcagttct accgtacaac cctaacataa ccattcttaa tttactatt tatattatcc	5100
taactactac cgcattccta ctactcaact taaactccag caccacgacc ctactactat	5160
ctcgcacctg aaacaagcta acatgactaa cacccttaat tccatccacc ctctctccc	5220
taggaggcct gccccgcta accggctttt tgcccaaagtg ggccattatc gaagaattca	5280
caaaaaacaa tagcctcatc atccccacca tcatagccac catcacctc cttaacctct	5340
acttctacct acgcctaatc tactccacct caatcacact actccccata tctaacaacg	5400
taaaaataaa atgacagttt gaacatacaa aaccaccccc attcctcccc acactcatcg	5460
cccttaccac gctactccta cctatctccc cttttatact aataatctta tagaaattta	5520
ggttaaatac agaccaagag ctttcaaagc cctcagtaag ttgcaatact taatttctgt	5580

aacagctaag gactgcaaaa cccactctg catcaactga acgcaaatca gccactttaa	5640
ttaagctaag cccttactag accaatggga cttaaaccga caaactta gttacagct	5700
aagcacccta atcaactggc ttcaatctac ttctcccgcc gccgggaaaa aaggcgggag	5760
aagccccggc aggtttgaag ctgcttcttc gaatttgcaa ttcaatatga aaatcacctc	5820
ggagctggta aaaagaggcc taaccctgt ctttagattt acagtccaat gcttcactca	5880
gccattttac ctcaccccca ctgatgttcg ccgaccgttg actattctct acaaaccaca	5940
aagacattgg aacactatac ctattattcg gcgcatgagc tggagtccta ggcacagctc	6000
taagcctcct tattcgagcc gagctgggcc agccaggcaa ctttctaggt aacgaccaca	6060
tctacaacgt tatcgtcaca gccatgcat ttgtaataat cttcttcata gtaataccca	6120
tcataatcgg aggctttggc aactgactag ttcccctaataatcgggtgcc cccgatatgg	6180
cgtttccccg cataaacaac ataagcttct gactcttacc tccctctctc ctactcctgc	6240
tcgcatctgc tatagtggag gccggagcag gaacagggtg aacagtctac cctcccttag	6300
cagggaacta ctcccaccct ggagcctccg tagacctaac catcttctcc ttacacctag	6360
cagggtgtctc ctctatctta ggggccatca atttcatcac aacaattatc aatataaaac	6420
cccctgccat aaccaatac caaacgcccc tcttcgtctg atccgtccta atcacagcag	6480
tcctacttct cctatctctc ccagtcctag ctgctggcat cactatacta ctaacagacc	6540
gcaacctcaa caccaccttc ttcgaccccg ccggaggagg agacccatt ctataccaac	6600
acctattctg atttttcggt caccctgaag tttatattct tatcctacca ggcttcggaa	6660
taatctcca tattgtaact tactactccg gaaaaaaga accatttgga tacataggta	6720
tggtctgagc tatgatatac attggcttcc tagggtttat cgtgtgagca caccatatac	6780
ttacagtagg aatagacgta gacacacgag catatttcac ctccgctacc ataatcatcg	6840
ctatccccac cggcgtcaaa gtatttagct gactcgccac actccacgga agcaatatga	6900
aatgatctgc tgcagtgtc tgagccctag gattcatctt tcttttcacc gtaggtggcc	6960
tgactggcat tgtattagca aatcatcac tagacatcgt actacacgac acgtactacg	7020
tttagccca cttccactat gtcctatcaa taggagctgt atttgccatc ataggaggct	7080
tcattcactg atttccccta ttctcaggct acaccctaga ccaaacctac gccaaaatcc	7140
atttcactat catattcatc ggcgtaaatc taactttctt cccacaacac tttctcggcc	7200

tatccggaat gccccgacgt tactcggact accccgatgc atacaccaca tgaaacatcc	7260
tatcatctgt aggctcattc atttctctaa cagcagtaat attaataatt ttcatgattt	7320
gagaagcctt cgcttcgaag cgaaaagtcc taatagtaga agaaccctcc ataaacctgg	7380
agtgactata tggatgcccc ccaccctacc acacattcga agaacccgta tacataaaat	7440
ctagacaaaa aaggaaggaa tcgaaccccc caaagctggg ttcaagccaa ccccatggcc	7500
tccatgactt tttcaaaaag gtattagaaa aaccatttca taactttgtc aaagttaaatt	7560
tataggctaa atcctatata tcttaatggc acatgcagcg caagtaggtc tacaagacgc	7620
tacttcccct atcatagaag agcttatcac ctttcatgat cagccctca taatcatttt	7680
ccttatctgc ttcctagtcc tgtatgccct tttcctaaca ctcaacaaca aactaactaa	7740
tactaacatc tcagacgctc aggaaataga aaccgtctga actatcctgc ccgcatcat	7800
cctagtcctc atcgccctcc catccctacg catcctttac ataacagacg aggtcaacga	7860
tccctccctt accatcaaatt caattggcca ccaatggtag tgaacctacg agtacaccga	7920
ctacggcgga ctaatcttca actcctacat acttccccca ttattcctag aaccaggcga	7980
cctgcgactc cttgacgttg acaatcgagt agtactcccg attgaagccc ccattcgtat	8040
aataattaca tcacaagacg tcttgactc atgagctgtc cccacattag gcttaaaaac	8100
agatgcaatt cccggacgct taaaccaaac cactttcacc gctacacgac cgggggtata	8160
ctacgggtcaa tgctctgaaa tctgtggagc aaaccacagt ttcatgcca tcgtcctaga	8220
attaattccc ctaaaaatct ttgaaatagg gcccgatatt accctatagc accccctcta	8280
ccccctctag agcccactgt aaagctaact tagcattaac cttttaagtt aaagattaag	8340
agaaccaaca cctctttaca gtgaaatgcc ccaactaaat actaccgtat ggcccacat	8400
aattaccccc atactcctta cactattcct catcacccaa ctaaaaatat taaacacaaa	8460
ctaccaccta cctccctcac caaagcccat aaaaataaaa aattataaca aaccctgaga	8520
acaaaaatga acgaaaatct gttcgcttca ttcattgccc ccacaatcct aggcctaccc	8580
gccgcagtac tgatcattct atttccccct ctattgatcc ccacctcaa atatctcatc	8640
aacaaccgac taatcaccac ccaacaatga ctaatcaaac taacctcaaa acaaatgata	8700
accatacaca aactaaagg acgaacctga tctcttatac tagtatcctt aatcattttt	8760
attgccacaa ctaacctcct cggactcctg cctcactcat ttacaccaac cacccaacta	8820

tctataaacc tagccatggc catcccccta tgagcgggca cagtgattat aggctttcgc	8880
tctaagatta aaaatgcctt agcccacttc ttaccacaag gcacacctac accccttatc	8940
cccatactag ttattatcga aaccatcagc ctactcattc aaccaatagc cctggccgta	9000
cgcctaaccg ctaacattac tgcaggccac ctactcatgc acctaattgg aagcgccacc	9060
ctagcaatat caaccattaa ccttccctct acacttatca tcttcacaat tctaattcta	9120
ctgactatcc tagaaatcgc tgtcgcctta atccaagcct acgttttcac acttctagta	9180
agcctctacc tgcacgacaa cacataatga cccaccaatc acatgcctat catatagtaa	9240
aaccagccc atgacccta acaggggccc tctcagccct cctaattgacc tccggcctag	9300
ccatgtgatt tcacttccac tccataacgc tcctcatact aggcctacta accaacacac	9360
taaccatata ccaatgatgg cgcgatgtaa cacgagaaag cacataccaa ggccaccaca	9420
caccacctgt ccaaaaaggc cttcgatacg ggataatcct atttattacc tcagaagttt	9480
ttttcttcgc aggatTTTTTc tgagcctttt accactccag cctagcccct accccccaat	9540
taggagggca ctggcccccac acaggcatca ccccgctaaa tcccctagaa gtcccactcc	9600
taaacacatc cgtattactc gcatcaggag tatcaatcac ctgagctcac catagtctaa	9660
tagaaaacaa ccgaaaccaa ataattcaag cactgcttat tacaatttta ctgggtctct	9720
attttaccct cctacaagcc tcagagtact tcgagtctcc cttcaccatt tccgacggca	9780
tctacggctc aacatttttt gtagccacag gcttccacgg acttcacgtc attattggct	9840
caactttcct cactatctgc ttcattccgc aactaatatt tcactttaca tccaaacatc	9900
actttggctt cgaagccgcc gcctgatact ggcatTTTgt agatgtgggt tgactatttc	9960
tgtatgtctc catctattga tgagggtctt actcttttag tataaatagt accgttaact	10020
tccaattaac tagttttgac aacattcaaa aaagagtaat aaacttcgcc ttaattttaa	10080
taatcaacac cctcctagcc ttactactaa taattattac attttgacta ccacaactca	10140
acggctacat agaaaaatcc accccttacg agtgcggctt cgaccctata tccccgcc	10200
gcgtcccttt ctccataaaa ttcttcttag tagctattac cttcttatta tttgatctag	10260
aaattgccct ctttttacct ctacatgag ccctacaaac aactaacctg ccactaatag	10320
ttatgtcatc cctcttatta atcatcatcc tagccctaag tctggcctat gagtgactac	10380
aaaaaggatt agactgaacc gaattggtat atagtttaaa caaaacgaat gatttcgact	10440

cattaaatta tgataatcat atttaccaaa tgccccctcat ttacataaat attataactag	10500
catttaccat ctactttcta ggaataactag tatatcgctc acacctcata tcctccctac	10560
tatgcctaga aggaataata ctatcgctgt tcattatagc tactctcata accctcaaca	10620
cccactccct cttagccaat attgtgccta ttgccatact agtcctttgcc gcctgcgaag	10680
cagcgggtggg cctagcccta ctagtctcaa tctccaacac atatggccta gactacgtac	10740
ataacctaaa cctactccaa tgctaaaact aatcggtcca acaattatat tactaccact	10800
gacatgactt tccaaaaaac acataatttg aatcaacaca accacccaca gcctaattat	10860
tagcatcatc cctctactat tttttaacca aatcaacaac aacctattta gctgttcccc	10920
aaccttttcc tccgaccccc taacaacccc cctcctaata ctaactacct gactcctacc	10980
cctcacaatc atggcaagcc aacgccactt atccagtga ccactatcac gaaaaaaact	11040
ctacctctct atactaatct ccctacaaat ctctttaatt ataacattca cagccacaga	11100
actaatcata ttttatatct tcttcgaaac cacacttatc cccaccttgg ctatcatcac	11160
ccgatgaggc aaccagccag aacgcctgaa cgcaggcaca tacttcctat tctacaccct	11220
agtaggctcc cttcccctac tcatcgact aatttacact cacaacaccc taggctcact	11280
aaacattcta ctactcactc tcaactgcca agaactatca aactcctgag ccaacaactt	11340
aatatgacta gcttacacaa tagcttttat agtaaagata cctctttacg gactccactt	11400
atgactccct aaagcccatg tcgaagcccc catcgctggg tcaatagtac ttgccgcagt	11460
actcttaaaa ctaggcggct atggtataat acgcctcaca ctcatctca accccctgac	11520
aaaacacata gcctaccct tccttgact atccctatga ggcataatta taacaagctc	11580
catctgccta cgacaaacag acctaaaatc gctcattgca tactcttcaa tcagccacat	11640
agccctcgta gtaacagcca ttctcatcca aacccccctga agcttcaccg gcgcagtcac	11700
tctcataatc gccacgggc ttacatctc attactattc tgcctagcaa actcaaacta	11760
cgaacgcact cacagtcgca tcataatcct ctctcaagga cttcaaactc tactcccact	11820
aatagctttt tgatgacttc tagcaagcct cgctaacctc gccttaccct ccactattaa	11880
cctactggga gaactctctg tgctagtaac cacgttctcc tgatcaaata tcactctcct	11940
acttacagga ctcaacatac tagtcacagc cctatactcc ctctacatat ttaccacaac	12000
acaatggggc tcactcacc accacattaa caacataaaa ccctcattca cacgagaaaa	12060

caccctcatg	ttcatacacc	tatcccccat	tctcctccta	tcctcaacc	ccgacatcat	12120
taccggggtt	tcctcttgta	aatatagttt	aacaaaaaca	tcagattgtg	aatctgacaa	12180
cagaggctta	cgaccctta	tttaccgaga	aagctcacia	gaactgctaa	ctcatgcccc	12240
catgtctaac	aacatggctt	tctcaacttt	taaaggataa	cagctatcca	ttgggtcttag	12300
gccccaaaaa	ttttggtgca	actccaaata	aaagtaataa	ccatgcacac	tactataacc	12360
accctaacc	tgacttcct	aattcccccc	atccttacca	ccctcgtaa	ccctaacaaa	12420
aaaaactcat	acccccatta	tgtaaaatcc	attgtcgcat	ccacctttat	tatcagtctc	12480
ttccccacaa	caatattcat	gtgcctagac	caagaagtta	ttatctcgaa	ctgacactga	12540
gccacaacc	aaacaacca	gctctcccta	agcttcaaac	tagactactt	ctccataata	12600
ttcatccctg	tagcattgtt	cgttacatgg	tccatcatag	aattctcact	gtgatataata	12660
aactcagacc	caaacattaa	tcagttcttc	aaatatctac	tcctcttcct	aattaccata	12720
ctaactcttag	ttaccgctaa	caacctattc	caactgttca	tcggctgaga	gggcgtagga	12780
attatatacct	tcttgctcat	cagttgatga	tacgcccagag	cagatgcaa	cacagcagcc	12840
attcaagcaa	tcctatacaa	ccgtatcggc	gatatacggtt	tcctcctcgc	cttagcatga	12900
tttatcctac	actccaactc	atgagacca	caacaaatag	cccttctaaa	cgctaatacca	12960
agcctcacc	cactactagg	cctcctccta	gcagcagcag	gcaaatacagc	ccaattaggt	13020
ctccaccct	gactcccctc	agccatagaa	ggccccacc	cagtctcagc	cctactccac	13080
tcaagcacta	tagttgtagc	aggaatcttc	ttactcatcc	gcttccacc	cctagcagaa	13140
aatagcccac	taatccaaac	tctaactacta	tgcttaggcg	ctatcaccac	tctgttcgca	13200
gcagtctgcg	cccttacaca	aaatgacatc	aaaaaaatcg	tagccttctc	cacttcaagt	13260
caactaggac	tcataatagt	tacaatcggc	atcaaccaac	cacacctagc	attcctgcac	13320
atctgtaccc	acgccttctt	caaagccata	ctatttatgt	gctccgggtc	catcatccac	13380
aaccttaaca	atgaacaaga	tattcgaaaa	ataggaggac	tactcaaaac	catacctctc	13440
acttcaacct	ccctcaccat	tggcagccta	gcattagcag	gaataccttt	cctcacaggt	13500
ttctactcca	aagaccacat	catcgaaacc	gcaaacatat	catacacaaa	cgcttgagcc	13560
ctatctatta	ctctcatcgc	tacctccctg	acaagcgcct	atagcactcg	aataattctt	13620
ctcaccctaa	caggtcaacc	tcgcttcccc	acccttacta	acattaacga	aaataacccc	13680

accctactaa	accccattaa	acgcctggca	gccggaagcc	tattcgcagg	atttctcatt	13740
actaacaaca	tttccccgc	atcccccttc	caaacaacaa	tccccctcta	cctaaaactc	13800
acagccctcg	ctgtcacttt	cctaggactt	ctaacagccc	tagacctcaa	ctacctaacc	13860
aacaaactta	aaataaaaatc	cccactatgc	acatttttatt	tctccaacat	actcggattc	13920
taccctagca	tcacacaccg	cacaatcccc	tatctaggcc	ttcttacgag	ccaaaacctg	13980
cccctactcc	tcctagacct	aacctgacta	gaaaagctat	tacctaaaac	aatttcacag	14040
caccaaactt	ccacctccat	catcacctca	acccaaaaag	gcataattaa	actttacttc	14100
ctctctttct	tcttcccact	catcctaacc	ctactcctaa	tcacataacc	tattcccccg	14160
agcaatctca	attacaatat	atacaccaac	aaacaatggt	caaccagtaa	ctactactaa	14220
tcaacgcca	taatcatata	aagccccgc	accaatagga	tcctcccgaa	tcaaccctga	14280
cccctctcct	tcataaatta	ttcagcttcc	tacactatta	aagttttacca	caaccaccac	14340
cccatcatac	tctttcaccc	acagcaccaa	tcctacctcc	atcgctaacc	ccactaaaac	14400
actaccaag	acctcaaccc	ctgacccccca	tgccctcagga	tactcctcaa	tagccatcgc	14460
tgtagtatat	caaagacaa	ccatcattcc	ccctaaataa	attaaaaaaaa	ctattaaacc	14520
catataacct	cccccaaat	tcagaataat	aacacacccg	accacaccgc	taacaatcaa	14580
tactaaacct	ccataaatag	gagaaggctt	agaagaaaac	cccacaaacc	ccattactaa	14640
accacactc	aacagaaaca	aagcatacat	cattattctc	gcacggacta	caaccacgac	14700
caatgatatg	aaaaaccatc	gttgtatttc	aactacaaga	acaccaatga	ccccaatagc	14760
caaaactaac	cccctaataa	aattaattaa	ccactcattc	atcgacctcc	ccaccccatc	14820
caacatctcc	gcatgatgaa	acttcggctc	actccttggc	gcctgcctga	tcctccaaat	14880
caccacagga	ctattcctag	ccatgcacta	ctcaccagac	gcctcaaccg	ccttttcatc	14940
aatcgccac	atcactcgag	acgtaaatta	tggctgaatc	atccgctacc	ttcacgcaa	15000
tggcgcctca	atattcttta	tctgcctctt	cctacacatc	gggcgaggcc	tatattacgg	15060
atcatttctc	tactcagaaa	cctgaaacat	cggcattatc	ctcctgcttg	caactatagc	15120
aacagccttc	ataggctatg	tcctcccgtg	aggccaaata	tcattctgag	gggccacagt	15180
aattacaaac	ttactatccg	ccatcccata	cattggggaca	gacctagttc	aatgaatctg	15240
aggaggctac	tcagtagaca	gtcccaccct	cacacgattc	tttacctttc	acttcatctt	15300

gcccttcatt attgcagccc tagcaacact ccacctccta ttcttgacg aaacgggatc	15360
aaacaacccc ctaggaatca cctcccattc cgataaaatc accttccacc cttactacac	15420
aatcaaagac gccctcgggt tactttctctt ctttctctcc ttaatgacat taacactatt	15480
ctcaccagac ctcttaggcg acccagacaa ttatacccta gcccaaccct taaacacccc	15540
tccccacatc aagcccgaat gatatttcct attgcctac acaattctcc gatccgtccc	15600
taacaaacta ggaggcgtcc ttgccttatt actatccatc ctcatcctag caataatccc	15660
catcctccat atatccaaac aacaaagcat aatatttcgc ccactaagcc aatcacttta	15720
ttgactccta gccgcagacc tcctcattct aacctgaatc ggaggacaac cagtaagcta	15780
cccttttacc atcattggac aagtagcatc cgtactatac ttcacaacaa tcctaatcct	15840
aataccaact atctccctaa ttgaaaacaa aataactcaa tgggcctgtc cttgtagtat	15900
aaactaatac accagtcttg taaaccggag atgaaaacct ttttccaagg acaaatcaga	15960
gaaaaagtct ttaactccac cattagcacc caaagctaag attctaattt aaactattct	16020
ctgttctttc atggggaagc agatttgggt accaccaag tattgactca cccatcaaca	16080
accgctatgt atttcgtaca ttactgccag ccaccatgaa tattgtacgg taccataaat	16140
acttgaccac ctgtagtaca taaaaacca atccacatca aaaccccctc cccatgctta	16200
caagcaagta cagcaatcaa ccctcaacta tcacacatca actgcaactc caaagccacc	16260
cctcaccac taggatacca acaaacctac ccacccttaa cagtacatag tacataaagc	16320
catttaccgt acatagcaca ttacagtcaa atcccttctc gtcccatgg atgaccccc	16380
tcagataggg gtcccttgac caccatcctc cgtgaaatca atatcccgca caagagtgt	16440
actctcctcg ctccgggccc ataacacttg ggggtagcta aagtgaactg tatccgacat	16500
ctggttccta cttcagggtc ataaagccta aatagcccac acgttcccct taaataagac	16560
atcacgatg	16569

<210> 2
 <211> 783
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (81)..(82)

<223> Positions 81 and 82 comprises the joining position between mutated left gene and mutated right gene

<400> 2

atggcccacc ataattaccc ccatactcct tacactattc ctcatcaccc aactaaaaat	60
attaaacaca aactaccacc tacctccctc accattggca gcctagcatt agcaggaata	120
cctttcctca caggtttcta ctccaaagac cacatcatcg aaaccgcaaa catatcatac	180
acaaacgcct gagccctatc tattactctc atcgctacct ccctgacaag cgcctatagc	240
actcgaataa ttcttctcac cctaacaggt caacctcgct tccccaccct tactaacatt	300
aacgaaaata accccaccct actaaacccc attaaacgcc tggcagccgg aagcctattc	360
gcaggatttc tcattactaa caacatttcc cccgcatccc ctttccaaac aacaatcccc	420
ctctacctaa aactcacagc cctcgctgtc actttcctag gatttctaac agccctagac	480
ctcaactacc taaccaacaa acttaaaata aaatccccac tatgcacatt ttatttctcc	540
aacatactcg gattctaccc tagcatcaca caccgcacaa tcccctatct aggccttctt	600
acgagccaaa acctgcccct actcctccta gacctaacct gactagaaaa gctattacct	660
aaaacaattt cacagcacca aatctccacc tccatcatca cctcaacca aaaaggcata	720
attaaacttt acttcctctc tttcttcttc ccactcatcc taaccctact cctaatacaca	780
taa	783

<210> 3

<211> 565

<212> DNA

<213> Human

<220>

<221> misc_feature

<222> (407)..(408)

<223> Positions 407 and 408 comprises the joining position between mutated left gene and mutated right gene

<400> 3

atggcacatg cagcgcaagt aggtctacaa gacgctactt cccctatcat agaagagctt	60
atcacctttc atgatcacgc cctcataatc attttcctta tctgcttcct agtcctgtat	120
gcccttttcc taacactcac aacaaaacta actaatacta acatctcaga cgctcaggaa	180
atagaaaccg tctgaactat cctgcccgcc atcatcctag tcctcatcgc cctcccatcc	240

ctacgcatcc tttacataac agacgaggtc aacgatccct cccttaccat caaatcaatt	300
ggccaccaat ggtactgaac ctacgagtac accgactacg gcggactaat cttcaactcc	360
tacatacttc cccattatt cctagaacca ggcgacctgc gactcctagc cgagacctc	420
ctcattctaa cctgaatcgg aggacaacca gtaagctacc cttttaccat cattggacaa	480
gtagcatccg tactatactt cacaacaatc ctaatcctaa taccaactat ctccctaatt	540
gaaaacaaaa tactcaaatg ggcct	565

<210> 4
 <211> 1905
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (665)..(666)
 <223> Positions 665 and 666 comprises the joining position between
 mutated left gene and mutated right gene

<400> 4	
atgaacgaaa atctgttcgc ttcattcatt gccccacaa tcctaggcct acccgccgca	60
gtactgatca ttctatttcc ccctctattg atccccacct ccaaatatct catcaacaac	120
cgactaatca ccaccaaca atgactaatc aaactaacct caaaacaaat gataaccata	180
cacaacacta aaggacgaac ctgatctctt atactagtat ccttaatcat ttttattgcc	240
acaactaacc tcctcggact cctgcctcac tcattttacac caaccacca actatctata	300
aacctagcca tggccatccc cttatgagcg ggcacagtga ttataggcct tcgctctaag	360
attaaaaatg ccctagccca cttcttacca caaggcacac ctacaccct tatccccata	420
ctagttatta tcgaaaccat cagcctactc attcaaccaa tagccctggc cgtacgccta	480
accgctaaca ttactgcagg ccacctactc atgcacctaa ttggaagcgc caccctagca	540
atatcaacca ttaaccttcc ctctacactt atcatcttca caattctaatt tctactgact	600
atcctagaaa tcgctgtcgc cttaatccaa gcctacgttt tcacacttct agtaagcctc	660
tacctacact ccaactcatg agaccacaaa caaatagccc ttctaaacgc taatccaagc	720
ctcaccac tactaggcct cctcctagca gcagcaggca aatcagccca attaggtctc	780
caccctgac tcccctcagc catagaaggc cccacccag tctcagccct actccactca	840

agcactatag ttgtagcagg aatctttctta ctcatccgct tccaccccct agcagaaaaat	900
agcccactaa tccaaactct aacactatgc ttaggcgcta tcaccactct gttcgcagca	960
gtctgcgccc ttacacaaaa tgacatcaaa aaaatcgtag ctttctccac ttcaagtcaa	1020
ctaggactca taatagttac aatcggcatc aaccaaccac acctagcatt cctgcacatc	1080
tgtaccacag ctttcttcaa agccatacta tttatgtgct ccgggtccat catccacaac	1140
cttaacaatg aacaagatat tcgaaaaata ggaggactac tcaaaaccat acctctcact	1200
tcaacctccc tcaccattgg cagcctagca ttagcaggaa tacctttcct cacaggtttc	1260
tactccaaag accacatcat cgaaaccgca aacatatcat acacaaacgc ctgagcccta	1320
tctattactc tcatcgctac ctccctgaca agcgcctata gcactcgaat aattcttctc	1380
accctaacag gtcaacctcg cttccccacc cttactaaca ttaacgaaaa taaccccacc	1440
ctactaaacc ccattaaacg cctggcagcc ggaagcctat tcgcaggatt tctcattact	1500
aacaacattt cccccgcatc ccccttccaa acaacaatcc ccctctacct aaaactcaca	1560
gccctcgctg tcactttcct aggacttcta acagccctag acctcaacta cctaaccaac	1620
aaacttaaaa taaaatcccc actatgcaca ttttatttct ccaacatact cggattctac	1680
cctagcatca cacaccgcac aatcccctat ctaggccttc ttacgagcca aaacctgccc	1740
ctactcctcc tagacctaac ctgactagaa aagctattac ctaaaacaat ttcacagcac	1800
caaatctcca cctccatcat cacctcaacc caaaaaggca taattaaaact ttacttcctc	1860
tctttcttct tcccactcat cctaacccta ctctaataca cataa	1905

<210> 5
 <211> 1905
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (662)..(663)
 <223> Positions 662 and 663 comprises the joining position between
 mutated left gene and mutated right gene

<400> 5	
atgaacgaaa atctgttcgc ttcatcatt gccccacaa tcctaggcct acccgccgca	60
gtactgatca ttctatttcc ccctctattg atccccacct ccaaatatct catcaacaac	120

cgactaatca ccaccaaca atgactaatc aaactaacct caaaacaaat gataaccata	180
cacaacacta aaggacgaac ctgatctctt atactagtat ccttaatcat ttttattgcc	240
acaactaacc tcctcggact cctgcctcac tcattttacac caaccaccca actatctata	300
aacctagcca tggccatccc cttatgagcg ggcacagtga ttataggcctt tcgctctaag	360
attaaaaatg ccctagccca cttctttacca caaggcacac ctacaccctt tatccccata	420
ctagttatta tcgaaaccat cagcctactc attcaaccaa tagccctggc cgtacgccta	480
accgctaaca ttactgcagg ccacctactc atgcacctaa ttggaagcgc caccctagca	540
atatcaacca ttaaccttcc ctctacactt atcatcttca caattctaata tctactgact	600
atcctagaaa tcgctgtcgc cttaatccaa gcctacgttt tcacacttct agtaagcctc	660
tacctacact ccaactcatg agaccacaaa caaatagccc ttctaaacgc taatccaagc	720
ctcacccac tactaggcct cctcctagca gcagcaggca aatcagccca attaggtctc	780
caccctgac tcccctcagc catagaaggc cccaccccag tctcagccct actccactca	840
agcactatag ttgtagcagg aatctttctta ctcatccgct tccaccccct agcagaaaat	900
agcccactaa tccaaactct aacactatgc ttaggcgcta tcaccactct gttcgcagca	960
gtctgcgccc ttacacaaaa tgacatcaaa aaaatcgtag ctttctccac ttcaagtcaa	1020
ctaggactca taatagttac aatcggcatc aaccaaccac acctagcatt cctgcacatc	1080
tgtaccacag ctttcttcaa agccatacta tttatgtgct ccgggtccat catccacaac	1140
cttaacaatg aacaagatat tcgaaaaata ggaggactac tcaaaacat acctctcact	1200
tcaacctccc tcaccattgg cagcctagca ttagcaggaa tacctttcct cacaggtttc	1260
tactccaaag accacatcat cgaaaccgca aacatatcat acacaaacgc ctgagcccta	1320
tctattactc tcatcgctac ctccctgaca agcgcctata gcactcgaat aattcttctc	1380
accctaacag gtcaacctcg cttccccacc cttactaaca ttaacgaaaa taaccccacc	1440
ctactaaacc ccattaaacg cctggcagcc ggaagcctat tcgcaggatt tctcattact	1500
aacaacattt ccccgcatc ccccttccaa acaacaatcc ccctctacct aaaactcaca	1560
gccctcgctg tcactttcct aggacttcta acagccctag acctcaacta cctaaccaac	1620
aaacttaaaa taaaatcccc actatgcaca ttttatttct ccaacatact cggattctac	1680
cctagcatca cacaccgcac aatcccctat ctaggccttc ttacgagcca aaacctgccc	1740

ctactcctcc tagacctaac ctgactagaa aagctattac ctaaaacaat ttcacagcac	1800
caaatctcca cctccatcat cacctcaacc caaaaaggca taattaaact ttacttcctc	1860
tctttcttct tcccactcat cctaacccta ctcctaata cataa	1905

<210> 6
 <211> 1629
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (309)..(310)
 <223> Positions 309 and 310 comprises the joining position between
 mutated left gene and mutated right gene

<400> 6	
ataaacttcg ccttaatttt aataatcaac accctcctag ccttactact aataattatt	60
acattttgac taccacaact caacggctac atagaaaaat ccaccctta cgagtgcggc	120
ttcgacccta tatccccgc cgcgctcct ttctccataa aattcttctt agtagctatt	180
accttcttat tatttgatct agaaattgcc ctccttttac ccctaccatg agccctacaa	240
acaactaacc tgccactaat agttatgtca tccctcttat taatcatcat cctagcccta	300
agtctggcca acacagcagc cattcaagca atcctataca accgtatcgg cgatatcgg	360
ttcatcctcg ccttagcatg atttataccta cactccaact catgagaccc acaacaaata	420
gcccttctaa acgctaatcc aagcctcacc ccactactag gcctcctcct agcagcagca	480
ggcaaatcag cccaattagg tctccacccc tgactcccct cagccataga aggccccacc	540
ccagtctcag ccctactcca ctcaagcact atagttgtag caggaatctt cttactcatc	600
cgcttccacc ccctagcaga aaatagccca ctaatccaaa ctctaact atgcttaggc	660
gctatcacca ctctgttcgc agcagtctgc gcccttacac aaaatgacat caaaaaaatc	720
gtagccttct ccacttcaag tcaactagga ctcataatag ttacaatcgg catcaaccaa	780
ccacacctag cattcctgca catctgtacc cagccttct tcaaagccat actatttatg	840
tgctccgggt ccatcatcca caaccttaac aatgaacaag atattcgaaa aataggagga	900
ctactcaaaa ccatacctct cacttcaacc tccctcacca ttggcagcct agcattagca	960
ggaatacctt tcctcacagg tttctactcc aaagaccaca tcatcgaaac cgcaaacata	1020

tcatacacaa acgcctgagc cctatctatt actctcatcg ctacctcct gacaagcgcc	1080
tatagcactc gaataattct tctcacccta acaggtcaac ctcgcttccc cacccttact	1140
aacattaacg aaaataaccc caccctacta aaccccattha aacgcctggc agccggaagc	1200
ctattcgag gatttctcat tactaacaac atttcccccg catccccctt ccaaacaaca	1260
atccccctct acctaaaact cacagccctc gctgtcactt tcctaggact tctaacagcc	1320
ctagacctca actacctaac caacaaactt aaaataaaat cccactatg cacattttat	1380
ttctccaaca tactcggatt ctaccctagc atcacacacc gcacaatccc ctatctaggc	1440
cttcttacga gccaaaacct gccctactc ctcttagacc taacctgact agaaaagcta	1500
ttacctaata caatttcaca gcaccaaata tccacctcca tcatcacctc aacccaaaaa	1560
ggcataatta aactttactt cctctctttc ttcttccac tcatcctaac cctactccta	1620
atcacataa	1629

<210> 7
 <211> 1692
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (357)..(358)
 <223> Positions 357 and 358 comprises the joining position between
 mutated left gene and mutated right gene

<400> 7	
atgttcgccg accgttgact atttcttaca aaccacaaag acattggaac actataccta	60
ttattcggcg catgagctgg agtcctaggc acagctctaa gcctccttat tcgagccgag	120
ctgggccagc caggcaacct tctaggtaac gaccacatct acaacgttat cgtcacagcc	180
catgcatttg taataatctt cttcatagta ataccatca taatcggagg ctttggcaac	240
tgactagttc ccctaataat cgggtgcccc gatatggcgt ttccccgcat aaacaacata	300
agcttctgac tcttacctcc ctctctccta ctctgctcg catctgctat agtggaggcc	360
cgagcagatg ccaacacagc agccattcaa gcaatcctat acaaccgtat cggcgatatc	420
ggtttcatcc tcgccttagc atgatttatc ctacactcca actcatgaga cccacaacaa	480
atagcccttc taaacgctaa tccaagcctc accccactac taggcctcct cctagcagca	540

gcaggcaaat cagcccaatt aggtctccac ccctgactcc cctcagccat agaaggcccc	600
accccagtct cagccctact ccactcaagc actatagttg tagcaggaat cttcttactc	660
atccgcttcc accccctagc agaaaatagc ccactaatcc aaactctaac actatgctta	720
ggcgctatca ccactctggt cgagcagtc tgcgccctta cacaaaatga catcaaaaaa	780
atcgtagcct tctccacttc aagtcaacta ggactcataa tagttacaat cggcatcaac	840
caaccacacc tagcattcct gcacatctgt acccagcct tcttcaaagc catactattt	900
atgtgctccg ggtccatcat ccacaacctt aacaatgaac aagatattcg aaaaatagga	960
ggactactca aaaccatacc tctcacttca acctccctca ccattggcag cctagcatta	1020
gcaggaatac ctttcctcac aggtttctac tccaaagacc acatcatcga aaccgcaaac	1080
atatcataca caaacgcctg agccctatct attacttca tcgctacctc cctgacaagc	1140
gcctatagca ctcgaataat tcttctcacc ctaacaggtc aacctcgctt cccaccctt	1200
actaacatta acgaaaataa cccacccta ctaaacccta ttaaagcctt ggagccgga	1260
agcctattcg caggatttct cattactaac aacatttccc ccgcatcccc cttccaaaca	1320
acaatcccc tctacctaaa actcacagcc ctgctgtca ctttcctagg acttctaaca	1380
gccctagacc tcaactacct aaccaacaaa cttaaaataa aatccccact atgcacattt	1440
tatttctcca acatactcgg attctaccct agcatcacac accgcacaat cccctatcta	1500
ggccttctta cgagccaaaa cctgccccta ctctcctag acctaacctg actagaaaag	1560
ctattaccta aaacaatttc acagcaccaa atctccacct ccatcatcac ctcaacccaa	1620
aaaggcataa ttaaaacttta cttcctctct ttcttcttcc cactcatcct aaccctactc	1680
ctaatcacat aa	1692

<210> 8
 <211> 572
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (387)..(388)
 <223> Positions 387 and 388 comprises the joining position between
 mutated left gene and mutated right gene

<400> 8

atggcacatg cagcgcaagt aggtctacaa gacgctactt cccctatcat agaagagctt	60
atcacctttc atgatacagc cctcataatc attttcctta tctgcttcct agtcctgtat	120
gcccttttcc taacactcac aacaaaacta actaatacta acatctcaga cgctcaggaa	180
atagaaaccg tctgaactat cctgcccgcc atcatcctag tcctcatcgc cctcccatcc	240
ctacgcatcc tttacataac agacgaggtc aacgatccct cccttaccat caaatcaatt	300
ggccaccaat ggtactgaac ctacgagtac accgactacg gcggactaat cttcaactcc	360
tacatacttc cccattatt cctagaacag gccacctact catgcaccta attggaagcg	420
ccaccctagc aatatcaacc attaaccttc cctctacact tatcatcttc acaattctaa	480
ttctactgac tatcctagaa atcgctgtcg ccttaatcca agcctacgtt ttcacacttc	540
tagtaagcct ctacctgcac gacaacacat aa	572

<210> 9
 <211> 652
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (560)..(561)
 <223> Positions 560 and 561 comprises the joining position between
 mutated left gene and mutated right gene

<400> 9	
atgaacgaaa atctgttcgc ttcattcatt gccccacaa tcctaggcct acccgccgca	60
gtactgatca ttctatttcc ccctctattg atccccacct ccaaatatct catcaacaac	120
cgactaatca ccaccaaca atgactaatc aaactaacct caaaacaaat gataaccata	180
cacaacacta aaggacgaac ctgatctctt atactagtat ccttaatcat ttttattgcc	240
acaactaacc tcctcggact cctgcctcac tcatttacac caaccacca actatctata	300
aacctagcca tggccatccc cttatgagcg ggcacagtga ttataggctt tcgctctaag	360
attaaaaatg ccctagccca cttcttacca caaggcacac ctacaccctt tatccccata	420
ctagttatta tcgaaacat cagcctactc attcaaccaa tagccctggc cgtacgccta	480
accgctaaca ttactgcagg ccacctactc atgcacctaa ttggaagcgc caccctagca	540
atatcaacca ttaaccttcc actaatagtt atgtcatccc tcttattaat catcatccta	600

gccctaagtc tggcctatga gtgactacaa aaaggattag actgaaccga at 652

<210> 10
<211> 1453
<212> DNA
<213> Human

<220>
<221> misc_feature
<222> (553)..(554)
<223> Positions 553 and 554 comprises the joining position between
mutated left gene and mutated right gene

<400> 10
atgaacgaaa atctgttcgc ttcattcatt gccccacaa tcctaggcct acccgccgca 60
gtactgatca ttctatttcc ccctctattg atccccacct ccaaatatct catcaacaac 120
cgactaatca ccaccaaca atgactaatc aaactaacct caaaacaaat gataaccata 180
cacaacacta aaggacgaac ctgatctctt atactagtat ccttaatcat ttttattgcc 240
acaactaacc tcctcggact cctgcctcac tcatttacac caaccacca actatctata 300
aacctagcca tggccatccc cttatgagcg ggcacagtga ttataggctt tcgctctaag 360
attaaaaatg ccctagccca cttcttacca caaggcacac ctacaccctt tatccccata 420
ctagttatta tcgaaacat cagcctactc attcaaccaa tagccctggc cgtacgccta 480
accgctaaca ttactgcagg ccacctactc atgcacctaa ttggaagcgc caccctagca 540
atatcaacca ttaaccttca cgccaatggc gcctcaatat tctttatctg cctcttccta 600
cacatcgggc gaggcctata ttacggatca tttctctact cagaaacctg aaacatcggc 660
attatcctcc tgcttgcaac tatagcaaca gccttcatag gctatgtcct cccgtgaggc 720
caaatatcat tctgaggggc cacagtaatt acaaacttac tatccgcat cccatacatt 780
gggacagacc tagttcaatg aatctgagga ggctactcag tagacagtcc caccctcaca 840
cgattcttta cttttcactt catcttgccc ttcattattg cagccctagc aacactccac 900
ctcctattct tgcacgaaac gggatcaaac aacccctag gaatcacctc ccattccgat 960
aaaatcacct tccaccctta ctacacaatc aaagacgccc tcggcttact tctcttcctt 1020
ctctccttaa tgacattaac actattctca ccagacctcc taggcgaccc agacaattat 1080
accctagcca accccttaaa caccctccc cacatcaagc ccgaatgata tttcctattc 1140

gcctacacaa ttctccgata cgtccctaac aaactaggag gcgtccttgc cctattacta	1200
tccatcctca tcctagcaat aatccccatc ctccatatat ccaaacaaca aagcataata	1260
tttcgcccac taagccaatc actttattga ctccatagccg cagacctcct cattctaacc	1320
tgaatcggag gacaaccagt aagctaccct ttaccatca ttggacaagt agcatccgta	1380
ctatacttca caacaatcct aatcctaata ccaactatct ccctaattga aaacaaaata	1440
ctcaaattggg cct	1453

<210> 11
 <211> 1705
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (1357)..(1358)
 <223> Positions 1357 and 1358 comprises the joining position between
 mutated left gene and mutated right gene

<400> 11	
atgttcgccg accgttgact atttcttaca aaccacaaag acattggaac actataccta	60
ttattcggcg catgagctgg agtcctaggc acagctctaa gcctccttat tcgagccgag	120
ctgggccagc caggcaacct tctaggtaac gaccacatct acaacgttat cgtcacagcc	180
catgcatttg taataatctt cttcatagta atacccatca taatcggagg ctttggcaac	240
tgactagttc ccctaataat cggtgccccc gatatggcgt ttccccgcat aaacaacata	300
agcttctgac tcttacctcc ctctctccta ctctgctcg catctgctat agtggaggcc	360
ggagcaggaa caggttgaac agtctaccct cccttagcag ggaactactc ccaccctgga	420
gcctccgtag acctaaccat cttctcctta cacctagcag gtgtctcctc tatcttaggg	480
gccatcaatt tcatcacaac aattatcaat ataaaacccc ctgccataac ccaataccea	540
acgcccctct tcgtctgata cgtcctaatac acagcagtcc tacttctcct atctctccca	600
gtcctagctg ctggcatcac tatactacta acagaccgca acctcaacac caccttcttc	660
gaccccgccg gaggaggaga cccatttcta taccaacacc tattctgatt tttcggtcac	720
cctgaagttt atattcttat cctaccaggc ttcggaataa tctcccatat tgtaacttac	780
tactccggaa aaaaagaacc atttggatac ataggtatgg tctgagctat gatatcaatt	840

ggcttcctag ggtttatcgt gtgagcacac catatattta cagtaggaat agacgtagac	900
acacgagcat atttcacctc cgctaccata atcatcgcta tccccaccgg cgtcaaagta	960
tttagctgac tcgccacact ccacggaagc aatatgaaat gatctgctgc agtgctctga	1020
gccctaggat tcatctttct tttcaccgta ggtggcctga ctggcattgt attagcaaac	1080
tcatcactag acatcgctact acacgacacg tactacgttg tagccactt ccactatgtc	1140
ctatcaatag gagctgtatt tgccatcata ggaggcttca ttcactgatt tcccctattc	1200
tcaggctaca ccctagacca aacctacgcc aaaatccatt tcactatcat attcatcggc	1260
gtaaatctaa ctttcttccc acaacacttt ctcggcctat ccggaatgcc ccgacgttac	1320
tcggactacc ccgatgcata caccacatga aacatccctc cccacatcaa gcccgaatga	1380
tatttcctat tcgcctacac aattctccga tccgtcccta acaaactagg aggcgtcctt	1440
gccctattac tatccatcct catcctagca ataatcccca tcctccatat atccaaacaa	1500
caaagcataa tatttcgccc actaagccaa tcactttatt gactcctagc cgacagacctc	1560
ctcattctaa cctgaatcgg aggacaacca gtaagctacc cttttaccat cattggacaa	1620
gtagcatccg tactatactt cacaacaatc ctaatcctaa taccaactat ctccctaatt	1680
gaaaacaaaa tactcaaatg ggcct	1705

<210> 12
 <211> 1361
 <212> DNA
 <213> Human

<220>
 <221> misc_feature
 <222> (66)..(67)
 <223> Positions 66 and 67 comprises the joining position between
 mutated left gene and mutated right gene

<400> 12	
atgccccaac taaatactac cgtatggccc accataatta ccccatact ccttacacta	60
ttcctcacca cccacagcct aatttttagca tcatccctct actatTTTTT aaccaaataca	120
acaacaacct atttagctgt tccccaacct tttcctccga ccccctaaca acccccctcc	180
taatactaac tacctgactc ctaccctca caatcatggc aagccaacgc cacttatcca	240
gtgaaccact atcacgaaaa aaactctacc tctctatact aatctcccta caaatctcct	300

taattataac attcacagcc acagaactaa tcatatttta tatcttcttc gaaaccacac	360
ttatccccac cttggctatc atcacccgat gaggcaacca gccagaacgc ctgaacgcag	420
gcacatactt cctattctac accctagtag gctcccttcc cctactcatc gcactaattt	480
acactcacia caccctaggg tcactaaaca ttctactact cactctcact gccaagaac	540
tatcaaactc ctgagccaac aacttaatag actagcttac acaatagctt ttatagtaaa	600
gatacctctt tacggactcc acttatgact ccctaaagcc catgtcgaag ccccatcgc	660
tgggtcaata gtacttgccg cagtactctt aaaactaggc ggctatggta taatacgcct	720
cacactcatt ctcaaccccc tgacaaaaca catagcctac cccttccttg tactatccct	780
atgaggcata attataacaa gctccatctg cctacgacaa acagacctaa aatcgctcat	840
tgcatactct tcaatcagcc acatagccct cgtagtaaca gccattctca tccaaacccc	900
ctgaagcttc accggcgcag tcatttctcat aatcgccac gggcttacat cctcattact	960
attctgccta gcaaactcaa actacgaacg cactcacagt cgcatacataa tcctctctca	1020
aggacttcaa actctactcc cactaatagc tttttgatga cttctagcaa gcctcgctaa	1080
cctcgcctta cccccacta ttaacctact gggagaactc tctgtgctag taaccacgtt	1140
ctcctgatca aatatcactc tcctacttac aggactcaac atactagtca cagccctata	1200
ctccctctac atattttacca caacacaatg gggctcactc acccaccaca ttaacaacat	1260
aaaaccctca ttcacacgag aaaacaccct catgttcata cacctatccc ccattctcct	1320
cctatccctc aaccccgaca tcattaccgg gttttcctct t	1361

<210> 13
 <211> 783
 <212> RNA
 <213> Human

<400> 13	
auggccacc auaauuaccc ccuauacuccu uacacuaauuc cucaucaccc aacuaaaaau	60
auuaaacaca aacuaccacc uaccuccuc accauuggca gccuagcauu agcaggaaua	120
ccuuuccuca cagguuucua cuccaaagac cacaucaucg aaaccgcaaa cauaucauac	180
acaaacgccu gagcccuauu uauuacucuc aucgcuaccu ccugacaag cgccuauagc	240
acucgaauaa uucuucucac ccuaacaggu caaccucgcu uccccacccu uacuaacauu	300
aacgaaaaua accccacccu acuaaacccc auuaaacgcc uggcagccgg aagccuauuc	360

gcaggauuuc ucauuacuaa caacauuucc cccgcauccc ccuuccaaac aacaaucccc	420
cucuaccuaa aacucacagc ccucgcuguc acuuuccuag gacuucuaac agcccuagac	480
cucaacuacc uaaccaaca acuuaaaaua aaauccccac uaugcacauu uuauuucucc	540
aacauacucg gauucuaccc uagcaucaca caccgcacaa uccccuaucu aggccuucuu	600
acgagccaaa accugccccc acuccuccua gaccuaaccu gacuagaaaa gcuauuaccu	660
aaaacaauuu cacagcacca aaucuccacc uccaucauca ccucaaccca aaaaggcaua	720
auuaaacuuu acuuccucuc uuucuucuuc ccacucaucc uaaccuacu ccuaaucaca	780
uaa	783

<210> 14
 <211> 565
 <212> RNA
 <213> Human

<400> 14 auggcacaug cagcgcaagu aggucuacaa gacgcuacuu ccccuaucau agaagagcuu	60
aucaccuuuc augaucacgc ccucauaaauc auuuuccuua ucugcuuccu aguccuguau	120
gccccuuuucc uaacacucac aacaaaacua acuaauacua acaucucaga cgcucaggaa	180
auagaaaccg ucugaacuau ccugcccgcc aucauccuag uccucaucgc ccucccaucc	240
cuacgcaucc uuuacauaac agacgagguc aacgaucuu cccuuaccu caaaucaauu	300
ggccaccaau gguacugaac cuacgaguac accgacuacg gcggacuau cuucaacucc	360
uacauacuuc ccccauuauu ccuagaacca ggcgaccugc gacuccuagc cgcagaccuc	420
cucauucuaa ccugaaucgg aggacaacca guaagcuacc cuuuuaccu cauuggacaa	480
guagcauccg uacuauacuu cacaacaau cuaauccuaa uaccaacuau cucccuauuu	540
gaaaacaaaa uacucaaaug ggccu	565

<210> 15
 <211> 1905
 <212> RNA
 <213> Human

<400> 15 augaacgaaa aucuguucgc uucauucuu gccccacaa uccuaggccu acccgccgca	60
guacugauca uucuaauucc ccucuaauug auccccaccu ccaauauacu caucaacaac	120

cgacuaauca	ccaccaaca	augacuaauc	aaacuaaccu	caaaacaaau	gauaaccaua	180
cacaacacua	aaggacgaac	cugaucucuu	auacuaguau	ccuuaaucau	uuuuauugcc	240
acaacuaacc	uccucggacu	ccugccucac	ucauuuacac	caaccaccca	acuaucuaua	300
aaccuagcca	uggccaucuu	cuaugagcg	ggcacaguga	uuauaggcuu	ucgcucuaag	360
auuaaaaaug	cccuagccca	cuucuuacca	caaggcacac	cuacaccccu	uaucuccaua	420
cuaguuaaua	ucgaaaccu	cagccuacuc	auucaacca	uagcccuggc	cguacgccua	480
accgcuaaca	uuacugcagg	ccaccuacuc	augcaccuaa	uuggaagcgc	cacccuagca	540
auaucaacca	uuaaccuucc	cucuacacuu	aucaucuua	caauucuaau	ucuacugacu	600
auccuagaaa	ucgcugucgc	cuuaauccaa	gccuacguuu	ucacacuucu	aguaagccuc	660
uaccuacacu	ccaacucaug	agaccacaa	caaauagccc	uucuaaacgc	uaauccaagc	720
cucacccac	uacuaggccu	ccuccuagca	gcagcaggca	aaucagccca	auuaggucuc	780
cacccugac	uccccucagc	cauagaaggc	cccaccccag	ucucagcccu	acuccacuca	840
agcacuauag	uuguagcagg	aaucuucuua	cucauccgcu	uccacccccc	agcagaaaau	900
agcccacuaa	uccaaacucu	aacacuaugc	uuaggcgcua	ucaccacucu	guucgcagca	960
gucugcgcgc	uuacacaaaa	ugacaucaaa	aaaaucguag	ccuucuccac	uucaagucaa	1020
cuaggacuca	uaauaguua	aaucggcauc	aaccaaccac	accuagcauu	ccugcacau	1080
uguaccacg	ccuucuuaa	agccauacua	uuuauugcu	ccggguccau	cauccacaac	1140
cuuaacaau	aacaagauu	ucgaaaaaua	ggaggacuac	uacaaaacca	accucucacu	1200
ucaaccuccc	ucaccauugg	cagccuagca	uuagcaggaa	uaccuuuccu	cacagguuuc	1260
uacuccaaag	accacaucau	cgaaacgcga	aacauaucu	acacaaacgc	cugagcccua	1320
ucuauuacuc	ucaucgcuac	cucccugaca	agcgccuaua	gcacucgaau	aauucuucuc	1380
accuaacag	gucaaccucg	cuuccccacc	cuuacuaaca	uuacgaaaa	uaaccccacc	1440
cuacuaaacc	ccauuaaacg	ccuggcagcc	ggaagccuau	ucgcaggauu	ucucauuacu	1500
aacaacauuu	cccccgcauc	cccuuccaa	acaacaaucc	cccucuaccu	aaaacucaca	1560
gcccucgcug	ucacuuccu	aggacuucua	acagcccuag	accucaacua	ccuaaccaac	1620
aaacuuaaaa	uaaaaucucc	acuaugcaca	uuuuauuuu	ccaacauacu	cggauucuac	1680
ccuagcauca	cacaccgcac	aaucuccuau	cuaggccuuc	uuacgagcca	aaaccugccc	1740

cuacuccucc	uagaccuaac	cugacuagaa	aagcuauuac	cuaaaacaau	uucacagcac	1800
caaaucucca	ccuccaucau	caccucaacc	caaaaaggca	uaauuaaacu	uuacuuccuc	1860
ucuuucuucu	ucccacucau	ccuaaccua	cuccuaauca	cauaa		1905

<210> 16
 <211> 1905
 <212> RNA
 <213> Human

<400> 16		
augaacgaaa	aucuguucgc	uucauucauu gccccacaa uccuaggccu acccgccgca 60
guacugauca	uucuauuucc	cccucuaauug auccccaccu ccaaaauaucu caucaacaac 120
cgacuaauca	ccaccaaca	augacuaauc aaacuaaccu caaaacaaau gauaaccuaa 180
cacaacacua	aaggacgaac	cugaucucuu auacuaguau ccuaaaucau uuuuauugcc 240
acaacuaacc	uccucggacu	ccugccucac ucauuuacac caaccacca acuaucuaa 300
aaccuagcca	uggccaucuu	cuaugagcg ggcacaguga uuauaggcuu ugcucuaag 360
auuaaaaaug	ccuagccca	cuucuuacca caaggcacac cuacaccccu uaucccaua 420
cuaguuaaua	ucgaaaccau	cagccuacuc auucaacca uagcccuggc cguacgccua 480
accgcuaaca	uuacugcagg	ccaccuacuc augcaccuaa uuggaagcgc caccuagca 540
auaucaacca	uuaaccuucc	cucuacacuu aucaucuua caauucuaau ucuacugacu 600
auccuagaaa	ucgcugucgc	cuuaauccaa gccuacguuu ucacacuucu aguaaggcuc 660
uaccuacacu	ccaacucaug	agaccacaa caaaugccc uucuaaacgc uaauccaagc 720
cucacccac	uacuaggccu	ccuccuagca gcagcaggca aaucagcca auuaggucuc 780
cacccugac	uccccucagc	cauagaaggc cccacccag ucucagccu acuccacua 840
agcacuauag	uuguagcagg	aaucuucuua cucauccgc uccaccccu agcagaaaau 900
agcccacuaa	uccaaacucu	aacacuaugc uuaggcgua ucaccacucu guucgcagca 960
gucugcgccc	uuacacaaaa	ugacaucaaa aaaauuguag ccuuccac uucaagucua 1020
cuaggacuca	uaauaguua	aaucggcauc aaccaaccac accuagcau ccugcacauc 1080
uguaccacg	ccuucuuaa	agccauacua uuuauugcu ccggguccau cauccacaac 1140
cuuaacaau	aacaagauu	ucgaaaaaau ggaggacuac ucaaaaccau accucucacu 1200

ucaaccuccc	ucaccauugg	cagccuagca	uuagcaggaa	uaccuuuccu	cacagguuuc	1260
uacuccaaag	accacaucau	cgaaacgcga	aacauaucau	acacaaacgc	cugagcccua	1320
ucuauuacuc	ucaucgcuac	cucccugaca	agcgccuaua	gcacucgaau	aauucuucuc	1380
accuaacag	gucaaccucg	cuuccccacc	cuuacuaaca	uuaacgaaaa	uaaccccacc	1440
cuacuaaacc	ccauuaaacg	ccuggcagcc	ggaagccuau	ucgcaggauu	ucucauuacu	1500
aacaacauuu	cccccgcauc	ccccuuccaa	acaacaaucc	cccucuaccu	aaaacucaca	1560
gcccucgcug	ucacuuuccu	aggacuucua	acagcccuag	accucaacua	ccuaaccaac	1620
aaacuuaaaa	uaaaaucccc	acuaugcaca	uuuuauuuuc	ccaacauacu	cggauucua	1680
ccuagcauca	cacaccgcac	aauccccuau	cuaggccuuc	uuacgagcca	aaaccugccc	1740
cuacuccucc	uagaccuaac	cugacuagaa	aagcuauuac	cuaaaacaau	uucacagcac	1800
caaaucucca	ccuccaucau	caccucaacc	caaaaaggca	uaauuaaacu	uuacuuccuc	1860
ucuuucuucu	ucccacucau	ccuaaccua	cuccuaauca	cauaa		1905

<210> 17
 <211> 1629
 <212> RNA
 <213> Human

<400> 17	
auaaacuucg	ccuuaauuuu
aaauaaucaac	acccuccuag
ccuuacuacu	aaauauuuuu
60	
acauuuugac	uaccacaacu
caacggcuac	auagaaaaau
ccaccccuua	cgaugcgggc
120	
uucgaccua	uaucggcg
ccgcuuccu	uucuccauaa
aaauucuucu	aguagcuauu
180	
accuucuau	uaauugaucu
agaaauugcc	cuccuuuuac
cccuaccaug	agcccuacaa
240	
acaacuaacc	ugccacuaau
aguuauguca	ucccucuau
uaaucaucau	ccuagcccua
300	
agucuggcca	acacagcagc
cauuaagca	auccuauaca
accguaucgg	cgauaucggu
360	
uucauccucg	ccuagcaug
auuuauccua	cacuccaacu
caugagaccc	acaacaaaua
420	
gcccuuucaa	acgcuaaucc
aagccucacc	ccacuacuag
gccuccuccu	agcagcagca
480	
ggcaauucag	cccauuuagg
ucuccacccc	ugacuccccu
cagccauaga	aggccccacc
540	
ccagucucag	cccuacucca
cucaagcacu	auaguuguag
caggaaucuu	cuuacucauc
600	
cgcuccacc	cccuagcaga
aaauagccca	cuaauccaaa
cucuaacacu	augcuuaggc
660	
gcuaucacca	cucuguucgc
agcagucugc	gcccuuacac
aaaaugacau	caaaaaaau
720	

guagccuucu ccacuucaag ucaacuagga cucauauag uuacaaucgg caucaaccaa	780
ccacaccuag cauuccugca caucuguacc cacgccuucu ucaaagccau acuauuuuag	840
ugcuccgggu ccaucaucca caaccuuaac aaugaacaag auauucgaaa aaaggagga	900
cuacucaaaa ccuauaccucu cacuuaacc ucccucacca uuggcagccu agcauuagca	960
ggaauaccuu uccucacagg uuucuacucc aaagaccaca ucaucgaaac cgcaaaca	1020
ucauacacaa acgccugagc ccuaucuauu acucucaucg cuaccucccu gacaagcgcc	1080
uauagcacuc gaauaaauucu ucucacccua acaggucaac cucgcuuccc caccuuacu	1140
aacauuaacg aaaauaaccc caccuacua aaccccauua aacgccuggc agccggaagc	1200
cuauucgcag gauuucucau uacuaacaac auuucccccg caucccccuu ccaaacaaca	1260
aucccccucu accuaaaacu cacagcccuc gcugucacuu uccuaggacu ucuaacagcc	1320
cuagaccuca acuaccuaac caacaaacu uaaaauaaaau ccccacuaug cacauuuuau	1380
uucuccaaca uacucggauu cuaccuagc aucacacacc gcacaauccc cuaucuaggc	1440
cuucuuacga gccaaaaccu gcccucacuc cuccuagacc uaaccugacu agaaaagcua	1500
uuaccuaaaa cauuucaca gcaccaaaucc uccaccucca ucaucaccuc aacccaaaaa	1560
ggcauaauua aacuuuacuu ccucucuuc uucuuccac ucauccuaac ccuacuccua	1620
aucacauaa	1629

<210> 18
 <211> 1692
 <212> RNA
 <213> Human

<400> 18 auguucgccg accguugacu auucucuaca aaccacaaag acauuggaac acuauaccua	60
uuauucggcg caugagcugg aguuccuaggc acagcucuaa gccuccuuau ucgagccgag	120
cugggccagc caggcaaccu ucuagguaac gaccacau cuacaacguuau cgucacagcc	180
caugcauuug uaauaauuu cuucauagua auaccuca uaaucggagg cuuuggcaac	240
ugacuaguuc ccuaauauu cggugccccc gauauggcgu uuccccgcau aaacaacaua	300
agcuucugac ucuuaccucc cucucuccua cuccugcucg caucugcuau aguggaggcc	360
cgagcagaug ccaacacagc agccauuca gcaauccuau acaaccguau cggcgauauc	420

gguuucaucc	ucgccuuagc	augauuuau	cuacacucca	acucaugaga	cccacaacaa	480
auagcccuuc	uaaacgcuaa	uccaagccuc	accccacuac	uaggccuccu	ccuagcagca	540
gcaggcaaa	cagcccaau	aggucuccac	cccugacucc	ccucagccau	agaaggcccc	600
accccagucu	cagcccuacu	ccacucaagc	acuauaguug	uagcaggaau	cuucuuacuc	660
auccgcuucc	accccuagc	agaaaauagc	ccacuaaucc	aaacucuaac	acuaugcuua	720
ggcgcuauca	ccacucuguu	cgcagcaguc	ugcgcccuua	cacaaaauga	caucaaaaaa	780
aucguagccu	ucuccacuuc	aagucaacua	ggacucauaa	uaguuacaau	cggcaucaac	840
caaccacacc	uagcauuccu	gcacaucugu	acccacgccu	ucuucaaagc	cauacuauuu	900
augugcuccg	gguccaucau	ccacaaccuu	aacaaugaac	aagauauucg	aaaaauagga	960
ggacuacuca	aaaccauacc	ucucacuuca	accucccuca	ccauuggcag	ccuagcauua	1020
gcaggaauac	cuuuccucac	agguuucuac	uccaaagacc	acaucaucga	aaccgcaaac	1080
auaucauaca	caaacgccug	agcccuaucu	auuacucuca	ucgcuaccuc	ccugacaagc	1140
gccuauagca	cucgaauaau	ucuucucacc	cuaacagguc	aaccucgcuu	ccccacccuu	1200
acuaacauua	acgaaaauaa	ccccacccua	cuaaacccca	uuaaacgccu	ggcagccgga	1260
agccuauucg	caggauuucu	cauuacuaac	aacauuuccc	ccgcaucccc	cuuccaaaca	1320
acaauccccc	ucuaccuaaa	acucacagcc	cucgcuguca	cuuuccuagg	acuucuaaca	1380
gcccuaagacc	ucaacuaccu	aaccaacaaa	cuuaaaauaa	aauccccacu	augcacauuu	1440
uauuucucca	acauacucgg	auucuacccu	agcaucacac	accgcacaau	ccccuaucua	1500
ggccuucuuu	cgagccaaaa	ccugccccua	cuccuccuag	accuaaccug	acuagaaaag	1560
cuauuaccua	aaacaauuuc	acagcaccaa	aucuccaccu	ccaucaucac	cucaacccaa	1620
aaaggcauaa	uuaaacuuua	cuuccucucu	uucuucuucc	cacucauccu	aaccuacuc	1680
cuaaucacau	aa					1692

<210> 19
 <211> 572
 <212> RNA
 <213> Human

<400> 19						
auggcacaug	cagcgcaagu	aggucuacaa	gacgcuacuu	ccccuaucau	agaagagcuu	60
aucaccuuuc	augaucacgc	ccucauaauc	auuuuccuua	ucugcuuccu	aguuccuguau	120

gccccuuuucc uaacacucac aacaaaaacua acuaauacua acaucucaga cgcucaggaa	180
auagaaaccg ucugaacuau ccugcccgcc aucauccuag uccucaucgc ccucccaucc	240
cuacgcaucc uuuacauaac agacgagguc aacgaucucu cccuuaccuau caaaucaauu	300
ggccaccaau gguacugaac cuacgaguac accgacuacg gcggacuauu cuucaacucc	360
uacauacuuc ccccauuuuu ccuagaacag gccaccuacu caugcaccua auuggaagcg	420
ccaccuagc aauaucaacc auuaaccuuc ccucuacacu uaucaucuuc acaauucuaa	480
uucuacugac uauccuagaa aucgcugucg ccuaaucca agccuacguu uucacacuuc	540
uaguaagccu cuaccugcac gacaacacau aa	572

<210> 20
 <211> 652
 <212> RNA
 <213> Human

<400> 20	
augaacgaaa aucuguucgc uucauucuu gccccacaa uccuaggccu acccgccgca	60
guacugauca uucuauuucc cccucuauug auccccaccu ccaaauaucu caucaacaac	120
cgacuaauca ccaccaaca augacuaauc aaacuaaccu caaaacaaau gauaaccuaa	180
cacaacacua aaggacgaac cugaucucuu auacuaguau ccuaaucau uuuuuauugcc	240
acaacuaacc uccucggacu ccugccucac ucauuuacac caaccaccca acuaucuaa	300
aaccuagcca uggccaucuu cuuaugagcg ggcacaguga uuauaggcuu ugcucuaag	360
auuaaaaaug cccuagccca cuucuuaacca caaggcacac cuacaccccu uaucccca	420
cuaguuaaua ucgaaaccu cagccuacuc auucaacca uagcccuggc cguacgccua	480
accgcuaaca uuacugcagg ccaccuacuc augcaccuaa uuggaagcg caccuagca	540
auaucaacca uuaaccuucc acuaauagu auucauucc ucuuauuaa caucauccu	600
gcccuaaguc uggccuaua gugacuaca aaaggauuag acugaaccga au	652

<210> 21
 <211> 1453
 <212> RNA
 <213> Human

<400> 21	
augaacgaaa aucuguucgc uucauucuu gccccacaa uccuaggccu acccgccgca	60

guacugauca uucuauuuucc cccucuaauug auccccaccu ccaaauaucu caucaacaac	120
cgacuaauca ccaccaaca augacuaauc aaacuaaccu caaaacaaau gauaaccaua	180
cacaacacua aaggacgaac cugaucucuu auacuaguau ccuaaucau uuuuauugcc	240
acaacuaacc uccucggacu ccugccucac ucauuuacac caaccacca acuaucuaa	300
aaccuagcca uggccauccc cuuagagcg ggcacaguga uuauaggcuu ugcucuaag	360
auuaaaaug cccuagccca cuucuuacca caaggcacac cuacaccccu uaucccaua	420
cuaguuaaua ucgaaaccu cagccuacuc auucaacca uagcccuggc cguacgccua	480
accgcuaaca uuacugcagg ccaccuacuc augcaccuaa uuggaagcgc caccuagca	540
auaucaacca uuaaccuua cgccauggc gccucaauau ucuuuauucg ccucuuccua	600
cacaucgggc gaggccuaa uuacggauca uuucucuacu cagaaaccug aaacaucggc	660
auuauccucc ugcuugcaac uauagcaaca gccuucuaug gcuauguccu cccgugaggc	720
caaauaucu ucugaggggc cacaguaau acaaacuac uauccgccau cccauacauu	780
gggacagacc uaguucaaug aaucugagga ggcuaucucag uagacagucc caccucaca	840
cgauucuuua ccuucacuu caucugccc uucauuauug cagcccuagc aacacuccac	900
cuccuauucu ugcacgaaac gggaucaaac aaccccuag gaucaccuc ccuuccgau	960
aaaauaccu uccaccuuu cuacacaau aaagacgcc ucggcuuacu ucucuuccuu	1020
cucuccuuaa ugacauaac acuaauuca ccagaccucc uaggcgaccc agacaauuau	1080
accuagcca accccuuaa caccucucc cacaucaagc ccgaugaua uuuccuauuc	1140
gccuacaca uucuccgauc cguccuaac aaacuaggag gcguccuugc ccuauuacua	1200
uccauccua uccuagcau aaucaccau cuccauauau ccaaacaaca aagcauaaua	1260
uuucgcccac uaagccaau acuuuauuga cuccuagccg cagaccuccu cauucuaacc	1320
ugaucggag gacaaccagu aagcuaccu uuuaaccau uuggacaagu agcauccgua	1380
cuauacuua caacaauccu aauccuaau ccaacuauu cccuaauuga aaacaaaaua	1440
cucaauuggg ccu	1453

<210> 22
 <211> 1705
 <212> RNA
 <213> Human

<400> 22

auguucgccg accguugacu auucucuaca aaccacaaag acauuggaac acuaauaccua	60
uuauucggcg caugagcugg aguccuaggc acagcucuaa gccuccuuau ucgagccgag	120
cugggccagc caggcaaccu ucuagguaac gaccacau cuuacguuau cgucacagcc	180
caugcauuug uaauaauuu cuucauagua auacccauca uauucggagg cuuuggcaac	240
ugacuaguuc cccuaauaau cggugccccc gauauggcgu uucggcgcau aaacaacaua	300
agcuucugac ucuuaccucc cucucuccua cuccugcucg caucugcuau aguggaggcc	360
ggagcaggaa cagguugaac agucuacccu cccuuagcag ggaacuacuc ccaccucgga	420
gccuccguag accuaaccu cuucuccua caccuagcag gugucuccuc uauuuaggg	480
gccaucauu ucaucacaac aaauaucaau auaaaacccc cugccauaac ccaauaccaa	540
acgccccucu ucgucugauc cguccuaa ucagcagucc uacuucuccu aucucuccca	600
guccuagcug cuggcaucac uauacuacua acagaccgca accucaacac caccuucuc	660
gaccccgccg gaggaggaga ccccauucua uaccaacacc uauucugauu uuucggucac	720
ccugaaguuu auauucuuau ccuaccaggc uucggauaau ucucccauau uguaacuuac	780
uacuccggaa aaaaagaacc auuuggauac auagguaugg ucugagcuau gauaucauu	840
ggcuuccuag gguuuaucu gugagcacac cauauuuua caguaggau agacguagac	900
acacgagcau auuucaccuc cgcuaccua aucaucgcu uccccaccgg cgucaaagua	960
uuuagcugac ucgccacacu ccacggaagc aaauagaaau gaucugcugc agugcucuga	1020
gcccuaaggau ucaucuuucu uuucaccgua gguggccuga cuggcauugu auuagcaaac	1080
ucaucacuag acaucguacu acacgacag uacuacguug uagcccacuu ccacuaguc	1140
cuaucaauag gagcuguauu ugccaucaua ggaggcuuca uucacugauu uccccuauuc	1200
ucaggcuaca cccuagacca aaccuacgcc aaaauccauu ucacuauc auucaucggc	1260
guaaaucuaa cuuucuuucc acaacacuuu cucggccuau ccggaugcc ccgacguuac	1320
ucggacuacc ccgaugcaua caccacauga aacaucuccc cccacaucaa gcccgauga	1380
uauuuuccuau ucgccuacac aaucuccga uccguccua acaaacuagg aggcguccu	1440
gcccuaauac uauccaucu cauccuagca auaucccca uccuccauau auccaacaa	1500
caaagcauaa uauuucgccc acuaagccaa ucacuuuuu gacuccuagc cgagaccuc	1560

cucauucuaa ccugaaucgg aggacaacca guaagcuacc cuuuuaccuau cauuggacaa	1620
guagcauccg uacuauacuu cacaacaau cuaauccuaa uaccaacuau cucccuauuu	1680
gaaaacaaaa uacucaaaug ggccu	1705

<210> 23
 <211> 1362
 <212> RNA
 <213> Human

<400> 23	
augccccaac uaaauacuac cguaugggccc accauaauua cccccaauacu ccuuacacua	60
uuccucacca cccacagccu aauuauuagc aucaucccuc uacuauuuuu uaaccaaauc	120
aacaacaacc uauuuagcug uuccccaacc uuuuccuccg acccccuaac aacccccuc	180
cuaauacuaa cuaccugacu ccuaccccuc acaaucaugg caagccaacg ccacuuauc	240
agugaaccac uaucacgaaa aaaacucuac cucucuauac uaaucucccu aaaaucucc	300
uuauuuauaa cauucacagc cacagaacua aucauauuuu auaucuucuu cgaaaccaca	360
cuuaucccca ccuuggcuau caucacccga ugaggcaacc agccagaacg ccugaacgca	420
ggcacauacu uccuauucua caccuagua ggcucccuuc cccuacucau cgcacuaauu	480
uacacucaca acacccuagg cucacuaaac auucuacuac ucacucucac ugcccaagaa	540
cuaucaaacu ccugagccaa caacuuaaua gacuagcuua cacaauagcu uuuauaguaa	600
agauaccucu uuacggacuc cacuuauagc ucccuaaaagc ccaugucgaa gccccaucg	660
cugggucaau aguacuugcc gcaguacucu uaaaacuagg cggcuauuggu auauuacgcc	720
ucacacucau ucucaacccc cugacaaaac acauagccua ccccuuccuu guacuauccc	780
uauagggcua aauuauaaca agcuccaucu gccuacgaca aacagaccua aaucgcua	840
uugcauacuc uucaaucagc cacauagccc ucguaguaac agccauucuc auccaaaccc	900
ccugaagcuu caccggcgca gucauucua uaaucgccc cgggcuuaca uccucauuac	960
uauucugccu agcaaacuca aacuacgaac gcacucacag ucgcaucaua auccucucuc	1020
aaggacuua aacucucuc ccacuaauag cuuuuugaug acuuuagca agccucgcu	1080
accucgccuu acccccacu auuaaccuac ugggagaacu cucugugcua guaaccacgu	1140
ucuccugauc aaauaucacu cuccuacuua caggacuaa cauacuaguc acagcccuau	1200
acucccucua cauauuuacc acaacacaau ggggcucacu caccaccac auuaacaaca	1260

uaaaaccuc auucacacga gaaaacaccc ucauguucau acaccuaucc ccgauucucc 1320

uccuaucccu caaccccgac aucauuaccg gguuuuccuc uu 1362

<210> 24
<211> 260
<212> PRT
<213> Human

<400> 24

Met Ala His His Asn Tyr Pro His Thr Pro Tyr Thr Ile Pro His His
1 5 10 15

Pro Thr Lys Asn Ile Lys His Lys Leu Pro Pro Thr Ser Leu Thr Ile
20 25 30

Gly Ser Leu Ala Leu Ala Gly Met Pro Phe Leu Thr Gly Phe Tyr Ser
35 40 45

Lys Asp His Ile Ile Glu Thr Ala Asn Met Ser Tyr Thr Asn Ala Trp
50 55 60

Ala Leu Ser Ile Thr Leu Ile Ala Thr Ser Leu Thr Ser Ala Tyr Ser
65 70 75 80

Thr Arg Met Ile Leu Leu Thr Leu Thr Gly Gln Pro Arg Phe Pro Thr
85 90 95

Leu Thr Asn Ile Asn Glu Asn Asn Pro Thr Leu Leu Asn Pro Ile Lys
100 105 110

Arg Leu Ala Ala Gly Ser Leu Phe Ala Gly Phe Leu Ile Thr Asn Asn
115 120 125

Ile Ser Pro Ala Ser Pro Phe Gln Thr Thr Ile Pro Leu Tyr Leu Lys
130 135 140

Leu Thr Ala Leu Ala Val Thr Phe Leu Gly Leu Leu Thr Ala Leu Asp
145 150 155 160

Leu Asn Tyr Leu Thr Asn Lys Leu Lys Met Lys Ser Pro Leu Cys Thr

165

170

175

Phe Tyr Phe Ser Asn Met Leu Gly Phe Tyr Pro Ser Ile Thr His Arg
 180 185 190

Thr Ile Pro Tyr Leu Gly Leu Leu Thr Ser Gln Asn Leu Pro Leu Leu
 195 200 205

Leu Leu Asp Leu Thr Trp Leu Glu Lys Leu Leu Pro Lys Thr Ile Ser
 210 215 220

Gln His Gln Ile Ser Thr Ser Ile Ile Thr Ser Thr Gln Lys Gly Met
 225 230 235 240

Ile Lys Leu Tyr Phe Leu Ser Phe Phe Phe Pro Leu Ile Leu Thr Leu
 245 250 255

Leu Leu Ile Thr
 260

<210> 25
 <211> 189
 <212> PRT
 <213> Human

<220>
 <221> misc_feature
 <222> (189)..(189)
 <223> Xaa can be any naturally occurring amino acid

<400> 25

Met Ala His Ala Ala Gln Val Gly Leu Gln Asp Ala Thr Ser Pro Ile
 1 5 10 15

Met Glu Glu Leu Ile Thr Phe His Asp His Ala Leu Met Ile Ile Phe
 20 25 30

Leu Ile Cys Phe Leu Val Leu Tyr Ala Leu Phe Leu Thr Leu Thr Thr
 35 40 45

Lys Leu Thr Asn Thr Asn Ile Ser Asp Ala Gln Glu Met Glu Thr Val
 50 55 60

Trp Thr Ile Leu Pro Ala Ile Ile Leu Val Leu Ile Ala Leu Pro Ser
65 70 75 80

Leu Arg Ile Leu Tyr Met Thr Asp Glu Val Asn Asp Pro Ser Leu Thr
85 90 95

Ile Lys Ser Ile Gly His Gln Trp Tyr Trp Thr Tyr Glu Tyr Thr Asp
100 105 110

Tyr Gly Gly Leu Ile Phe Asn Ser Tyr Met Leu Pro Pro Leu Phe Leu
115 120 125

Glu Pro Gly Asp Leu Arg Leu Leu Ala Ala Asp Leu Leu Ile Leu Thr
130 135 140

Trp Ile Gly Gly Gln Pro Val Ser Tyr Pro Phe Thr Ile Ile Gly Gln
145 150 155 160

Val Ala Ser Val Leu Tyr Phe Thr Thr Ile Leu Ile Leu Met Pro Thr
165 170 175

Ile Ser Leu Ile Glu Asn Lys Met Leu Lys Trp Ala Xaa
180 185

<210> 26
<211> 634
<212> PRT
<213> Human

<400> 26

Met Asn Glu Asn Leu Phe Ala Ser Phe Ile Ala Pro Thr Ile Leu Gly
1 5 10 15

Leu Pro Ala Ala Val Leu Ile Ile Leu Phe Pro Pro Leu Leu Ile Pro
20 25 30

Thr Ser Lys Tyr Leu Ile Asn Asn Arg Leu Ile Thr Thr Gln Gln Trp
35 40 45

Leu Ile Lys Leu Thr Ser Lys Gln Met Met Thr Met His Asn Thr Lys

50

55

60

Gly Arg Thr Trp Ser Leu Met Leu Val Ser Leu Ile Ile Phe Ile Ala
65 70 75 80

Thr Thr Asn Leu Leu Gly Leu Leu Pro His Ser Phe Thr Pro Thr Thr
85 90 95

Gln Leu Ser Met Asn Leu Ala Met Ala Ile Pro Leu Trp Ala Gly Thr
100 105 110

Val Ile Met Gly Phe Arg Ser Lys Ile Lys Asn Ala Leu Ala His Phe
115 120 125

Leu Pro Gln Gly Thr Pro Thr Pro Leu Ile Pro Met Leu Val Ile Ile
130 135 140

Glu Thr Ile Ser Leu Leu Ile Gln Pro Met Ala Leu Ala Val Arg Leu
145 150 155 160

Thr Ala Asn Ile Thr Ala Gly His Leu Leu Met His Leu Ile Gly Ser
165 170 175

Ala Thr Leu Ala Met Ser Thr Ile Asn Leu Pro Ser Thr Leu Ile Ile
180 185 190

Phe Thr Ile Leu Ile Leu Leu Thr Ile Leu Glu Ile Ala Val Ala Leu
195 200 205

Ile Gln Ala Tyr Val Phe Thr Leu Leu Val Ser Leu Tyr Leu His Ser
210 215 220

Asn Ser Trp Asp Pro Gln Gln Met Ala Leu Leu Asn Ala Asn Pro Ser
225 230 235 240

Leu Thr Pro Leu Leu Gly Leu Leu Leu Ala Ala Ala Gly Lys Ser Ala
245 250 255

Gln Leu Gly Leu His Pro Trp Leu Pro Ser Ala Met Glu Gly Pro Thr
260 265 270

Pro Val Ser Ala Leu Leu His Ser Ser Thr Met Val Val Ala Gly Ile
275 280 285

Phe Leu Leu Ile Arg Phe His Pro Leu Ala Glu Asn Ser Pro Leu Ile
290 295 300

Gln Thr Leu Thr Leu Cys Leu Gly Ala Ile Thr Thr Leu Phe Ala Ala
305 310 315 320

Val Cys Ala Leu Thr Gln Asn Asp Ile Lys Lys Ile Val Ala Phe Ser
325 330 335

Thr Ser Ser Gln Leu Gly Leu Met Met Val Thr Ile Gly Ile Asn Gln
340 345 350

Pro His Leu Ala Phe Leu His Ile Cys Thr His Ala Phe Phe Lys Ala
355 360 365

Met Leu Phe Met Cys Ser Gly Ser Ile Ile His Asn Leu Asn Asn Glu
370 375 380

Gln Asp Ile Arg Lys Met Gly Gly Leu Leu Lys Thr Met Pro Leu Thr
385 390 395 400

Ser Thr Ser Leu Thr Ile Gly Ser Leu Ala Leu Ala Gly Met Pro Phe
405 410 415

Leu Thr Gly Phe Tyr Ser Lys Asp His Ile Ile Glu Thr Ala Asn Met
420 425 430

Ser Tyr Thr Asn Ala Trp Ala Leu Ser Ile Thr Leu Ile Ala Thr Ser
435 440 445

Leu Thr Ser Ala Tyr Ser Thr Arg Met Ile Leu Leu Thr Leu Thr Gly
450 455 460

Gln Pro Arg Phe Pro Thr Leu Thr Asn Ile Asn Glu Asn Asn Pro Thr
465 470 475 480

Leu Leu Asn Pro Ile Lys Arg Leu Ala Ala Gly Ser Leu Phe Ala Gly

485

490

495

Phe Leu Ile Thr Asn Asn Ile Ser Pro Ala Ser Pro Phe Gln Thr Thr
 500 505 510

Ile Pro Leu Tyr Leu Lys Leu Thr Ala Leu Ala Val Thr Phe Leu Gly
 515 520 525

Leu Leu Thr Ala Leu Asp Leu Asn Tyr Leu Thr Asn Lys Leu Lys Met
 530 535 540

Lys Ser Pro Leu Cys Thr Phe Tyr Phe Ser Asn Met Leu Gly Phe Tyr
 545 550 555 560

Pro Ser Ile Thr His Arg Thr Ile Pro Tyr Leu Gly Leu Leu Thr Ser
 565 570 575

Gln Asn Leu Pro Leu Leu Leu Leu Asp Leu Thr Trp Leu Glu Lys Leu
 580 585 590

Leu Pro Lys Thr Ile Ser Gln His Gln Ile Ser Thr Ser Ile Ile Thr
 595 600 605

Ser Thr Gln Lys Gly Met Ile Lys Leu Tyr Phe Leu Ser Phe Phe Phe
 610 615 620

Pro Leu Ile Leu Thr Leu Leu Leu Ile Thr
 625 630

<210> 27
 <211> 634
 <212> PRT
 <213> Human

<400> 27

Met Asn Glu Asn Leu Phe Ala Ser Phe Ile Ala Pro Thr Ile Leu Gly
 1 5 10 15

Leu Pro Ala Ala Val Leu Ile Ile Leu Phe Pro Pro Leu Leu Ile Pro
 20 25 30

Thr Ser Lys Tyr Leu Ile Asn Asn Arg Leu Ile Thr Thr Gln Gln Trp
35 40 45

Leu Ile Lys Leu Thr Ser Lys Gln Met Met Thr Met His Asn Thr Lys
50 55 60

Gly Arg Thr Trp Ser Leu Met Leu Val Ser Leu Ile Ile Phe Ile Ala
65 70 75 80

Thr Thr Asn Leu Leu Gly Leu Leu Pro His Ser Phe Thr Pro Thr Thr
85 90 95

Gln Leu Ser Met Asn Leu Ala Met Ala Ile Pro Leu Trp Ala Gly Thr
100 105 110

Val Ile Met Gly Phe Arg Ser Lys Ile Lys Asn Ala Leu Ala His Phe
115 120 125

Leu Pro Gln Gly Thr Pro Thr Pro Leu Ile Pro Met Leu Val Ile Ile
130 135 140

Glu Thr Ile Ser Leu Leu Ile Gln Pro Met Ala Leu Ala Val Arg Leu
145 150 155 160

Thr Ala Asn Ile Thr Ala Gly His Leu Leu Met His Leu Ile Gly Ser
165 170 175

Ala Thr Leu Ala Met Ser Thr Ile Asn Leu Pro Ser Thr Leu Ile Ile
180 185 190

Phe Thr Ile Leu Ile Leu Leu Thr Ile Leu Glu Ile Ala Val Ala Leu
195 200 205

Ile Gln Ala Tyr Val Phe Thr Leu Leu Val Ser Leu Tyr Leu His Ser
210 215 220

Asn Ser Trp Asp Pro Gln Gln Met Ala Leu Leu Asn Ala Asn Pro Ser
225 230 235 240

Leu Thr Pro Leu Leu Gly Leu Leu Leu Ala Ala Ala Gly Lys Ser Ala
245 250 255

Gln Leu Gly Leu His Pro Trp Leu Pro Ser Ala Met Glu Gly Pro Thr
260 265 270

Pro Val Ser Ala Leu Leu His Ser Ser Thr Met Val Val Ala Gly Ile
275 280 285

Phe Leu Leu Ile Arg Phe His Pro Leu Ala Glu Asn Ser Pro Leu Ile
290 295 300

Gln Thr Leu Thr Leu Cys Leu Gly Ala Ile Thr Thr Leu Phe Ala Ala
305 310 315 320

Val Cys Ala Leu Thr Gln Asn Asp Ile Lys Lys Ile Val Ala Phe Ser
325 330 335

Thr Ser Ser Gln Leu Gly Leu Met Met Val Thr Ile Gly Ile Asn Gln
340 345 350

Pro His Leu Ala Phe Leu His Ile Cys Thr His Ala Phe Phe Lys Ala
355 360 365

Met Leu Phe Met Cys Ser Gly Ser Ile Ile His Asn Leu Asn Asn Glu
370 375 380

Gln Asp Ile Arg Lys Met Gly Gly Leu Leu Lys Thr Met Pro Leu Thr
385 390 395 400

Ser Thr Ser Leu Thr Ile Gly Ser Leu Ala Leu Ala Gly Met Pro Phe
405 410 415

Leu Thr Gly Phe Tyr Ser Lys Asp His Ile Ile Glu Thr Ala Asn Met
420 425 430

Ser Tyr Thr Asn Ala Trp Ala Leu Ser Ile Thr Leu Ile Ala Thr Ser
435 440 445

Leu Thr Ser Ala Tyr Ser Thr Arg Met Ile Leu Leu Thr Leu Thr Gly
450 455 460

Gln Pro Arg Phe Pro Thr Leu Thr Asn Ile Asn Glu Asn Asn Pro Thr
465 470 475 480

Leu Leu Asn Pro Ile Lys Arg Leu Ala Ala Gly Ser Leu Phe Ala Gly
485 490 495

Phe Leu Ile Thr Asn Asn Ile Ser Pro Ala Ser Pro Phe Gln Thr Thr
500 505 510

Ile Pro Leu Tyr Leu Lys Leu Thr Ala Leu Ala Val Thr Phe Leu Gly
515 520 525

Leu Leu Thr Ala Leu Asp Leu Asn Tyr Leu Thr Asn Lys Leu Lys Met
530 535 540

Lys Ser Pro Leu Cys Thr Phe Tyr Phe Ser Asn Met Leu Gly Phe Tyr
545 550 555 560

Pro Ser Ile Thr His Arg Thr Ile Pro Tyr Leu Gly Leu Leu Thr Ser
565 570 575

Gln Asn Leu Pro Leu Leu Leu Leu Asp Leu Thr Trp Leu Glu Lys Leu
580 585 590

Leu Pro Lys Thr Ile Ser Gln His Gln Ile Ser Thr Ser Ile Ile Thr
595 600 605

Ser Thr Gln Lys Gly Met Ile Lys Leu Tyr Phe Leu Ser Phe Phe Phe
610 615 620

Pro Leu Ile Leu Thr Leu Leu Leu Ile Thr
625 630

<210> 28
<211> 542
<212> PRT
<213> Human

<400> 28

Met Asn Phe Ala Leu Ile Leu Met Ile Asn Thr Leu Leu Ala Leu Leu
1 5 10 15

Leu Met Ile Ile Thr Phe Trp Leu Pro Gln Leu Asn Gly Tyr Met Glu
20 25 30

Lys Ser Thr Pro Tyr Glu Cys Gly Phe Asp Pro Met Ser Pro Ala Arg
35 40 45

Val Pro Phe Ser Met Lys Phe Phe Leu Val Ala Ile Thr Phe Leu Leu
50 55 60

Phe Asp Leu Glu Ile Ala Leu Leu Leu Pro Leu Pro Trp Ala Leu Gln
65 70 75 80

Thr Thr Asn Leu Pro Leu Met Val Met Ser Ser Leu Leu Leu Ile Ile
85 90 95

Ile Leu Ala Leu Ser Leu Ala Asn Thr Ala Ala Ile Gln Ala Ile Leu
100 105 110

Tyr Asn Arg Ile Gly Asp Ile Gly Phe Ile Leu Ala Leu Ala Trp Phe
115 120 125

Ile Leu His Ser Asn Ser Trp Asp Pro Gln Gln Met Ala Leu Leu Asn
130 135 140

Ala Asn Pro Ser Leu Thr Pro Leu Leu Gly Leu Leu Leu Ala Ala Ala
145 150 155 160

Gly Lys Ser Ala Gln Leu Gly Leu His Pro Trp Leu Pro Ser Ala Met
165 170 175

Glu Gly Pro Thr Pro Val Ser Ala Leu Leu His Ser Ser Thr Met Val
180 185 190

Val Ala Gly Ile Phe Leu Leu Ile Arg Phe His Pro Leu Ala Glu Asn
195 200 205

Ser Pro Leu Ile Gln Thr Leu Thr Leu Cys Leu Gly Ala Ile Thr Thr
210 215 220

Leu Phe Ala Ala Val Cys Ala Leu Thr Gln Asn Asp Ile Lys Lys Ile

225		230		235		240									
Val	Ala	Phe	Ser	Thr	Ser	Ser	Gln	Leu	Gly	Leu	Met	Met	Val	Thr	Ile
				245					250					255	
Gly	Ile	Asn	Gln	Pro	His	Leu	Ala	Phe	Leu	His	Ile	Cys	Thr	His	Ala
			260					265					270		
Phe	Phe	Lys	Ala	Met	Leu	Phe	Met	Cys	Ser	Gly	Ser	Ile	Ile	His	Asn
		275					280					285			
Leu	Asn	Asn	Glu	Gln	Asp	Ile	Arg	Lys	Met	Gly	Gly	Leu	Leu	Lys	Thr
	290					295					300				
Met	Pro	Leu	Thr	Ser	Thr	Ser	Leu	Thr	Ile	Gly	Ser	Leu	Ala	Leu	Ala
305					310					315					320
Gly	Met	Pro	Phe	Leu	Thr	Gly	Phe	Tyr	Ser	Lys	Asp	His	Ile	Ile	Glu
				325					330					335	
Thr	Ala	Asn	Met	Ser	Tyr	Thr	Asn	Ala	Trp	Ala	Leu	Ser	Ile	Thr	Leu
			340					345					350		
Ile	Ala	Thr	Ser	Leu	Thr	Ser	Ala	Tyr	Ser	Thr	Arg	Met	Ile	Leu	Leu
		355					360					365			
Thr	Leu	Thr	Gly	Gln	Pro	Arg	Phe	Pro	Thr	Leu	Thr	Asn	Ile	Asn	Glu
	370					375						380			
Asn	Asn	Pro	Thr	Leu	Leu	Asn	Pro	Ile	Lys	Arg	Leu	Ala	Ala	Gly	Ser
385					390					395					400
Leu	Phe	Ala	Gly	Phe	Leu	Ile	Thr	Asn	Asn	Ile	Ser	Pro	Ala	Ser	Pro
				405					410					415	
Phe	Gln	Thr	Thr	Ile	Pro	Leu	Tyr	Leu	Lys	Leu	Thr	Ala	Leu	Ala	Val
			420					425					430		
Thr	Phe	Leu	Gly	Leu	Leu	Thr	Ala	Leu	Asp	Leu	Asn	Tyr	Leu	Thr	Asn
		435					440					445			

Lys Leu Lys Met Lys Ser Pro Leu Cys Thr Phe Tyr Phe Ser Asn Met
450 455 460

Leu Gly Phe Tyr Pro Ser Ile Thr His Arg Thr Ile Pro Tyr Leu Gly
465 470 475 480

Leu Leu Thr Ser Gln Asn Leu Pro Leu Leu Leu Asp Leu Thr Trp
485 490 495

Leu Glu Lys Leu Leu Pro Lys Thr Ile Ser Gln His Gln Ile Ser Thr
500 505 510

Ser Ile Ile Thr Ser Thr Gln Lys Gly Met Ile Lys Leu Tyr Phe Leu
515 520 525

Ser Phe Phe Phe Pro Leu Ile Leu Thr Leu Leu Leu Ile Thr
530 535 540

<210> 29
<211> 563
<212> PRT
<213> Human

<400> 29

Met Phe Ala Asp Arg Trp Leu Phe Ser Thr Asn His Lys Asp Ile Gly
1 5 10 15

Thr Leu Tyr Leu Leu Phe Gly Ala Trp Ala Gly Val Leu Gly Thr Ala
20 25 30

Leu Ser Leu Leu Ile Arg Ala Glu Leu Gly Gln Pro Gly Asn Leu Leu
35 40 45

Gly Asn Asp His Ile Tyr Asn Val Ile Val Thr Ala His Ala Phe Val
50 55 60

Met Ile Phe Phe Met Val Met Pro Ile Met Ile Gly Gly Phe Gly Asn
65 70 75 80

Trp Leu Val Pro Leu Met Ile Gly Ala Pro Asp Met Ala Phe Pro Arg
85 90 95

Met Asn Asn Met Ser Phe Trp Leu Leu Pro Pro Ser Leu Leu Leu Leu
100 105 110

Leu Ala Ser Ala Met Val Glu Ala Arg Ala Asp Ala Asn Thr Ala Ala
115 120 125

Ile Gln Ala Ile Leu Tyr Asn Arg Ile Gly Asp Ile Gly Phe Ile Leu
130 135 140

Ala Leu Ala Trp Phe Ile Leu His Ser Asn Ser Trp Asp Pro Gln Gln
145 150 155 160

Met Ala Leu Leu Asn Ala Asn Pro Ser Leu Thr Pro Leu Leu Gly Leu
165 170 175

Leu Leu Ala Ala Ala Gly Lys Ser Ala Gln Leu Gly Leu His Pro Trp
180 185 190

Leu Pro Ser Ala Met Glu Gly Pro Thr Pro Val Ser Ala Leu Leu His
195 200 205

Ser Ser Thr Met Val Val Ala Gly Ile Phe Leu Leu Ile Arg Phe His
210 215 220

Pro Leu Ala Glu Asn Ser Pro Leu Ile Gln Thr Leu Thr Leu Cys Leu
225 230 235 240

Gly Ala Ile Thr Thr Leu Phe Ala Ala Val Cys Ala Leu Thr Gln Asn
245 250 255

Asp Ile Lys Lys Ile Val Ala Phe Ser Thr Ser Ser Gln Leu Gly Leu
260 265 270

Met Met Val Thr Ile Gly Ile Asn Gln Pro His Leu Ala Phe Leu His
275 280 285

Ile Cys Thr His Ala Phe Phe Lys Ala Met Leu Phe Met Cys Ser Gly
290 295 300

Ser Ile Ile His Asn Leu Asn Asn Glu Gln Asp Ile Arg Lys Met Gly
305 310 315 320

Gly Leu Leu Lys Thr Met Pro Leu Thr Ser Thr Ser Leu Thr Ile Gly
325 330 335

Ser Leu Ala Leu Ala Gly Met Pro Phe Leu Thr Gly Phe Tyr Ser Lys
340 345 350

Asp His Ile Ile Glu Thr Ala Asn Met Ser Tyr Thr Asn Ala Trp Ala
355 360 365

Leu Ser Ile Thr Leu Ile Ala Thr Ser Leu Thr Ser Ala Tyr Ser Thr
370 375 380

Arg Met Ile Leu Leu Thr Leu Thr Gly Gln Pro Arg Phe Pro Thr Leu
385 390 395 400

Thr Asn Ile Asn Glu Asn Asn Pro Thr Leu Leu Asn Pro Ile Lys Arg
405 410 415

Leu Ala Ala Gly Ser Leu Phe Ala Gly Phe Leu Ile Thr Asn Asn Ile
420 425 430

Ser Pro Ala Ser Pro Phe Gln Thr Thr Ile Pro Leu Tyr Leu Lys Leu
435 440 445

Thr Ala Leu Ala Val Thr Phe Leu Gly Leu Leu Thr Ala Leu Asp Leu
450 455 460

Asn Tyr Leu Thr Asn Lys Leu Lys Met Lys Ser Pro Leu Cys Thr Phe
465 470 475 480

Tyr Phe Ser Asn Met Leu Gly Phe Tyr Pro Ser Ile Thr His Arg Thr
485 490 495

Ile Pro Tyr Leu Gly Leu Leu Thr Ser Gln Asn Leu Pro Leu Leu Leu
500 505 510

Leu Asp Leu Thr Trp Leu Glu Lys Leu Leu Pro Lys Thr Ile Ser Gln
515 520 525

His Gln Ile Ser Thr Ser Ile Ile Thr Ser Thr Gln Lys Gly Met Ile
530 535 540

Lys Leu Tyr Phe Leu Ser Phe Phe Phe Pro Leu Ile Leu Thr Leu Leu
545 550 555 560

Leu Ile Thr

<210> 30
<211> 136
<212> PRT
<213> Human

<400> 30

Met Ala His Ala Ala Gln Val Gly Leu Gln Asp Ala Thr Ser Pro Ile
1 5 10 15

Met Glu Glu Leu Ile Thr Phe His Asp His Ala Leu Met Ile Ile Phe
20 25 30

Leu Ile Cys Phe Leu Val Leu Tyr Ala Leu Phe Leu Thr Leu Thr Thr
35 40 45

Lys Leu Thr Asn Thr Asn Ile Ser Asp Ala Gln Glu Met Glu Thr Val
50 55 60

Trp Thr Ile Leu Pro Ala Ile Ile Leu Val Leu Ile Ala Leu Pro Ser
65 70 75 80

Leu Arg Ile Leu Tyr Met Thr Asp Glu Val Asn Asp Pro Ser Leu Thr
85 90 95

Ile Lys Ser Ile Gly His Gln Trp Tyr Trp Thr Tyr Glu Tyr Thr Asp
100 105 110

Tyr Gly Gly Leu Ile Phe Asn Ser Tyr Met Leu Pro Pro Leu Phe Leu
115 120 125

Glu Gln Ala Thr Tyr Ser Cys Thr

130

135

<210> 31
 <211> 217
 <212> PRT
 <213> Human

<400> 31

Met Asn Glu Asn Leu Phe Ala Ser Phe Ile Ala Pro Thr Ile Leu Gly
 1 5 10 15

Leu Pro Ala Ala Val Leu Ile Ile Leu Phe Pro Pro Leu Leu Ile Pro
 20 25 30

Thr Ser Lys Tyr Leu Ile Asn Asn Arg Leu Ile Thr Thr Gln Gln Trp
 35 40 45

Leu Ile Lys Leu Thr Ser Lys Gln Met Met Thr Met His Asn Thr Lys
 50 55 60

Gly Arg Thr Trp Ser Leu Met Leu Val Ser Leu Ile Ile Phe Ile Ala
 65 70 75 80

Thr Thr Asn Leu Leu Gly Leu Leu Pro His Ser Phe Thr Pro Thr Thr
 85 90 95

Gln Leu Ser Met Asn Leu Ala Met Ala Ile Pro Leu Trp Ala Gly Thr
 100 105 110

Val Ile Met Gly Phe Arg Ser Lys Ile Lys Asn Ala Leu Ala His Phe
 115 120 125

Leu Pro Gln Gly Thr Pro Thr Pro Leu Ile Pro Met Leu Val Ile Ile
 130 135 140

Glu Thr Ile Ser Leu Leu Ile Gln Pro Met Ala Leu Ala Val Arg Leu
 145 150 155 160

Thr Ala Asn Ile Thr Ala Gly His Leu Leu Met His Leu Ile Gly Ser
 165 170 175

Ala Thr Leu Ala Met Ser Thr Ile Asn Leu Pro Leu Met Val Met Ser
180 185 190

Ser Leu Leu Leu Ile Ile Ile Leu Ala Leu Ser Leu Ala Tyr Glu Trp
195 200 205

Leu Gln Lys Gly Leu Asp Trp Thr Glu
210 215

<210> 32
<211> 484
<212> PRT
<213> Human

<400> 32

Met Asn Glu Asn Leu Phe Ala Ser Phe Ile Ala Pro Thr Ile Leu Gly
1 5 10 15

Leu Pro Ala Ala Val Leu Ile Ile Leu Phe Pro Pro Leu Leu Ile Pro
20 25 30

Thr Ser Lys Tyr Leu Ile Asn Asn Arg Leu Ile Thr Thr Gln Gln Trp
35 40 45

Leu Ile Lys Leu Thr Ser Lys Gln Met Met Thr Met His Asn Thr Lys
50 55 60

Gly Arg Thr Trp Ser Leu Met Leu Val Ser Leu Ile Ile Phe Ile Ala
65 70 75 80

Thr Thr Asn Leu Leu Gly Leu Leu Pro His Ser Phe Thr Pro Thr Thr
85 90 95

Gln Leu Ser Met Asn Leu Ala Met Ala Ile Pro Leu Trp Ala Gly Thr
100 105 110

Val Ile Met Gly Phe Arg Ser Lys Ile Lys Asn Ala Leu Ala His Phe
115 120 125

Leu Pro Gln Gly Thr Pro Thr Pro Leu Ile Pro Met Leu Val Ile Ile
130 135 140

Glu Thr Ile Ser Leu Leu Ile Gln Pro Met Ala Leu Ala Val Arg Leu
145 150 155 160

Thr Ala Asn Ile Thr Ala Gly His Leu Leu Met His Leu Ile Gly Ser
165 170 175

Ala Thr Leu Ala Met Ser Thr Ile Asn Leu His Ala Asn Gly Ala Ser
180 185 190

Met Phe Phe Ile Cys Leu Phe Leu His Ile Gly Arg Gly Leu Tyr Tyr
195 200 205

Gly Ser Phe Leu Tyr Ser Glu Thr Trp Asn Ile Gly Ile Ile Leu Leu
210 215 220

Leu Ala Thr Met Ala Thr Ala Phe Met Gly Tyr Val Leu Pro Trp Gly
225 230 235 240

Gln Met Ser Phe Trp Gly Ala Thr Val Ile Thr Asn Leu Leu Ser Ala
245 250 255

Ile Pro Tyr Ile Gly Thr Asp Leu Val Gln Trp Ile Trp Gly Gly Tyr
260 265 270

Ser Val Asp Ser Pro Thr Leu Thr Arg Phe Phe Thr Phe His Phe Ile
275 280 285

Leu Pro Phe Ile Ile Ala Ala Leu Ala Thr Leu His Leu Leu Phe Leu
290 295 300

His Glu Thr Gly Ser Asn Asn Pro Leu Gly Ile Thr Ser His Ser Asp
305 310 315 320

Lys Ile Thr Phe His Pro Tyr Tyr Thr Ile Lys Asp Ala Leu Gly Leu
325 330 335

Leu Leu Phe Leu Leu Ser Leu Met Thr Leu Thr Leu Phe Ser Pro Asp
340 345 350

Leu Leu Gly Asp Pro Asp Asn Tyr Thr Leu Ala Asn Pro Leu Asn Thr

355

360

365

Pro Pro His Ile Lys Pro Glu Trp Tyr Phe Leu Phe Ala Tyr Thr Ile
 370 375 380

Leu Arg Ser Val Pro Asn Lys Leu Gly Gly Val Leu Ala Leu Leu Leu
 385 390 395 400

Ser Ile Leu Ile Leu Ala Met Ile Pro Ile Leu His Met Ser Lys Gln
 405 410 415

Gln Ser Met Met Phe Arg Pro Leu Ser Gln Ser Leu Tyr Trp Leu Leu
 420 425 430

Ala Ala Asp Leu Leu Ile Leu Thr Trp Ile Gly Gly Gln Pro Val Ser
 435 440 445

Tyr Pro Phe Thr Ile Ile Gly Gln Val Ala Ser Val Leu Tyr Phe Thr
 450 455 460

Thr Ile Leu Ile Leu Met Pro Thr Ile Ser Leu Ile Glu Asn Lys Met
 465 470 475 480

Leu Lys Trp Ala

<210> 33
 <211> 568
 <212> PRT
 <213> Human

<400> 33

Met Phe Ala Asp Arg Trp Leu Phe Ser Thr Asn His Lys Asp Ile Gly
 1 5 10 15

Thr Leu Tyr Leu Leu Phe Gly Ala Trp Ala Gly Val Leu Gly Thr Ala
 20 25 30

Leu Ser Leu Leu Ile Arg Ala Glu Leu Gly Gln Pro Gly Asn Leu Leu
 35 40 45

Gly Asn Asp His Ile Tyr Asn Val Ile Val Thr Ala His Ala Phe Val
50 55 60

Met Ile Phe Phe Met Val Met Pro Ile Met Ile Gly Gly Phe Gly Asn
65 70 75 80

Trp Leu Val Pro Leu Met Ile Gly Ala Pro Asp Met Ala Phe Pro Arg
85 90 95

Met Asn Asn Met Ser Phe Trp Leu Leu Pro Pro Ser Leu Leu Leu Leu
100 105 110

Leu Ala Ser Ala Met Val Glu Ala Gly Ala Gly Thr Gly Trp Thr Val
115 120 125

Tyr Pro Pro Leu Ala Gly Asn Tyr Ser His Pro Gly Ala Ser Val Asp
130 135 140

Leu Thr Ile Phe Ser Leu His Leu Ala Gly Val Ser Ser Ile Leu Gly
145 150 155 160

Ala Ile Asn Phe Ile Thr Thr Ile Ile Asn Met Lys Pro Pro Ala Met
165 170 175

Thr Gln Tyr Gln Thr Pro Leu Phe Val Trp Ser Val Leu Ile Thr Ala
180 185 190

Val Leu Leu Leu Leu Ser Leu Pro Val Leu Ala Ala Gly Ile Thr Met
195 200 205

Leu Leu Thr Asp Arg Asn Leu Asn Thr Thr Phe Phe Asp Pro Ala Gly
210 215 220

Gly Gly Asp Pro Ile Leu Tyr Gln His Leu Phe Trp Phe Phe Gly His
225 230 235 240

Pro Glu Val Tyr Ile Leu Ile Leu Pro Gly Phe Gly Met Ile Ser His
245 250 255

Ile Val Thr Tyr Tyr Ser Gly Lys Lys Glu Pro Phe Gly Tyr Met Gly
260 265 270

Met Val Trp Ala Met Met Ser Ile Gly Phe Leu Gly Phe Ile Val Trp
275 280 285

Ala His His Met Phe Thr Val Gly Met Asp Val Asp Thr Arg Ala Tyr
290 295 300

Phe Thr Ser Ala Thr Met Ile Ile Ala Ile Pro Thr Gly Val Lys Val
305 310 315 320

Phe Ser Trp Leu Ala Thr Leu His Gly Ser Asn Met Lys Trp Ser Ala
325 330 335

Ala Val Leu Trp Ala Leu Gly Phe Ile Phe Leu Phe Thr Val Gly Gly
340 345 350

Leu Thr Gly Ile Val Leu Ala Asn Ser Ser Leu Asp Ile Val Leu His
355 360 365

Asp Thr Tyr Tyr Val Val Ala His Phe His Tyr Val Leu Ser Met Gly
370 375 380

Ala Val Phe Ala Ile Met Gly Gly Phe Ile His Trp Phe Pro Leu Phe
385 390 395 400

Ser Gly Tyr Thr Leu Asp Gln Thr Tyr Ala Lys Ile His Phe Thr Ile
405 410 415

Met Phe Ile Gly Val Asn Leu Thr Phe Phe Pro Gln His Phe Leu Gly
420 425 430

Leu Ser Gly Met Pro Arg Arg Tyr Ser Asp Tyr Pro Asp Ala Tyr Thr
435 440 445

Thr Trp Asn Ile Pro Pro His Ile Lys Pro Glu Trp Tyr Phe Leu Phe
450 455 460

Ala Tyr Thr Ile Leu Arg Ser Val Pro Asn Lys Leu Gly Gly Val Leu
465 470 475 480

Ala Leu Leu Leu Ser Ile Leu Ile Leu Ala Met Ile Pro Ile Leu His
485 490 495

Met Ser Lys Gln Gln Ser Met Met Phe Arg Pro Leu Ser Gln Ser Leu
500 505 510

Tyr Trp Leu Leu Ala Ala Asp Leu Leu Ile Leu Thr Trp Ile Gly Gly
515 520 525

Gln Pro Val Ser Tyr Pro Phe Thr Ile Ile Gly Gln Val Ala Ser Val
530 535 540

Leu Tyr Phe Thr Thr Ile Leu Ile Leu Met Pro Thr Ile Ser Leu Ile
545 550 555 560

Glu Asn Lys Met Leu Lys Trp Ala
565

<210> 34
<211> 191
<212> PRT
<213> Human

<400> 34

Met Pro Gln Leu Asn Thr Thr Val Trp Pro Thr Met Ile Thr Pro Met
1 5 10 15

Leu Leu Thr Leu Phe Leu Thr Thr His Ser Leu Ile Ile Ser Ile Ile
20 25 30

Pro Leu Leu Phe Phe Asn Gln Ile Asn Asn Asn Leu Phe Ser Cys Ser
35 40 45

Pro Thr Phe Ser Ser Asp Pro Leu Thr Thr Pro Leu Leu Met Leu Thr
50 55 60

Thr Trp Leu Leu Pro Leu Thr Ile Met Ala Ser Gln Arg His Leu Ser
65 70 75 80

Ser Glu Pro Leu Ser Arg Lys Lys Leu Tyr Leu Ser Met Leu Ile Ser
85 90 95

Leu Gln Ile Ser Leu Ile Met Thr Phe Thr Ala Thr Glu Leu Ile Met
100 105 110

Phe Tyr Ile Phe Phe Glu Thr Thr Leu Ile Pro Thr Leu Ala Ile Ile
115 120 125

Thr Arg Trp Gly Asn Gln Pro Glu Arg Leu Asn Ala Gly Thr Tyr Phe
130 135 140

Leu Phe Tyr Thr Leu Val Gly Ser Leu Pro Leu Leu Ile Ala Leu Ile
145 150 155 160

Tyr Thr His Asn Thr Leu Gly Ser Leu Asn Ile Leu Leu Leu Thr Leu
165 170 175

Thr Ala Gln Glu Leu Ser Asn Ser Trp Ala Asn Asn Leu Met Asp
180 185 190

<210> 35
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Forward Primer

<400> 35
tctaccccct ctagagccca ctgt 24

<210> 36
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Reverse Primer

<400> 36
ctaggctgcc aatggtgagg gaggt 25

<210> 37
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
 <223> Primer

 <400> 37
 tgcgactcct agccgcagac ctcctcattc 30

<210> 38
 <211> 28
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer

 <400> 38
 ggtaccctaaa tctgcttccc catgaaag 28

<210> 39
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer

 <400> 39
 tgcgactcct agccgcagac 20

<210> 40
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer

 <400> 40
 cgccatcatc ctagtcctca tcgc 24

<210> 41
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer

 <400> 41
 gaatgaggag gtctgcggct aggagtc 27

<210> 42
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 42
gtaagcctct acctacactc caactc

26

<210> 43
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 43
gcggatgagt aagaagattc ctg

23

<210> 44
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 44
ggagacctaa ttgggctgat ttg

23

<210> 45
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 45
ggccgtacgc ctaaccgcta

20

<210> 46
<211> 29
<212> DNA
<213> Artificial Sequence

<220>		
<223>	Primer	
<400>	46	
gttgtgggtc tcatgagttg gagtgtagg		29
<210>	47	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	47	
ccctggccgt acgcctaacc		20
<210>	48	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	48	
atttggtgtg ggtctcatga gttggagtgt agg		33
<210>	49	
<211>	25	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	49	
ccctaagtct ggccaacaca gcagc		25
<210>	50	
<211>	22	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	50	
gggtggagac ctaattgggc tg		22

<210>	51	
<211>	25	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	51	
	acaacgttat cgtcacagcc catgc	25
<210>	52	
<211>	27	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	52	
	gattgcttga atggctgctg tgttggc	27
<210>	53	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	53	
	cgtctgaact atcctgccccg c	21
<210>	54	
<211>	27	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	54	
	caattaggtg catgagtagg tggcctg	27
<210>	55	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	

<220>		
<223>	Primer	
<400>	55	
aaggcacacc	tacacccctt	20
<210>	56	
<211>	28	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	56	
gagggatgac	ataactatta gtggcagg	28
<210>	57	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	57	
aaccaatagc	cctggccgta	20
<210>	58	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	58	
gagggatgac	ataactatta gtggcaggtt a	31
<210>	59	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	59	
ctatagcacc	ccctctaccc c	21

<210> 60
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 60
gatgctaata attaggctgt gggtggt

27

<210> 61
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 61
acagtgaaat gcccctaacta

20

<210> 62
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 62
gctcaggcgt ttgtgtatga

20

<210> 63
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 63
caacgatccc tcccttacca

20

<210> 64
<211> 21
<212> DNA
<213> Artificial Sequence

<220>		
<223>	Primer	
<400>	64	
	agtacggatg ctacttgtcc a	21
<210>	65	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	65	
	gaagcgccac cctagcaata	20
<210>	66	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	66	
	ggtgaggctt ggattagcgt	20
<210>	67	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	67	
	aatccacccc ttacgagtgc	20
<210>	68	
<211>	21	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Primer	
<400>	68	
	acaacgttat cgtcacagcc c	21

<210> 69
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 69
gtgaggcttg gattagcggtt 20

<210> 70
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 70
ctgaacctac gagtacaccg a 21

<210> 71
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 71
gtgtgaaaac gtaggcttgg a 21

<210> 72
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 72
tcgaaaccat cagcctactc a 21

<210> 73
<211> 23
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer

<400> 73

ccaattcggg tcagtcta cct

23

<210> 74

<211> 773

<212> DNA

<213> Human

<400> 74

atgaacgaaa atctgttcgc ttcattcatt gccccacaa tcctaggcct acccgccgca 60

gtactgatca ttctatttcc cctctattg atccccacct ccaaatactt catcaacaac 120

cgactaatca ccaccaaca atgactatca aactaacctc aaaacaaatg ataaccatac 180

acaacactaa aggacgaacc tgatctctta tactagtatc cttaatcatt tttattgcca 240

caactaacct cctcggactc ctgcctcact catttacacc aaccaccaa ctatctataa 300

acctagccat ggccatcccc ttatgagcgg gcacagtgat tataggcttt cgctctaaga 360

ttaaaaatgc cctagcccac ttcttaccac aaggcacacc tacaccctt atccccatac 420

tagttattat cgaaaccatc agcctactca ttcaaccaac agccctagac ctcaactacc 480

taaccaacaa acttaaaata aaatccccac tatgcacatt ttatttctcc aacatactcg 540

gattctaccc tagcatcaca caccgcacaa tcccctatct aggccttctt acgagccaaa 600

acctgcccct actcctccta gacctaacct gactagaaaa gctattacct aaaacaattt 660

cacagcacca aatctccacc tccatcatca cctcaaccca aaaaggcata attaaacttt 720

acttcctctc tttcttcttc ccactcatcc taaccctact cctaatacaca taa 773

<210> 75

<211> 993

<212> DNA

<213> Human

<400> 75

attaatcccc tggcccaacc cgtcatctac tctaccatct ttgcaggcac actcatcaca 60

gcgctaagct cgcactgatt ttttacctga gtaggcctag aaataaacat gctagctttt 120

attccagttc taaccaaaaa aataaacctt cgttccacag aagctgccat caagtatttc 180

ctcacgcaag caaccgatc cataatcctt ctaatagcta tcctcttcaa caatatactc 240

tccggacaat gaaccataac caatactacc aatcaatact catcattaat aatcataata	300
gctatagcaa taaaactagg aatagccccc tttcacttct gagtcccaga ggttacccaa	360
ggcaccctc tgacatccgg cctgcttctt ctacatgac aaaaactagc ccccatctca	420
atcatatacc aaatctctcc ctactaaac gtaagccttc tcctcactct ctcaatctta	480
tccatcatag caggcagttg aggtggatta aaccaaacc agctacgcaa aatcttagca	540
tactcctcaa ttaccacat aggatgaata atagcagttc taccgtacaa ccctaacata	600
accattctta atttaactat ttatattatc ctaactacta ccgcattcct actactcaac	660
ttaaactcca gcaccacgac cctactacta tctcgcacct gaaacaagct aacatgacta	720
acacccttaa ttccatccac cctcctctcc ctaggaggcc tgccccgct aaccggcttt	780
ttgccc aaat gggccattat cgaagaattc acaaaaaaca atagcctcat catccccacc	840
atcatagcca ccatcacct ccttaacctc tacttctacc tacgcctaata ctactccacc	900
tccatcatca cctcaacc caaaaggcata attaaacttt acttcctctc tttcttcttc	960
ccactcatcc taaccctact cctaatacaca taa	993

<210> 76
 <211> 773
 <212> RNA
 <213> Human

<400> 76	
augaacgaaa aucuguucgc uucauucauu gccccacaa uccuaggccu acccgccgca	60
guacugauca uucuaauucc ccucuaauug auccccaccu ccaaauaucu caucaacaac	120
cgacuaauca ccaccaaca augacuauca aacuaaccuc aaaacaaaug auaaccuac	180
acaacacuaa aggacgaacc ugaucucuua uacuaguauc cuuaaucauu uuuaugcca	240
caacuaaccu ccucggacuc cugccucacu cauuuacacc aaccacccaa cuaucuauaa	300
accuagccau ggccaucucc uuaugagcgg gcacagugau uauaggcuuu cgcucuaaga	360
uuaaaaaugc ccuagcccac uucuuaccac aaggcacacc uacaccccuu auccccaauac	420
uaguuaauau cgaaaccauc agccuacuca uucaaccaac agcccuagac cucaacuacc	480
uaaccaacaa acuuaaaaua aaaucccac uaugcacauu uuauuucucc aacauacucg	540
gauucuaccc uagcaucaca caccgcacaa ucccuaucu aggccuucuu acgagccaaa	600
accugcccuu acuccuccua gaccuaaccu gacuagaaaa gcuauuaccu aaaacaauuu	660

cacagcacca aaucuccacc uccaucauca ccucaaccca aaaaggcaua auuaaacuuu	720
acuuccucuc uuucuucuuc ccacucaucc uaaccuacu ccuaaucaca uaa	773

<210> 77
 <211> 993
 <212> RNA
 <213> Human

<400> 77	
auuaaucccc uggccaacc cgucaucuac ucuaccaucu uugcaggcac acucaucaca	60
gcgcuaagcu cgcacugauu uuuuaccuga guaggccuag aaauaaacau gcuagcuuuu	120
auuccaguuc uaaccaaaaa aaauaacccu cguuccacag aagcugccau caaguauuuc	180
cucacgcaag caaccgcauc cauaauccuu cuaauagcua uccucuuaa caauauacuc	240
uccggacaau gaaccauaac caauacuacc aaucaauacu caucauuau aaucauaaua	300
gcuauagcaa uaaaacuagg aaugccccc uuucacuucu gagucccaga gguuacccaa	360
ggcaccucc ucacauccgg ccugcuucuu cucacaugac aaaaacuagc ccccaucuca	420
aucauauacc aaucucucc cucacuaaac guaagccuuc uccucacucu cucaaucuua	480
uccaucauag caggcaguug agguggauua aaccaaacc agcuacgcaa aaucuuagca	540
uacuccuca uuaaccacau aggaugaaua auagcaguuc uaccguacaa cccuaacaua	600
accauucua auuaacuau uuauuuuau cuaacuacua ccgcauuccu acuacucaac	660
uuaaacucca gcaccacgac ccuacuacua ucucgcaccu gaaacaagcu aacaugacua	720
acaccuuua uuccauccac ccuccucucc cuaggaggcc ugcccccgcu aaccggcuuu	780
uugcccaaau gggccauuau cgaagaauuc aaaaaaaca auagccucau cauccccacc	840
aucauagcca ccaucacccu ccuaaaccuc uacuucuacc uacgccuau cuacuccacc	900
uccaucauca ccucaaccca aaaaggcaua auuaaacuuu acuuccucuc uuucuucuuc	960
ccacucaucc uaaccuacu ccuaaucaca uaa	993

<210> 78
 <211> 222
 <212> PRT
 <213> Human

<400> 78

Met Ser His Gln Gln Pro Thr Asn His His Pro Thr Met Thr Ile Lys
1 5 10 15

Leu Thr Ser Lys Gln Met Met Thr Met His Asn Thr Lys Gly Arg Thr
20 25 30

Trp Ser Leu Met Leu Val Ser Leu Ile Ile Phe Ile Ala Thr Thr Asn
35 40 45

Leu Leu Gly Leu Leu Pro His Ser Phe Thr Pro Thr Thr Gln Leu Ser
50 55 60

Met Asn Leu Ala Met Ala Ile Pro Leu Trp Ala Gly Thr Val Ile Met
65 70 75 80

Gly Phe Arg Ser Lys Ile Lys Asn Ala Leu Ala His Phe Leu Pro Gln
85 90 95

Gly Thr Pro Thr Pro Leu Ile Pro Met Leu Val Ile Ile Glu Thr Ile
100 105 110

Ser Leu Leu Ile Gln Pro Thr Ala Leu Asp Leu Asn Tyr Leu Thr Asn
115 120 125

Lys Leu Lys Met Lys Ser Pro Leu Cys Thr Phe Tyr Phe Ser Asn Met
130 135 140

Leu Gly Phe Tyr Pro Ser Ile Thr His Arg Thr Ile Pro Tyr Leu Gly
145 150 155 160

Leu Leu Thr Ser Gln Asn Leu Pro Leu Leu Leu Asp Leu Thr Trp
165 170 175

Leu Glu Lys Leu Leu Pro Lys Thr Ile Ser Gln His Gln Ile Ser Thr
180 185 190

Ser Ile Ile Thr Ser Thr Gln Lys Gly Met Ile Lys Leu Tyr Phe Leu
195 200 205

Ser Phe Phe Phe Pro Leu Ile Leu Thr Leu Leu Leu Ile Thr
210 215 220

<210> 79
<211> 296
<212> PRT
<213> Human

<400> 79

Met Asn Met Leu Ala Phe Ile Pro Val Leu Thr Lys Lys Met Asn Pro
1 5 10 15

Arg Ser Thr Glu Ala Ala Ile Lys Tyr Phe Leu Thr Gln Ala Thr Ala
20 25 30

Ser Met Ile Leu Leu Met Ala Ile Leu Phe Asn Asn Met Leu Ser Gly
35 40 45

Gln Trp Thr Met Thr Asn Thr Thr Asn Gln Tyr Ser Ser Leu Met Ile
50 55 60

Met Met Ala Met Ala Met Lys Leu Gly Met Ala Pro Phe His Phe Trp
65 70 75 80

Val Pro Glu Val Thr Gln Gly Thr Pro Leu Thr Ser Gly Leu Leu Leu
85 90 95

Leu Thr Trp Gln Lys Leu Ala Pro Ile Ser Ile Met Tyr Gln Ile Ser
100 105 110

Pro Ser Leu Asn Val Ser Leu Leu Leu Thr Leu Ser Ile Leu Ser Ile
115 120 125

Met Ala Gly Ser Trp Gly Gly Leu Asn Gln Thr Gln Leu Arg Lys Ile
130 135 140

Leu Ala Tyr Ser Ser Ile Thr His Met Gly Trp Met Met Ala Val Leu
145 150 155 160

Pro Tyr Asn Pro Asn Met Thr Ile Leu Asn Leu Thr Ile Tyr Ile Ile
165 170 175

Leu Thr Thr Thr Ala Phe Leu Leu Leu Asn Leu Asn Ser Ser Thr Thr

180

185

190

Thr Leu Leu Leu Ser Arg Thr Trp Asn Lys Leu Thr Trp Leu Thr Pro
 195 200 205

Leu Ile Pro Ser Thr Leu Leu Ser Leu Gly Gly Leu Pro Pro Leu Thr
 210 215 220

Gly Phe Leu Pro Lys Trp Ala Ile Ile Glu Glu Phe Thr Lys Asn Asn
 225 230 235 240

Ser Leu Ile Ile Pro Thr Ile Met Ala Thr Ile Thr Leu Leu Asn Leu
 245 250 255

Tyr Phe Tyr Leu Arg Leu Ile Tyr Ser Thr Ser Ile Ile Thr Ser Thr
 260 265 270

Gln Lys Gly Met Ile Lys Leu Tyr Phe Leu Ser Phe Phe Phe Pro Leu
 275 280 285

Ile Leu Thr Leu Leu Leu Ile Thr
 290 295

<210> 80
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Forward Primer

<400> 80
 tgccctagcc cacttcttac c

21

<210> 81
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Reverse (junction) Primer

<400> 81
 tagttgaggt ctagggctgt tggtt

25

<210> 82
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Forward Primer

<400> 82
gggccattat cgaagaattc acaaa 25

<210> 83
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Reverse (junction) Primer

<400> 83
gaggtgatga tggaggtgga gtag 24

<210> 84
<211> 259
<212> PRT
<213> Human

<400> 84

Met Ala Phe Met Val Lys Met Pro Leu Tyr Gly Leu His Leu Trp Leu
1 5 10 15

Pro Lys Ala His Val Glu Ala Pro Ile Ala Gly Ser Met Val Leu Ala
20 25 30

Ala Val Leu Leu Lys Leu Gly Gly Tyr Gly Met Met Arg Leu Thr Leu
35 40 45

Ile Leu Asn Pro Leu Thr Lys His Met Ala Tyr Pro Phe Leu Val Leu
50 55 60

Ser Leu Trp Gly Met Ile Met Thr Ser Ser Ile Cys Leu Arg Gln Thr
65 70 75 80

Asp Leu Lys Ser Leu Ile Ala Tyr Ser Ser Ile Ser His Met Ala Leu
85 90 95

Val Val Thr Ala Ile Leu Ile Gln Thr Pro Trp Ser Phe Thr Gly Ala
100 105 110

Val Ile Leu Met Ile Ala His Gly Leu Thr Ser Ser Leu Leu Phe Cys
115 120 125

Leu Ala Asn Ser Asn Tyr Glu Arg Thr His Ser Arg Ile Met Ile Leu
130 135 140

Ser Gln Gly Leu Gln Thr Leu Leu Pro Leu Met Ala Phe Trp Trp Leu
145 150 155 160

Leu Ala Ser Leu Ala Asn Leu Ala Leu Pro Pro Thr Ile Asn Leu Leu
165 170 175

Gly Glu Leu Ser Val Leu Val Thr Thr Phe Ser Trp Ser Asn Ile Thr
180 185 190

Leu Leu Leu Thr Gly Leu Asn Met Leu Val Thr Ala Leu Tyr Ser Leu
195 200 205

Tyr Met Phe Thr Thr Thr Gln Trp Gly Ser Leu Thr His His Ile Asn
210 215 220

Asn Met Lys Pro Ser Phe Thr Arg Glu Asn Thr Leu Met Phe Met His
225 230 235 240

Leu Ser Pro Ile Leu Leu Leu Ser Leu Asn Pro Asp Ile Ile Thr Gly
245 250 255

Phe Ser Ser