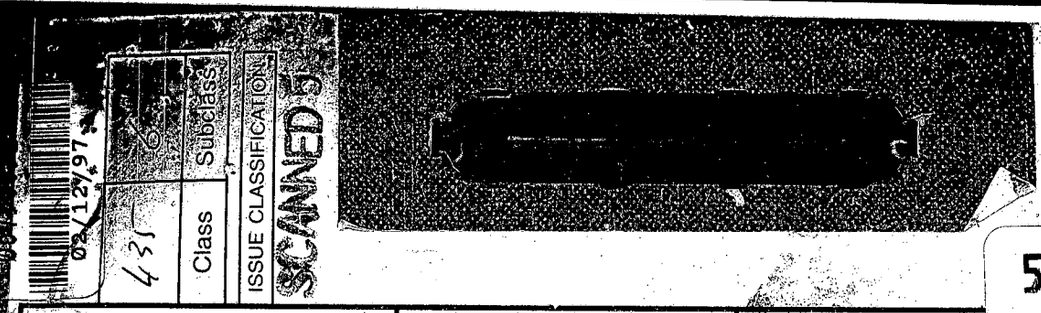


5750400

UTILITY SERIAL NUMBER	PATENT DATE	PATENT NUMBER
	MAY 12 1998	5750400

SERIAL NUMBER	FILING DATE	CLASS	SUBCLASS	GROUP ART UNIT	EXAMINER
08/798,691	02/12/97	435	6	1642	~~DAVIS~~ Huff

APPLICANTS

PATRICIA D. MURPHY, SLINGERLANDS, NY; ANTONETTE C. ALLEN, SEVERN, MD; CHRISTOPHER P. ALVARES, POTOMAC, MD; BRENDA S. CRITZ, FREDERICK, MD; SHERI J. OLSON, ARLINGTON, VA; DENISE B. SCHELTER, SILVER SPRING, MD; BIN ZENG, ROCKVILLE, MD.

CONTINUING DATA********
VERIFIED THIS APPLN IS A CIP OF 08/598,591 02/12/96 PAT 5,654,155

Davis

Certificate
FEB 17 2004
of Correction

NOTE-DISCLAIMER
The term of this patent shall not extend beyond the expiration date of Pat. No. 5,654,155.

FOREIGN APPLICATIONS********
VERIFIED

Davis

Certificate
FEB 1
of Corr

***** SMALL ENTITY *****

Foreign priority claimed / 35 USC 119 conditions met	AS FILED	STATE OR COUNTRY	SHEETS DRWGS.	TOTAL CLAIMS	INDEP. CLAIMS	FILING FEE RECEIVED	ATTORNEY'S DOCKET NO.
☐ yes ☒ no / ☐ yes ☒ no — Verified and Acknowledged Davis Examiner's initials	→	NY	1	27	13	$927.00	PA-0054CIP

ADDRESS

ALBERT P. HALLUIN
HOWERY & SIMON
1299 PENNSYLVANIA AVE., N.W.
WASHINGTON DC 20004

TITLE

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

U.S. DEPT. OF COMM./ PAT. & TM—PTO-436L (Rev.12-94)

PARTS OF APPLICATION FILED SEPARATELY

Applications Examiner

NOTICE OF ALLOWANCE MAILED		CLAIMS ALLOWED	
Feb. 17, 1998	Assistant Examiner ~~DAVIS~~	Total Claims: 8	Print Claim: 1

ISSUE FEE		DRAWING		
Amount Due: $660.00	Date Paid:	Sheets Drwg.: 1	Figs. Drwg.: 1	Print Fig.: NONE

Sheela Huff
SHEELA HUFF
PRIMARY EXAMINER
Primary Examiner

ISSUE BATCH NUMBER: A-52

PREPARED FOR ISSUE

FORM PTO-1267 (Rev. 11-92) U.S. Department of Commerce Patent and Trademark Office

DISCLAIMER LABEL

Application No. 08-798691

A terminal disclaimer has been entered and recorded under 35 U.S.C. 253 in this file.

DO NOT DESTROY

Form PTO-436A (Rev.8/92) SCAN 5 BW QC KC TP

ISSUE FEE IN FILE

61851 U.S. PTO
08798691
02/12/97

PATENT APPLICATION

08798691

APPROVED FOR LICENSE ☐
INITIALS MAR 1 79729

Date Entered or Counted — Briefed DAE

CONTENTS

Date Received or Mailed

1107

	1.	Application ____ papers.	
	2.	Re: LTR Signature	3/28/97
	3.	Pre Amdt A w/ disk	08-14-97
	4.	Raw Sequence Listing (OK)	9-24-97
	5.	Opsot P/A	9-30-97
	6.	Ex's Interview Summary	10-17-97
	7.	Info Disclosure Stmt	10-17-97
	8.	Terminal Drsl.	10-17-97
11-28-97	9.	Supp IDS	10-30-97
	10.	Raw Sequence Listing (OK)	9/8/97
	11.	pre amdt B	10/22/97
	12.	Supp. IDS	12/9/97
	13.	pre amdt C	12/9/97
	14.	Interview Summary	12/4/97
	15.	Examiner's Amdt	02-17-98
4/15/98 ~~2/19/98~~	16.	Issue Fee PTO Grant MAY 1 2 1998	3-4-98
	17.		
	18.	Reg Ltr	10/6/03
	19.	Director's Report	12/5/03
	20.		
	21.		
	22.		
	23.		
	24.		
	25.		
	26.		
	27.		
	28.		
	29.		
	30.		
	31.		
	32.		

Staple Issue ~~Slip~~ Here

POSITION		ID NO.	DATE
CLASSIFIER		21	3/17/97
EXAMINER		872	3/23
TYPIST			
VERIFIER			
CORPS CORR.			
SPEC. HAND			
FILE MAINT.		413	3/27/97
DRAFTING			

INDEX OF CLAIMS

Claim Final	Claim Original	Date
	1	
	2	
	3	
	4	
	5	
	6	
	7	
	8	
	9	
	10	
	11	
	12	
	13	
	14	
	15	
2	16	=
3	17	=
4	18	=
5	19	=
6	20	=
7	21	=
	22	
	23	
8	24	=
	25	
	26	
	27	
	28	
	29	
	30	
	31	
	32	
	33	
	34	
	35	
	36	
	37	
	38	
	39	
	40	
	41	
	42	
	43	
	44	
	45	
	46	
	47	
	48	
	49	
	50	

SYMBOLS

- ✓ ..... Rejected
- = ... Allowed
- - (Through numberal) Canceled
- + ... Restricted
- N ..... Non-elected
- I ..... Interference
- A ..... Appeal
- O ..... Objected

Claim Final	Claim Original	Date
	51	
	52	
	53	
	54	
	55	
	56	
	57	
	58	
	59	
	60	
	61	
	62	
	63	
	64	
	65	
	66	
	67	
	68	
	69	
	70	
	71	
	72	
	73	
	74	
	75	
	76	
	77	
	78	
	79	
	80	
	81	
	82	
	83	
	84	
	85	
	86	
	87	
	88	
	89	
	90	
	91	
	92	
	93	
	94	
	95	
	96	
	97	
	98	
	99	
	100	

[redacted]

SEARCHED

Class	Sub.	Date	Exmr.
435	6	12/20/97	TDais
	91.2	↓	↓
536	23.1		
	24.3		
	24.33		
435	6		
	91.1, 91.2		
	91.21		
536	23.1		
	23.5		
	24.3		
	24.33	2/10/98	SSH

SEARCH NOTES

	Date	Exmr.
MPSRCH, APS, Dialog	12/20/97	TDais
08/825487		
APS-RECA	2/10/98	SSH
Margaret Parrand Lila Fusco said it is allowable [redacted]	2/4/98	SSH

INTERFERENCE SEARCHED

Class	Sub.	Date	Exmr.
435	6	12/20/97	TDais
	91.2	↓	↓
536	23.1		
	24.3		
	24.33		
MPSRCH			
updated above			
435	91.1, 91.21		
536	23.5	2/10/98	SSH
[redacted]			

61851 U.S. PTO
08798691
02/12/97

PATENT APPLICATION SERIAL NO. _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

010 15-0609 03/10/97 08798691
01035 101 862.00CH 0054CIP

TO-1556
(5/87)

NEW PATENT APPLICATION TRANSMITTAL

Docket No. : PA-0054 CIP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner of Patents and Trademarks
Box Patent Application
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the application of

Inventor(s): PATRICIA D. MURPHY, ANTONETTE C. ALLEN, CHRISTOPHER P. ALVARES, BRENDA S. CRITZ, SHERI J. OLSON, DENISE B. SCHELTER, AND BIN ZENG

Title: CODING SEQUENCES OF THE HUMAN BRCA1 GENE

Enclosed are also:

_____ An assignment of the invention to ONCORMED, INC., along with a Recordation form Cover Sheet.
_____ An Information Disclosure Statement under 37 CFR 1.97.
__X__ Declaration, Petition and Power of Attorney.
__X__ Small Entity Status Declaration.

The filing fee is calculated as follows:

Basic fee	$385*
Total claims: 27-20= 7 X $11*=	$ 77
Independent claims:13-3 = 10 X $40*=	$400
Multiple dependent claims ($130* if any)	$
TOTAL: *Applicant has small entity status	$862

The Commissioner is hereby authorized to charge the filing fee, and any other fees which may be required, or credit any overpayment to Deposit Account No. 15-0609. This sheet is enclosed in duplicate.

Respectfully submitted,

R Thomas Gallegos

R. THOMAS GALLEGOS
Reg. No. 32,692
Attorney for Applicants
Codon Pharmaceuticals, Inc.
200 Perry Parkway
Gaithersburg, MD 20877
(301) 527-2051

Date: FEBRUARY 12, 1997

1xduplicate

A

PA-0054CIP

This Application is a Continuation-In Part of U.S.Application Serial No. 08/598,591 filed on February 12, 1996, now patented US 5,654,155.

TDavis 12/3/97

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

FIELD OF THE INVENTION

This invention relates to a gene which has been associated with breast and ovarian cancer where the gene is found to be mutated. More specifically, this invention relates to the three coding sequences of the BRCA1 gene BRCA1$^{(omi1)}$, BRCA1$^{(omi2)}$, and BRCA1$^{(omi3)}$) isolated from human subjects.

BACKGROUND OF THE INVENTION

It has been estimated that about 5-10% of breast cancer is inherited Rowell, S., *et al.*, *American Journal of Human Genetics* 55:861-865 (1994). Located on chromosome 17, BRCA1 is the first gene identified to be conferring increased risk for breast and ovarian cancer. Miki *et al.*, *Science* 266:66-71 (1994). Mutations in this "tumor suppressor" gene are thought to account for roughly 45% of inherited breast cancer and 80-90% of families with increased risk of early onset breast and ovarian cancer. Easton *et al.*, *American Journal of Human Genetics* 52:678-701 (1993).

Locating one or more mutations in the BRCA1 region of chromosome 17 provides a promising approach to reducing the high incidence and mortality associated with breast and ovarian cancer through the early detection of women at high risk. These women, once identified, can be targeted for more aggressive prevention programs. Screening is carried out by a variety of methods which include karyotyping, probe binding and DNA sequencing.

In DNA sequencing technology, genomic DNA is extracted from whole blood and the coding sequences of the BRCA1 gene are amplified. The coding sequences might be sequenced completely and the results are compared to the DNA sequence of the gene. Alternatively, the coding sequence of the sample gene may be compared to a panel of known mutations before completely sequencing the gene and comparing it to a normal sequence of the gene.

If a mutation in the BRCA1 coding sequence is found, it may be possible to provide the individual with increased expression of the gene through gene transfer therapy. It has been demonstrated that the gene transfer of the BRCA1 coding sequence into cancer cells inhibits their growth and reduces tumorigenesis of human cancer cells in nude mice. Jeffrey Holt and his colleagues conclude that the product of BRCA1 expression is a secreted tumor growth inhibitor, making BRCA1 an ideal gene for gene therapy studies. Transduction of only a moderate percentage of tumor cells apparently produces enough growth inhibitor to inhibit all tumor cells. Arteaga, CL, and JT Holt Cancer Research 56: 1098-1103 (1996), Holt, JT et al., Nature Genetics 12: 298-302 (1996).

The observation of Holt et al, that the BRCA1 growth inhibitor is a secreted protein leads to the possible use of injection of the growth inhibitor into the area of the tumor for tumor suppression.

The BRCA1 gene is divided into 24 separate exons. Exons 1 and 4 are noncoding, in that they are not part of the final functional BRCA1 protein product. The BRCA1 coding sequence spans roughly 5600 base pairs (bp). Each exon consists of 200-400 bp, except for exon 11 which contains about 3600 bp. To sequence the coding sequence of the BRCA1 gene, each exon is amplified separately and the resulting PCR products are sequenced in the forward and reverse directions. Because exon 11 is so large, we have divided it into twelve overlapping PCR fragments of roughly 350 bp each (segments "A" through "L" of BRCA1 exon 11).

Many mutations and polymorphisms have already been reported in the BRCA1 gene. A world wide web site has been built to facilitate the detection and characterization of alterations in breast cancer susceptibility genes. Such mutations in BRCA1 can be accessed through the Breast Cancer Information Core at: http://www.nchgr.nih.gov/dir/lab_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995).

The genetics of Breast/Ovarian Cancer Syndrome is autosomal dominant with reduced penetrance. In simple terms, this means that the syndrome runs through families such that both sexes can be carriers (only women get the disease but men can pass it on), all generations will likely have breast/ovarian or both diseases and sometimes in the same individual, occasionally women carriers either die young before they have the time to manifest disease (and yet offspring get it) or they never develop breast or ovarian cancer and die of old age (the latter people are said to have "reduced penetrance" because they never develop cancer). Pedigree analysis and genetic counseling is absolutely essential to the proper workup of a family prior to any lab work.

Until now, only a single coding sequence for the BRCA1 gene has been available for comparison to patient samples. That sequence is available as GenBank Accession Number U14680. There is a need in the art, therefore, to have available a coding sequence which is the BRCA1 coding sequence found in the majority of the population, a "consensus coding sequence", BRCA1$^{(omi1)}$ Seq. ID. NO. 1. A consensus coding sequence will make it possible for true mutations to be easily identified or differentiated from polymorphisms. Identification of mutations of the BRCA1 gene and protein would allow more widespread diagnostic screening for hereditary breast and ovarian cancer than is currently possible. Two additional coding sequences have been isolated and characterize. The BRCA1$^{(omi2)}$ SEQ. ID. NO.: 3, and BRCA1$^{(omi3)}$ SEQ. ID. NO.:5 coding sequences also have utility in diagnosis, gene therapy and in making therapeutic BRCA1 protein.

A coding sequence of the BRCA1 gene which occurs most commonly in the human gene pool is provided. The most commonly occurring coding sequence more accurately reflects the most likely sequence to be found in a subject. Use of the coding sequence BRCA1$^{(omi1)}$ SEQ. ID. NO.: 1, rather than the previously published BRCA1 sequence, will reduce the likelihood of misinterpreting a "sequence variation" found in the population (i.e. polymorphism) with a pathologic "mutation" (i.e. causes disease in the individual or puts the individual at a high risk of developing the disease). With large interest in breast cancer predisposition testing, misinterpretation is

particularly worrisome. People who already have breast cancer are asking the clinical question: "is my disease caused by a heritable genetic mutation?" The relatives of the those with breast cancer are asking the question: "Am I also a carrier of the mutation my relative has? Thus, is my risk increased, and should I undergo a more aggressive surveillance program."

SUMMARY OF THE INVENTION

The present invention is based on the isolation of three coding sequences of the BRCA1 gene found in human individuals.

It is an object of the invention to provide the most commonly occurring coding sequence of the BRCA1 gene.

It is another object of this invention to provide two other coding sequences of BRCA1 gene.

It is another object of the invention to provide three protein sequences coded for by three of the coding sequences of the BRCA1 gene.

It is another object of the invention to provide a list of the codon pairs which occur at each of seven polymorphic points on the BRCA1 gene.

It is another object of the invention to provide the rates of occurrence for the codons.

It is another object of the invention to provide a method wherein BRCA1, or parts thereof, is amplified with one or more oligonucleotide primers.

It is another object of this invention to provide a method of identifying individuals who carry no mutation(s) of the BRCA1 coding sequence and therefore have no increased genetic susceptibility to breast or ovarian cancer based on their BRCA1 genes.

It is another object of this invention to provide a method of identifying a mutation leading to an increased genetic susceptibility to breast or ovarian cancer.

There is a need in the art for a sequence of the BRCA1 gene and for the protein sequence of BRCA1 as well as for an accurate list of codons which occur at polymorphic points on a sequence.

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

a) identifying individuals having a BRCA1 gene with no coding mutations, who therefore cannot be said to have an increased genetic susceptibility to breast or ovarian cancer from their BRCA1 genes;

b) avoiding misinterpretation of polymorphisms found in the BRCA1 gene;

c) determining the presence of a previously unknown mutation in the BRCA1 gene.

d) identifying a mutation which increases the genetic susceptibility t o breast or ovarian cancer.

e) probing a human sample of the BRCA1 gene.

f) performing gene therapy.

g) for making a functioning tumor growth inhibitor protein coded for by one of the BRCA1^{omi} genes.

BRIEF DESCRIPTION OF THE FIGURE

As shown in FIGURE 1, the alternative alleles at polymorphic (non-mutation causing variations) sites along a chromosome can be represented as a "haplotype" within a gene such as BRCA1. The BRCA1$^{(omi1)}$ haplotype is shown in Figure 1 with dark shading (encompassing the alternative alleles found at nucleotide sites 2201, 2430, 2731, 3232, 3667, 4427, and 4956). For comparison, the haplotype that is in GenBank is shown with no shading. As can be seen from the figure, the common "consensus" haplotype is found intact in five separate chromosomes labeled with the OMI symbol (numbers 1-5 from left to right). Two additional haplotypes (BRCA1$^{(omi2)}$, and BRCA1$^{(omi3)}$ are represented with mixed dark and light shading (numbers 7 and 9 from left to right). In total, 7 of 10 haplotypes along the BRCA1 gene are unique.

DETAILED DESCRIPTION OF THE INVENTION

DEFINITIONS

The following definitions are provided for the purpose of understanding this invention.

"Breast and Ovarian cancer" is understood by those skilled in the art to include breast and ovarian cancer in women and also breast and prostate cancer in men. BRCA1 is associated genetic susceptibility to inherited breast and ovarian cancer in women and also breast and prostate cancer in men. Therefore, claims in this document which recite breast and/or ovarian cancer refer to breast, ovarian and prostate cancers in men and women.

" Coding sequence" or " DNA coding sequence"refers to those portions of a gene which, taken together, code for a peptide (protein), or which nucleic acid itself has function.

" Protein" or "peptide" refers to a sequence amino acids which has function.

"BRCA1$^{(omi)}$" refers collectively to the "BRCA1$^{(omi1)}$", "BRCA1$^{(omi2)}$" and "BRCA1$^{(omi3)}$" coding sequences.

"BRCA1$^{(omi1)}$" refers to SEQ. ID. NO.: 1, a coding sequence for the BRCA1 gene. The coding sequence was found by end to end sequencing of BRCA1 alleles from individuals randomly drawn from a Caucasian population found to have no family history of breast or ovarian cancer. The sequenced gene was found not to contain any mutations. BRCA1$^{(omi1)}$ was determined to be a consensus sequence by calculating the frequency with which the coding sequence occurred among the sample alleles sequenced.

"BRCA1$^{(omi2)}$" and "BRCA1$^{(omi3)}$" refer to SEQ. ID. NO.: 3, and SEQ. ID. NO.: 5 respectively. They are two additional coding sequences for the BRCA1 gene

which were also isolated from individuals randomly drawn from a Caucasian population found to have no family history of breast or ovarian cancer.

~~polymorphisms.~~

"Primer" as used herein refers to a sequence comprising about 20 or more nucleotides of the BRCA1 gene.

"Genetic susceptibility" refers to the susceptibility to breast or ovarian cancer due to the presence of a mutation in the BRCA1 gene.

A "target polynucleotide" refers to the nucleic acid sequence of interest *e.g.*, the BRCA1 encoding polynucleotide. Other primers which can be used for primer hybridization will be known or readily ascertainable to those of skill in the art.

"Consensus" means the most commonly occurring in the population.

"Consensus genomic sequence" means the allele of the target gene which occurs with the greatest frequency in a population of individuals having no family history of disease associated with the target gene.

"Substantially complementary to" refers to a probe or primer sequences which hybridize to the sequences provided under stringent conditions and/or sequences having sufficient homology with BRCA1 sequences, such that the allele specific oligonucleotide probe or primers hybridize to the BRCA1 sequences to which they are complimentary.

"Haplotype" refers to a series of alleles within a gene on a chromosome.

"Isolated" as used herein refers to substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which they may be associated. Such association is typically either in cellular material or in a synthesis medium.

"Mutation" refers to a base change or a gain or loss of base pair(s) in a DNA sequence, which results in a DNA sequence which codes for a non-functioning protein or a protein with substantially reduced or altered function.

"Polymorphism" refers to a base change which is not associated with known pathology.

"Tumor growth inhibitor protein" refers to the protein coded for by the BRCA1 gene. The functional protein is thought to suppress breast and ovarian tumor growth.

Sub B1

The invention in several of its embodiments includes:

1. An isolated consensus DNA sequence of the BRCA1 coding sequence as set forth in SEQ. ID. NO.: 1.

2. A consensus protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 2.

3. An isolated coding sequence of the BRCA1 gene as set forth in SEQ. ID. NO.: 3.

4. A protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 4 .

5. An isolated coding sequence of the BRCA1 gene as set forth in SEQ. ID. NO.: 5.

6. A protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 6.

7. A BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer which comprises an alternative pair of codons, AGC

and AGT, which occur at position 2201 at frequencies of about 35-45%, and from about 55-65%, respectively.

8. A BRCA1 gene according to Claim 7 wherein AGC occurs at a frequency of about 40%.

9. A set of at least two alternative codon pairs which occur at polymorphic positions in a BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer, wherein codon pairs are selected from the group consisting of:

- AGC and AGT at position 2201;
- TTG and CTG at position 2430;
- CCG and CTG at position 2731;
- GAA and GGA at position 3232;
- AAA and AGA at position 3667;
- TCT and TCC at position 4427; and
- AGT and GGT at position 4956.

Sub B1 cont

10. A set of at least two alternative codon pairs according to claim 9, wherein the codon pairs occur in the following frequencies, respectively, in a population of individuals free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about

35-45%, and from about 55-65%, respectively.

11 A set according to Claim 10 which is at least three codon pairs.

12 A set according to Claim 10 which is at least four codon pairs.

13. A set according to Claim 10 which is at least five codon pairs.

14. A set according to Claim 10 which is at least six codon pairs.

15 A set according to Claim 10 which is at least seven codon pairs.

16. A method of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with disease, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1(omi) DNA sequence, SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5;

(e) determining the presence or absence of each of the following polymorphic variation in said individual's BRCA1 coding sequence:

- AGC and AGT at position 2201,
- TTG and CTG at position 2430,
- CCG and CTG at position 2731,
- GAA and GGA at position 3232,
- AAA and AGA at position 3667,
- TCT and TCC at position 4427, and
- AGT and GGT at position 4956;

Sub B1 cont

(f) determining any sequence differences between said individual's BRCA1 coding sequences and SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5 wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427, and 4956, is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

17. A method of claim 16 wherein, codon variations occur at the following frequencies, respectively, in a population of individuals free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

18. A method according to claim 16 wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label a bioluminescent label, a chemiluminescent label, or an enzyme label.

19. A method of detecting a increased genetic susceptibility to breast and

ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence, SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5;

(e) determining any sequence differences between said individual's BRCA1 coding sequences and SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5; to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a base change which is not any one of the following:

- AGC and AGT at position 2201,
- TTG and CTG at position 2430,
- CCG and CTG at position 2731,
- GAA and GGA at position 3232,
- AAA and AGA at position 3667,
- TCT and TCC at position 4427, and
- AGT and GGT at position 4956 is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

20. A method of claim 19 wherein, codon variations occur at the following frequencies, respectively, in a population free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 40%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about

35-45%, and from about 55-65%, respectively;

- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

Sub B1 cont

21. A method according to claim 19 wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label a bioluminescent label, a chemiluminescent label, or an enzyme label.

22. A set of codon pairs, which occur at polymorphic positions in a BRCA1 gene with a BRCA1 coding sequence according to Claim 1, wherein said set of codon pairs is:

- AGC and AGT at position 2201;
- TTG and CTG at position 2430;
- CCG and CTG at position 2731;
- GAA and GGA at position 3232;
- AAA and AGA at position 3667;
- TCT and TCC at position 4427; and
- AGT and GGT at position 4956.

23. A set of at least two alternative codon pairs according to claim 22 wherein set of at least two alternative codon pairs occur at the following frequencies:

- at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

24. A BRCA1 coding sequence according to claim 1 wherein the codon pairs occur at the following frequencies:

- at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

25. A method of determining the consensus genomic sequence or consensus coding sequence for a target gene, comprising:

a) screening a number of individuals in a population for a family history which indicates inheritance of normal alleles for a target gene;

b) isolating at least one allele of the target gene from individuals found to

have a family history which indicates inheritance of normal alleles for a target gene;

c) sequencing each allele;

d) comparing the nucleic acid sequence of the genomic sequence or of the coding sequence of each allele of the target gene to determine similarities and differences in the nucleic acid sequence; and

e) determining which allele of the target gene occurs with the greatest frequency.

26. A method of performing gene therapy, comprising:

a) transfecting cancer cell *in vivo* with an effective amount of a vector transformed with a BRCA1 coding sequences of SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, or SEQ. ID. NO.: 5;

b) allowing the cells to take up the vector, and

c) measuring a reduction in tumor growth.

27. A method of performing protein therapy, comprising:

a) injecting into a patient, an effective amount of BRCA1 tumor growth inhibiting protein of SEQ. ID. NO.: 2, SEQ. ID. NO.: 4, or SEQ. ID. NO.: 6;

b) allowing the cells to take up the protein, and

c) measuring a reduction in tumor growth.

SEQUENCING

Any nucleic acid specimen, in purified or non-purified form, can be utilized as the starting nucleic acid or acids, providing it contains, or is suspected of containing, the specific nucleic acid sequence containing a polymorphic locus. Thus, the process may amplify, for example, DNA or RNA, including messenger RNA, wherein DNA or RNA may be single stranded or double stranded. In the event that RNA is to be used as a template, enzymes, and/or conditions optimal for reverse transcribing the template to DNA would be utilized. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture

of nucleic acids may also be employed, or the nucleic acids produced in a previous amplification reaction herein, using the same or different primers may be so utilized. See TABLE II. The specific nucleic acid sequence to be amplified, *i.e.,* the polymorphic locus, may be a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as contained in whole human DNA.

DNA utilized herein may be extracted from a body sample, such as blood, tissue material and the like by a variety of techniques such as that described by Maniatis, *et. al.* in *Molecular Cloning:A Laboratory Manual,* Cold Spring Harbor, NY, p 280-281, 1982). If the extracted sample is impure, it may be treated before amplification with an amount of a reagent effective to open the cells, or animal cell membranes of the sample, and to expose and/or separate the strand(s) of the nucleic acid(s). This lysing and nucleic acid denaturing step to expose and separate the strands will allow amplification to occur much more readily.

The deoxyribonucleotide triphosphates dATP, dCTP, dGTP, and dTTP are added to the synthesis mixture, either separately or together with the primers, in adequate amounts and the resulting solution is heated to about 90°-100°C from about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period, the solution is allowed to cool, which is preferable for the primer hybridization. To the cooled mixture is added an appropriate agent for effecting the primer extension reaction (called herein "agent for polymerization"), and the reaction is allowed to occur under conditions known in the art. The agent for polymerization may also be added together with the other reagents if it is heat stable. This synthesis (or amplification) reaction may occur at room temperature up to a temperature above which the agent for polymerization no longer functions. Thus, for example, if DNA polymerase is used as the agent, the temperature is generally no greater than about 40°C. Most conveniently the reaction occurs at room temperature.

The primers used to carry out this invention embrace oligonucleotides of

sufficient length and appropriate sequence to provide initiation of polymerization. Environmental conditions conducive to synthesis include the presence of nucleoside triphosphates and an agent for polymerization, such as DNA polymerase, and a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency in amplification, but may be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent for polymerization. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The oligonucleotide primer typically contains 12-20 or more nucleotides, although it may contain fewer nucleotides.

Primers used to carry out this invention are designed to be substantially complementary to each strand of the genomic locus to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands under conditions which allow the agent for polymerization to perform. In other words, the primers should have sufficient complementarity with the 5′ and 3′ sequences flanking the mutation to hybridize therewith and permit amplification of the genomic locus.

Oligonucleotide primers of the invention are employed in the amplification process which is an enzymatic chain reaction that produces exponential quantities of polymorphic locus relative to the number of reaction steps involved. Typically, one primer is complementary to the negative (-) strand of the polymorphic locus and the other is complementary to the positive (+) strand. Annealing the primers to denatured nucleic acid followed by extension with an enzyme, such as the large fragment of DNA polymerase I (Klenow) and nucleotides, results in newly synthesized + and - strands containing the target polymorphic locus sequence. Because these newly synthesized sequences are also templates, repeated cycles of denaturing, primer annealing, and extension results in exponential production of the region (*i.e.*, the target polymorphic locus sequence) defined by the primers. The product of the chain reaction is a discreet nucleic acid duplex with termini corresponding to the

ends of the specific primers employed.

The oligonucleotide primers of the invention may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage, et al., Tetrahedron Letters, 22:1859-1862, 1981. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066.

The agent for polymerization may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase, polymerase muteins, reverse transcriptase, other enzymes, including heat-stable enzymes (*e.i.*, those enzymes which perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation), such as *Taq* polymerase. Suitable enzyme will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each polymorphic locus nucleic acid strand. Generally, the synthesis will be initiated at the 3′ end of each primer and proceed in the 5′ direction along the template strand, until synthesis terminates, producing molecules of different lengths.

The newly synthesized strand and its complementary nucleic acid strand will form a double-stranded molecule under hybridizing conditions described above and this hybrid is used in subsequent steps of the process. In the next step, the newly synthesized double-stranded molecule is subjected to denaturing conditions using any of the procedures described above to provide single-stranded molecules.

The steps of denaturing, annealing, and extension product synthesis can be repeated as often as needed to amplify the target polymorphic locus nucleic acid sequence to the extent necessary for detection. The amount of the specific nucleic acid sequence produced will accumulate in an exponential fashion. Amplification is described in PCR. A Practical Approach, ILR Press, Eds. M. J.

McPherson, P. Quirke, and G. R. Taylor, 1992.

The amplification products may be detected by Southern blots analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing a very low level of the nucleic acid sequence of the polymorphic locus is amplified, and analyzed via a Southern blotting technique or similarly, using dot blot analysis. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal. Alternatively, probes used to detect the amplified products can be directly or indirectly detectably labeled, for example, with a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme. Those of ordinary skill in the art will know of other suitable labels for binding to the probe, or will be able to ascertain such, using routine experimentation.

Sequences amplified by the methods of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific DNA sequence such as PCR, oligomer restriction (Saiki, *et.al.*, *Bio/Technology*,3:1008-1012, 1985), allele-specific oligonucleotide (ASO) probe analysis (Conner, *et. al.*, *Proc. Natl. Acad. Sci.* U.S.A., 80:278, 1983), oligonucleotide ligation assays (OLAs) (Landgren, *et. al.*, *Science*,241:1007, 1988), and the like. Molecular techniques for DNA analysis have been reviewed (Landgren, *et. al.*, *Science*,242:229-237, 1988).

Preferably, the method of amplifying is by PCR, as described herein and as is commonly used by those of ordinary skill in the art. Alternative methods of amplification have been described and can also be employed as long as the BRCA1 locus amplified by PCR using primers of the invention is similarly amplified by the alternative means. Such alternative amplification systems include but are not limited to self-sustained sequence replication, which begins with a short sequence of RNA of interest and a T7 promoter. Reverse transcriptase copies the RNA into cDNA and degrades the RNA, followed by reverse transcriptase polymerizing a second strand of DNA. Another nucleic acid amplification technique is nucleic acid sequence-based amplification (NASBA) which uses reverse transcription and T7 RNA polymerase and

incorporates two primers to target its cycling scheme. NASBA can begin with either DNA or RNA and finish with either, and amplifies to 10^8 copies within 60 to 90 minutes. Alternatively, nucleic acid can be amplified by ligation activated transcription (LAT). LAT works from a single-stranded template with a single primer that is partially single-stranded and partially double-stranded. Amplification is initiated by ligating a cDNA to the promoter oligonucleotide and within a few hours, amplification is 10^8 to 10^9 fold. Another amplification system useful in the method of the invention is the QB Replicase System. The QB replicase system can be utilized by attaching an RNA sequence called MDV-1 to RNA complementary to a DNA sequence of interest. Upon mixing with a sample, the hybrid RNA finds its complement among the specimen's mRNAs and binds, activating the replicase to copy the tag-along sequence of interest. Another nucleic acid amplification technique, ligase chain reaction (LCR), works by using two differently labeled halves of a sequence of interest which are covalently bonded by ligase in the presence of the contiguous sequence in a sample, forming a new target. The repair chain reaction (RCR) nucleic acid amplification technique uses two complementary and target-specific oligonucleotide probe pairs, thermostable polymerase and ligase, and DNA nucleotides to geometrically amplify targeted sequences. A 2-base gap separates the oligonucleotide probe pairs, and the RCR fills and joins the gap, mimicking DNA repair. Nucleic acid amplification by strand displacement activation (SDA) utilizes a short primer containing a recognition site for *hincII* with short overhang on the 5′ end which binds to target DNA. A DNA polymerase fills in the part of the primer opposite the overhang with sulfur-containing adenine analogs. *HincII* is added but only cuts the unmodified DNA strand. A DNA polymerase that lacks 5′ exonuclease activity enters at the cite of the nick and begins to polymerize, displacing the initial primer strand downstream and building a new one which serves as more primer. SDA produces greater than 10^7-fold amplification in 2 hours at 37°C. Unlike PCR and LCR, SDA does not require instrumented Temperature cycling.

Another method is a process for amplifying nucleic acid sequences from a DNA or RNA template which may be purified or may exist in a mixture of

nucleic acids. The resulting nucleic acid sequences may be exact copies of the template, or may be modified. The process has advantages over PCR in that it increases the fidelity of copying a specific nucleic acid sequence, and it allows one to more efficiently detect a particular point mutation in a single assay. A target nucleic acid is amplified enzymatically while avoiding strand displacement. Three primers are used. A first primer is complementary to the first end of the target. A second primer is complementary to the second end of the target. A third primer which is similar to the first end of the target and which is substantially complementary to at least a portion of the first primer such that when the third primer is hybridized to the first primer, the position of the third primer complementary to the base at the 5′ end of the first primer contains a modification which substantially avoids strand displacement. This method is detailed in U.S. Patent 5,593,840 to Bhatnagar et al. 1997. Although PCR is the preferred method of amplification if the invention, these other methods can also be used to amplify the BRCA1 locus as described in the method of the invention.

The BRCA1$^{(omi)}$ DNA coding sequences were obtained by end to end sequencing of the BRCA1 alleles of five subjects in the manner described above followed by analysis of the data obtained. The data obtained provided us with the opportunity to evaluate seven previously published polymorphisms and to affirm or correct where necessary, the frequency of occurrence of alternative codons.

GENE THERAPY

The coding sequences can be used for gene therapy.

A variety of methods are known for gene transfer, any of which might be available for use.

Direct injection of Recombinant DNA in vivo

1. Direct injection of "naked" DNA directly with a syringe and needle into a specific tissue, infused through a vascular bed, or transferred through a catheter into endothelial cells.

2. Direct injection of DNA that is contained in artificially generated lipid vesicles.

3. Direct injection of DNA conjugated to a targeting structure, such as an

antibody.

4. Direct injection by particle bombardment, where the DNA is coated onto gold particles and shot into the cells.

Human Artificial Chromosomes

This novel gene delivery approach involves the use of human chromosomes that have been striped down to contain only the essential components for replication and the genes desired for transfer.

Receptor-Mediated Gene Transfer

DNA is linked to a targeting molecule that will bind to specific cell-surface receptors, inducing endocytosis and transfer of the DNA into mammalian cells. One such technique uses poly-L-lysine to link asialoglycoprotein to DNA. An adenovirus is also added to the complex to disrupt the lysosomes and thus allow the DNA to avoid degradation and move to the nucleus. Infusion of these particles intravenously has resulted in gene transfer into hepatocytes.

RECOMBINANT VIRUS VECTORS

Several vectors are used in gene therapy. Among them are the Moloney Murine Leukemia Virus (MoMLV) Vectors, the adenovirus vectors, the adeno-Associated Virus (AAV) vectors, the herpes simplex virus (HSV) vectors, the poxvirus vectors, and human immunodeficiency virus (HIV) vectors,

GENE REPLACEMENT AND REPAIR

The ideal genetic manipulation for treatment of a genetic disease would be the actual replacement of the defective gene with a normal copy of the gene. Homologous recombination is the term used for switching out a section of DNA and replacing it with a new piece. By this technique, the defective gene can be replaced with a normal gene which expresses a functioning BRCA1 tumor growth inhibitor protein.

A complete description of gene therapy can also be found in "Gene Therapy A Primer For Physicians 2d Ed. by Kenneth W. Culver, M.D. Publ. Mary Ann

15

Liebert Inc. (1996). Two Gene Therapy Protocols for BRCA1 are approved by the Recombinant DNA Advisory Committee for Jeffrey T. Holt *et al.*. They are listed as 9602-148, and 9603-149 and are available from the NIH. The isolated BRCA1 gene can be synthesized or constructed from amplification products and inserted into a vector such as the LXSN vector.

The BRCA1 amino acid and nucleic acid sequence may be used to make diagnostic probes and antibodies. Labeled diagnostic probes may be used by any hybridization method to determine the level of BRCA1 protein in serum or lysed cell suspension of a patient, or solid surface cell sample.

The BRCA1 amino acid sequence may be used to provide a level of protection for patients against risk of breast or ovarian cancer or to reduce the size of a tumor. Methods of making and extracting proteins are well known. Itakura *et al.* U.S. Patents 4,704,362, 5, 221, 619, and 5,583,013. BRCA1 has been shown to be secreted. Jensen, R.A. *et al. Nature Genetics* 12: 303-308 (1996).

EXAMPLE 1

Determination Of The Coding Sequence Of A BRCA1(omi) Gene From Five Individuals

MATERIALS AND METHODS

Approximately 150 volunteers were screened in order to identify individuals with no cancer history in their immediate family (i.e. first and second degree relatives). Each person was asked to fill out a hereditary cancer prescreening questionnaire See TABLE I below. Five of these were randomly chosen for end-to-end sequencing of their BRCA1 gene. A first degree relative is a parent, sibling, or offspring. A second degree relative is an aunt, uncle, grandparent, grandchild, niece, nephew, or half-sibling.

16

TTC X

TABLE I

Hereditary Cancer Pre-Screening Questionnaire

Part A: Answer the following questions about your family

1. To your knowledge, has anyone in your family been diagnosed with a very specific hereditary colon disease called Familial Adenomatous Polyposis (FAP)?
2. To your knowledge, have you or any aunt had breast cancer diagnosed before the age 35?
3. Have you had Inflammatory Bowel Disease, also called Crohn's Disease or Ulcerative Colitis, for more than 7 years?

Part B: Refer to the list of cancers below for your responses **only** to questions in Part B

Bladder Cancer	Lung Cancer	Pancreatic Cancer
Breast Cancer	Gastric Cancer	Prostate Cancer
Colon Cancer	Malignant Melanoma	Renal Cancer
Endometrial Cancer	Ovarian Cancer	Thyroid Cancer

4. Have your mother or father, your sisters or brothers or your children had any of the listed cancers?
5. Have there been diagnosed in your mother's brothers or sisters, or your mother's parents more than one of the cancers in the above list?
6. Have there been diagnosed in your father's brothers or sisters, or your father's parents more than one of the cancers in the above list?

Part C: Refer to the list of relatives below for responses **only** to questions in Part C

You	Your mother
Your sisters or brothers	Your mothers's sisters or brothers (maternal aunts and uncles)
Your children	Your mother's parents (maternal grandparents)

7. Have there been diagnosed in these relatives 2 or more identical types of cancer? Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
8. Is there a total of 4 or more of any cancers in the list of relatives above other than "simple" skin cancers?

Part D: Refer to the list of relatives below for responses **only** to questions in Part D.

You	Your father
Your sisters or brothers	Your fathers's sisters or brothers (paternal aunts and uncles)
Your children	Your father's parents (paternal grandparents)

9. Have there been diagnosed in these relatives 2 or more identical types of cancer? Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
10. Is there a total of 4 or more of any cancers in the list of relatives above other than "simple" skin cancers?

Genomic DNA was isolated from white blood cells of five subjects selected from analysis of their answers to the questions above. Dideoxy sequence analysis was performed following polymerase chain reaction amplification.

All exons of the BRCA1 gene were subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, *et al.*, Handbook of Techniques in Endocrine Research, p. 457-486, DePablo,F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye was attached for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated Model 377® sequencer. The software used for analysis of the resulting data was Sequence Navigator® software purchased through ABI.

1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of five subjects. Each of the five samples was sequenced end to end. Each sample was amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM $MgCl_2$), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer, 2.5 microliters reverse primer, and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The primers in Table II, below were used to carry out amplification of the various sections of the BRCA1 gene samples. The primers were synthesized on an DNA/RNA Model 394® Synthesizer.

18

TABLE II
BRCA1 PRIMERS AND SEQUENCING DATA

EXON	SEQUENCE		SEQ.ID NO.	MER	Mg++	SIZE
EXON 2	2F	5' GAA GTT GTC ATT TTA TAA ACC TTT-3'	7	24	1.6	~275
	2R	5' TGT CTT TTC TTC CCT AGT ATG T-3'	8	22		
EXON 3	3F	5' TCC TGA CAC AGC AGA CAT TTA-3'	9	21	1.4	~375
	3R	5' TTG GAT TTT CGT TCT CAC TTA-3'	10	21		
EXON 5	5F	5' CTC TTA AGG GCA GTT GTG AG-3'	11	20	1.2	~275
	5R	5' TTC CTA CTG TGG TTG CTT CC	12	20[1]		
EXON 6	6/7F	5' CTT ATT TTA GTG TCC TTA AAA GG-3'	13	23	1.6	~250
	6R	5' TTT CAT GGA CAG CAC TTG AGT G-3'	14	22		
EXON 7	7F	5' CAC AAC AAA GAG CAT ACA TAG GG-3'	15	23	1.6	~275
	6/7R	5' TCG GGT TCA CTC TGT AGA AG-3'	16	20		
EXON 8	8F1	5' TTC TCT TCA GGA GGA AAA GCA-3'	17	21	1.2	~270
	8R1	5' GCT GCC TAC CAC AAA TAC AAA-3'	18	21		
EXON 9	9F	5' CCA CAG TAG ATG CTC AGT AAATA-3'	19	23	1.2	~250
	9R	5' TAG GAA AAT ACC AGC TTC ATA GA-3'	20	23		
EXOM 10	10F	5' TGG TCA GCT TTC TGT AAT CG-3'	21	20	1.6	~250
	10R	5' GTA TCT ACC CAC TCT CTT CTT CAG-3'	22	24		
EXON 11A	11AF	5' CCA CCT CCA AGG TGT ATC A-3'	23	19	1.2	372
	11AR	5' TGT TAT GTT GGC TCC TTG CT-3'	24	20		
EXON 11B	11BF1	5' CAC TAA AGA CAG AAT GAA TCT A-3'	25	21	1.2	~400
	11BR1	5' GAA GAA CCA GAA TAT TCA TCT A-3'	26	21		
EXON 11C	11CF1	5' TGA TGG GGA GTC TGA ATC AA-3'	27	20	1.2	~400
	11CR1	5' TCT GCT TTC TTG ATA AAA TCC T-3'	28	22		
EXON 11D	11DF1	5' AGC GTC CCC TCA CAA ATA AA-3'	29	20	1.2	~400
	11DR1	5' TCA AGC GCA TGA ATA TGC CT-3'	30	20		
EXON 11E	11EF	5' GTA TAA GCA ATA TGG AAC TCG A-3'	31	22	1.2	388
	11ER	5' TTA AGT TCA CTG GTA TTT GAA CA-3'	32	23		
EXON 11F	11FF	5' GAC AGC GAT ACT TTC CCA GA-3'	33	20	1.2	382
	11FR	5' TGG AAC AAC CAT GAA TTA GTC-3'	34	21		
EXON 11G	11GF	5' GGA AGT TAG CAC TCT AGG GA-3'	35	20	1.2	423
	11GR	5' GCA GTG ATA TTA ACT GTC TGT A-3'	36	22		

[1] M13 tailed

EXON 11H	11HF	5' TGG GTC CTT AAA GAA ACA AAGT-3'	37	22	1.2	366
	11HR	5' TCA GGT GAC ATT GAA TCT TCC-3'	38	21		
EXON 11I	11IF	5' CCA CTT TTT CCC ATC AAG TCA-3'	39	21	1.2	377
	11IR	5' TCA GGA TGC TTA CAA TTA CTT C-3'	40	21		
EXON 11J	11JF	5' CAA AAT TGA ATG CTA TGC TTA GA-3'	41	23	1.2	377
	11JR	5' TCG GTA ACC CTG AGC CAA AT-3'	42	20		
EXON 11K	11KF	5' GCA AAAGCG TCC AGA AAG GA-3'	43	20	1.2	396
	11KR-1	5' TAT TTG CAG TCA AGT CTT CCA A-3'	44	22		
EXON 11L	11LF-1	5' GTA ATA TTG GCA AAG GCA TCT-3'	45	22	1.2	360
	11LR	5' TAA AAT GTG CTC CCC AAA AGC A-3'	46	22		
EXON 12	12F	5' GTC CTG CCA ATG AGA AGA AA-3'	47	20	1.2	~300
	12R	5' TGT CAG CAA ACC TAA GAA TGT-3'	48	21		
EXON 13	13F	5' AAT GGA AAG CTT CTC AAAGTA-3'	49	21	1.2	~325
	13R	5' ATG TTG GAG CTA GGT CCT TAC-3'	50	21		
EXON 14	14F	5' CTA ACC TGA ATT ATC ACT ATC A-3'	51	22	1.2	~310
	14R	5' GTG TAT AAATGC CTG TAT GCA-3'	52	21		
EXON 15	15F	5' TGG CTG CCC AGG AAG TAT G-3'	53	19	1.2	~375
	15R	5' AAC CAG AAT ATC TTT ATG TAG GA-3'	54	23		
EXON 16	16F	5' AAT TCT TAA CAG AGA CCA GAA C-3'	55	22	1.6	~550
	16R	5' AAA ACT CTT TCC AGA ATG TTG T-3'	56	22		
EXON 17	17F	5' GTG TAG AAC GTG CAG GAT TG-3'	57	20	1.2	~275
	17R	5' TCG CCT CAT GTG GTT TTA-3'	58	18		
EXON 18	18F	5' GGC TCT TTA GCT TCT TAG GAC-3'	59	21	1.2	~350
	18R	5' GAG ACC ATT TTC CCA GCA TC-3'	60	20		
EXON 19	19F	5' CTG TCA TTC TTC CTG TGC TC-3'	61	20	1.2	~250
	19R	5' CAT TGT TAA GGA AAG TGG TGC-3'	62	21		
EXON 20	20F	5' ATA TGA CGT GTC TGC TCC AC-3'	63	20	1.2	~425
	20R	5' GGG AAT CCA AAT TAC ACA GC-3'	64	20		
EXON 21	21F	5' AAG CTC TTC CTT TTT GAA AGT C-3'	65	22	1.6	~300
	21R	5' GTA GAG AAA TAG AAT AGC CTC T-3'	66	22		
EXON 22	22F	5' TCC CAT TGA GAG GTC TTG CT-3'	67	20	1.6	~300
	22R	5' GAG AAG ACT TCT GAG GCT AC-3'	68	20		
EXON 23	23F-1	5' TGA AGT GAC AGT TCC AGT AGT-3'	69	21	1.2	~250
	23R-1	5' CAT TTT AGC CAT TCA TTC AAC AA-3'	70	23		
EXON 24	24F	5' ATG AAT TGA CAC TAA TCT CTG C-3'	71	22	1.4	~285
	24R	5' GTA GCC AGG ACA GTA GAA GGA-3'	72	21		

Thirty-five cycles were performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time was increased to 5 minutes, and during the last cycle in which the extension time was increased to 5 minutes.

PCR products were purified using Qia-quick® PCR purification kits (Qiagen cat# 28104; Chatsworth, CA). Yield and purity of the PCR product determined spectrophotometrically at OD_{260} on a Beckman DU 650 spectrophotometer.

2. Dideoxy Sequence Analysis

Fluorescent dye was attached to PCR products for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated Model 377® sequencer. The software used for analysis of the resulting data was "Sequence Navigator® software" purchased through ABI.

3. RESULTS

Differences in the nucleic acids of the ten alleles from five individuals were found in seven locations on the gene. The changes and their positions are found on TABLE III, below.

TABLE III
PANEL TYPING

AMINO ACID CHANGE	NUCLEOTIDE CHANGE	1	2	3	4	5	FREQUENCY
SER(SER) (694)	11E	C/C	C/T	C/T	T/T	T/T	0.4 C 0.6 T
LEU(LEU) (771)	11F	T/T	C/T	C/T	C/C	C/C	0.4 T 0.6 C
PRO(LEU) (871)	11G	C/T	C/T	C/T	T/T	T/T	0.3 C 0.7 T
GLU(GLY) (1038)	11I	A/A	A/G	A/G	G/G	G/G	0.4 A 0.6 G
LYS(ARG) (1183)	11J	A/A	A/G	A/G	G/G	G/G	0.4 A 0.6 G
SER(SER) (1436)	13	T/T	T/T	T/C	C/C	C/C	0.5 T 0.5 C
SER(GLY) (1613)	16	A/A	A/G	A/G	G/G	G/G	0.4 A 0.6 G

Tables 3 and 4 depict one aspect of the invention, sets of at least two alternative codon pairs wherein the codon pairs occur in the following frequencies, respectively, in a population of individuals free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

The data show that for each of the samples. The BRCA1 gene is identical except in the region of seven polymorphisms. These polymorphic regions, together with their locations, the amino acid groups of each codon, the frequency of their occurrence and the amino acid coded for by each codon are found in TABLE IV below.

TABLE IV

CODON AND BASE CHANGES IN SEVEN POLYMORPHIC SITES OF BRCA1 GENE

SAMPLE NAME	BASE CHANGE	POSITION nt/aa	EXON	CODON CHANGE	AA CHANGE	PUBLISHED FREQUENCY[2]	FREQUENCY IN THIS STUDY
2,3,4,5	C-T	2201/694	11E	AGC(AGT)	SER-SER	UNPUBLISHED	C=40%
2,3,4,5	T-C	2430/771	11F	TTG(CTG)	LEU-LEU	T=67%[13]	T=40%
1,2,3,4,5	C-T	2731/871	11G	CCG(CTG)	PRO-LEU	C=34%[12]	C=30%

	2,3,4,5	A-G	3232/1038	11I	GAA(GGA)	GLU-GLY	A=67%[13]	A=40%
	2,3,4,5	A-G	3667/1183	11J	AAA(AGA)	LYS-ARG	A=68%[12]	A=40%
5	3,4,5	T-C	4427/1436	13	TCT(TCC)	SER-SER	T=67%[12]	T=50%
	2,3,4,5	A-G	4956/1613	16	AGT(GGT)	SER-GLY	A=67%[12]	A=40%

[2]Reference numbers correspond to the Table of References below.

10

EXAMPLE 2

Determination Of A Individual Using BRCA1$^{(OMI)}$ And The Seven Polymorphisms For Reference

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

a) identifying individuals having a BRCA1 gene, who are therefore have no elevated genetic susceptibility to breast or ovarian cancer from a BRCA1 mutation;

b) avoiding misinterpretation of polymorphisms found in the BRCA1 gene;

Sequencing is carried out as in EXAMPLE 1 using a blood sample from the patient in question. However, a BRCA1$^{(omi)}$ sequence is used for reference and the polymorphic sites are compared to the nucleic acid sequences listed above for codons at each polymorphic site. A sample is one which compares to a BRCA1$^{(omi)}$ sequence and contains one of the base variations which occur at each of the polymorphic sites. The codons which occur at each of the polymorphic sites are paired here reference.

- AGC and AGT at position 2201,
- TTG and CTG at position 2430,
- CCG and CTG at position 2731,
- GAA and GGA at position 3232,
- AAA and AGA at position 3667,
- TCT and TCC at position 4427, and
- AGT and GGT at position 4956.

The availability of these polymorphic pairs provides added assurance that one skilled in the art can correctly interpret the polymorphic variations without mistaking a variation for a mutation.

Exon 11 of the BRCA1 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, *et al.*, Handbook of Techniques in Endocrine Research, p. 457-486, DePablo,F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing is performed in

both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated Model 377® sequencer. The software used for analysis of the resulting data is "Sequence Navigator® software" purchased through ABI.

1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM $MgCl_2$), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA1-11K-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA1-11K-R, 10 micromolar solution),and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify a patient's sample BRCA1 gene are listed in Table II. The primers were synthesized on an DNA/RNA Model 394® Synthesizer. Thirty-five cycles are of amplification are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes, and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick® PCR purification kits (Qiagen, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product determined spectrophotometrically at OD_{260} on a Beckman DU 650 spectrophotometer.

2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated Model 377® sequencer. The software used for analysis of the resulting data is "Sequence Navigator® software" purchased through ABI. The BRCA1$^{(omi1)}$ SEQ. ID. NO.:1 sequence is entered into the Sequence Navigator® software as the Standard for comparison. The Sequence Navigator® software compares the sample sequence to the BRCA1$^{(omi1)}$ SEQ. ID. NO.:1 standard, base by base. The Sequence Navigator® software

highlights all differences between the BRCA1$^{(omi1)}$ SEQ. ID. NO.:1 DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA1$^{(omi1)}$ SEQ. ID. NO.:1 standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator® software and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known polymorphisms for position and base change. If the sample BRCA1 sequence matches the BRCA1$^{(omi1)}$ SEQ. ID. NO.:1 standard, with only variations within the known list of polymorphisms, it is interpreted as a gene sequence.

EXAMPLE 3

DETERMINING THE ABSENCE OF A MUTATION IN THE BRCA1 GENE USING BRCA1$^{(omi1)}$ AND SEVEN POLYMORPHISMS FOR REFERENCE

A person skilled in the art of genetic susceptibility testing will find the present invention useful for determining the presence of a known or previously unknown mutation in the BRCA1 gene. A list of mutations of BRCA1 is publicly available in the Breast Cancer Information Core at:
http://www.nchgr.nih.gov/dir/lab_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995).

Sequencing is carried out as in EXAMPLE 1 using a blood sample from the patient in question. However, a BRCA1$^{(omi)}$ sequence is used for reference and polymorphic sites are compared to the nucleic acid sequences listed above for codons at each polymorphic site. A sample is one which compares to the BRCA1$^{(omi2)}$ SEQ. ID. NO.: 3 sequence and contains one of the base variations which occur at each of the polymorphic sites. The codons which occur at each of the polymorphic sites are paired here reference.

- AGC and AGT at position 2201,
- TTG and CTG at position 2430,

- CCG and CTG at position 2731,
- GAA and GGA at position 3232,
- AAA and AGA at position 3667,
- TCT and TCC at position 4427, and
- AGT and GGT at position 4956.

The availability of these polymorphic pairs provides added assurance that one skilled in the art can correctly interpret the polymorphic variations without mistaking a variation for a mutation.

Exon 11 of the BRCA1 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, *et al.*, Handbook of Techniques in Endocrine Research, p. 457-486, DePablo,F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated Model 377® sequencer. The software used for analysis of the resulting data is "Sequence Navigator® software" purchased through ABI.

1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM $MgCl_2$), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA1-11K-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA1-11K-R, 10 micromolar solution),and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify a patient's sample BRCA1 gene are listed in Table II. The primers were synthesized on an DNA/RNA Model 394® Synthesizer. Thirty-five cycles are of amplification are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes, and during the

last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick® PCR purification kits (Qiagen, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product determined spectrophotometrically at OD_{260} on a Beckman DU 650 spectrophotometer.

2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated Model 377® sequencer. The software used for analysis of the resulting data is "Sequence Navigator® software" purchased through ABI. The BRCA1(omi2) SEQ. ID. NO.: 3 sequence is entered into the Sequence Navigator® software as the Standard for comparison. The Sequence Navigator® software compares the sample sequence to the BRCA1(omi2) SEQ. ID. NO.: 3 standard, base by base. The Sequence Navigator® software highlights all differences between the BRCA1(omi2) SEQ. ID. NO.: 3 DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA1(omi2) SEQ. ID. NO.: 3 standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator® software and printed on a color printer. The peaks are interpreted by the first primary technologist and also by a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known polymorphisms for position and base change. If the sample BRCA1 sequence matches the BRCA1(omi2) SEQ. ID. NO.: 3 standard, with only variations within the known list of polymorphisms, it is interpreted as a gene sequence.

29

EXAMPLE 4

DETERMINING THE PRESENCE OF A MUTATION IN THE BRCA1 GENE USING BRCA1(omi) AND SEVEN POLYMORPHISMS FOR REFERENCE

A person skilled in the art of genetic susceptibility testing will find the present invention useful for determining the presence of a known or previously unknown mutation in the BRCA1 gene. A list of mutations of BRCA1 is publicly available in the Breast Cancer Information Core at:

http://www.nchgr.nih.gov/dir/lab_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995). In this example, a mutation in exon 11 is characterized by amplifying the region of the mutation with a primer which matches the region of the mutation.

Exon 11 of the BRCA1 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, *et al.*, Handbook of Techniques in Endocrine Research, p. 457-486, DePablo,F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated Model 377® sequencer. The software used for analysis of the resulting data is "Sequence Navigator® software" purchased through ABI.

1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM $MgCl_2$), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA1-11K-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA1-11K-R, 10 micromolar solution),and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify segment K of exon 11 (where the mutation is found) are as follows:

30

BRCA1-11K-F: 5′-GCA AAA GCG TCC AGA AAG GA-3′ SEQ ID NO:69

BRCA1-11K-R: 5′-AGT CTT CCA ATT CAC TGC AC-3′ SEQ ID NO:70

The primers are synthesized on an DNA/RNA Model 394® Synthesizer.

Thirty-five cycles are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes, and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick® PCR purification kits (Qiagen, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product determined spectrophotometrically at OD_{260} on a Beckman DU 650 spectrophotometer.

2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator® Kit (Perkin-Elmer cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated Model 377® sequencer. The software used for analysis of the resulting data is "Sequence Navigator® software" purchased through ABI. The BRCA1(omi2) SEQ. ID. NO.: 3 sequence is entered into the Sequence Navigator® software as the Standard for comparison. The Sequence Navigator® software compares the sample sequence to the BRCA1(omi2) SEQ. ID. NO.: 3 standard, base by base. The Sequence Navigator® software highlights all differences between the BRCA1(omi2) SEQ. ID. NO.: 3 DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA1(omi2) SEQ. ID. NO.: 3 standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator® software and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known polymorphisms for position and base

change. Mutations are noted by the length of non-matching variation. Such a lengthy mismatch pattern occurs with deletions and substitutions.

3. Result

Using the above PCR amplification and standard fluorescent sequencing technology, The 3888delGA mutation may be found. The 3888delGA mutation The BRCA1 gene lies in segment "K" of exon 11. The DNA sequence results demonstrate the presence of a two base pair deletion at nucleotides 3888 and 3889 of the published BRCA1$^{(omi)}$ sequence. This mutation interrupts the reading frame of the BRCA1 transcript, resulting in the appearance of an in-frame terminator (TAG) at codon position 1265. This mutation is, therefore, predicted to result in a truncated, and most likely, non-functional protein. The formal name of the mutation will be 3888delGA. This mutation is named in accordance with the suggested nomenclature for naming mutations, Baudet, A *et al., Human Mutation* 2:245-248, (1993).

EXAMPLE 5

USE OF THE BRCA1$^{(omi1)}$ GENE THERAPY

The growth of ovarian, breast or prostate cancer can be arrested by increasing the expression of the BRCA1 gene where inadequate expression of that gene is responsible for hereditary ovarian, breast and prostate cancer. It has been demonstrated that transfection of BRCA1 into cancer cells inhibits their growth and reduces tumorigenesis. Gene therapy is performed on a patient to reduce the size of a tumor. The LXSN vector is transformed with any of the BRCA1$^{(omi1)}$ SEQ. ID. NO.:1, BRCA1$^{(omi2)}$ SEQ. ID. NO.:3, or BRCA1$^{(omi3)}$ SEQ. ID. NO.:5 coding region.

Vector

The LXSN vector is transformed with wildtype BRCA1$^{(omi1)}$ SEQ. ID. NO.:1 coding sequence. The LXSN-BRCA1$^{(omi1)}$ retroviral expression vector is constructed by cloning a *Sal*I-linkered BRCA1$^{(omi1)}$ cDNA (nucleotides 1-5711) into the *Xho*I site of the vector LXSN. Constructs are confirmed by DNA sequencing. Holt *et al. Nature Genetics* 12: 298-302 (1996).

32

Retroviral vectors are manufactured from viral producer cells using serum free and phenol-red free conditions and tested for sterility, absence of specific pathogens, and absence of replication-competent retrovirus by standard assays. Retrovirus is stored frozen in aliquots which have been tested.

Patients receive a complete physical exam, blood, and urine tests to determine overall health. They may also have a chest X-ray, electrocardiogram, and appropriate radiologic procedures to assess tumor stage.

Patients with metastatic ovarian cancer are treated with retroviral gene therapy by infusion of recombinant LXSN-BRCA1$^{(omi1)}$ retroviral vectors into peritoneal sites containing tumor, between 10^9 and 10^{10} viral particles per dose. Blood samples are drawn each day and tested for the presence of retroviral vector by sensitive polymerase chain reaction (PCR)-based assays. The fluid which is removed is analyzed to determine:

1. The percentage of cancer cells which are taking up the recombinant LXSN-BRCA1$^{(omi1)}$ retroviral vector combination. Successful transfer of BRCA1 gene into cancer cells is shown by both RT-PCR analysis and *in situ* hybridization.
RT-PCR is performed with by the method of Thompson *et al. Nature Genetics* 9: 444-450 (1995), using primers derived from BRCA1$^{(omi1)}$ SEQ. ID. NO.:1. Cell lysates are prepared and immunoblotting is performed by the method of Jensen *et al. Nature Genetics* 12: 303-308 1996) and Jensen *et al. Biochemistry* 31: 10887-10892 (1992).

2. Presence of programmed cell death using ApoTAG® *in situ* apoptosis detection kit (Oncor, Inc., Gaithersburg, Maryland) and DNA analysis.

3. Measurement of BRCA I gene expression by slide immunofluorescence or western blot.
Patients with measurable disease are also evaluated for a clinical response to LXSN-BRCAI, especially those that do not undergo a palliative intervention immediately after retroviral vector therapy. Fluid cytology, abdominal girth, CT scans of the abdomen, and local symptoms are followed.

33

For other sites of disease, conventional response criteria are used as follows:

1. Complete Response (CR), complete disappearance of all measurable lesions and of all signs and symptoms of disease for at least 4 weeks.

2. Partial Response (PR), decrease of at least 50% of the sum of the products of the 2 largest perpendicular diameters of all measurable lesions as determined by 2 observations not less than 4 weeks apart. To be considered a PR, no new lesions should have appeared during this period and none should have increased in size.

3. Stable Disease, less than 25% change in tumor volume from previous evaluations.

4. Progressive Disease, greater than 25% increase in tumor measurements from prior evaluations.

The number of doses depends upon the response to treatment.

For further information related to this gene therpay approach see in "BRCA1 Retroviral Gene Therapy for Ovarian Cancer" a Human Gene Transfer Protocol: NIH ORDA Registration #: 9603-149 Jeffrey Holt, JT, M.D. and Carlos L. Arteaga, M.D.

<u>TABLE OF REFERENCES</u>

1. Sanger, F., *et al., J. Mol. Biol.* 42:1617, (1980).
2. Beaucage, et al., Tetrahedron Letters 22:1859-1862, (1981).
3. Maniatis, *et. al.* in *Molecular Cloning:A Laboratory Manual,* Cold Spring Harbor, NY, p 280-281, (1982).
4. Conner, *et. al., Proc. Natl. Acad. Sci.* U.S.A. 80:278, (1983)
5. Saiki, *et.al.., Bio/Technology* 3:1008-1012, (1985)
6. Landgren, *et. al., Science* 241:1007,(1988)
7. Landgren, *et. al., Science* 242 :229-237, (1988).
8. PCR. A Practical Approach, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, (1992).
9. Easton *et al., American Journal of Human Genetics* 52:678-701, (1993).
10. U.S. Patent No. 4,458,066.
11. Rowell, S., *et al., American Journal of Human Genetics* 55:861-865, (1994)
12. Miki, Y. *et al., Science* 266:66-71, (1994).
13. Friedman, L. *et al., Nature Genetics* 8:399-404, (1994).

34

14. Baudet, A *et al., Human Mutation* 2:245-248, (1993).
15. Friend, S. *et al., Nature Genetics* 11:238, (1995).
16. Arteaga, CL and JT Holt *Cancer Research* 56:1098-1103 (1996).
17. Holt, JT *et al., Nature Genetics* 12:298-302 (1996).
18. Jensen, RA *etal., Nature Genetics* 12:303-308 (1996).
19. Steeg, P. *Nature Genetics* 12:223-225 (1996).
20. Thompson, ME *et al., Nature Genetics* 9: 444-450 (1995)
21. Holt, JT, and C. Arteaga, Gene Therapy Protocol ORDA #: 9603-149 ORDA approved Protocol for BRCA1 Gene Therapy .

"Breast and Ovarian cancer" is understood by those skilled in the art to include breast and ovarian cancer in women and also breast and prostate cancer in men. BRCA1 is associated genetic susceptibility to inherited breast and ovarian cancer in women and also breast and prostate cancer in men. Therefore, claims in this document which recite breast and/or ovarian cancer refer to breast, ovarian and prostate cancers in men and women. Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

35

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Schelter, Denise B.
Zeng, Bin

(ii) TITLE OF INVENTION: A Sequence of the Human BRCA1 Gene

(iii) NUMBER OF SEQUENCES: 78

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ONCORMED
(B) STREET: 200 PERRY PARKWAY
(C) CITY: GAITHERSBURG
(D) STATE: MD
(E) COUNTRY: USA
(F) ZIP: 20877

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: to be assigned
(B) FILING DATE: herewith
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: R. THOMAS GALLEGOS
(B) REGISTRATION NUMBER: 32,692

(C) REFERENCE/DOCKET NUMBER: PA-0054

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 301-527-2051
(B) TELEFAX: 301-208-6997

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CCAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT 900

ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA 960

GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA 1020

AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT 1080

GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG 1140

ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT 1200

CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA 1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG 1320

GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG 1380

AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA 1440

TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT 1500

TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC 1560

TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA 1620

AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG 1680

CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC 1740

AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT 1800

CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA 1860

AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC 1920

ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC 1980

ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA 2040

TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA 2100

GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA 2160

GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA 2220

AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT 2280

TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT 2340

CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG 2400

AAAGATCTGT AGAGAGTAGC AGTATTTCAC TGGTACCTGG TACTGATTAT GGCACTCAGG 2460

AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520

GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580

ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640

GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700

TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760

AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820

TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880

AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940

ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000

ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060

CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120

AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180

GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GGAGCCAGCT 3240

CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300

TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360

ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420

GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480

ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540

ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600

AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660

TCCAGAGAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720

GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780

AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840

CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900

TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960

AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020

GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080

CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140

TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200

TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260

ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320

AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380

ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCCTCT GCCCTTGAGG 4440

ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500

GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620

GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680

ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740

AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800

CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860

AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG 4980

ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040

CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100

AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160

CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220

TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280

AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340

TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400

TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460

AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520

GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580

TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1863 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1               5                  10                  15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
            20                  25                  30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
        35                  40                  45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
    50                  55                  60

Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65                  70                  75                  80

Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
                85                  90                  95

Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
            100                 105                 110

Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
        115                 120                 125

Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
    130                 135                 140

Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145                 150                 155                 160

Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
                165                 170                 175

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
            180                 185                 190
```

Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
195 200 205

Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
210 215 220

Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225 230 235 240

Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
245 250 255

His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
260 265 270

Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
275 280 285

Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290 295 300

Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305 310 315 320

Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325 330 335

Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu
340 345 350

Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu
355 360 365

Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu
370 375 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu

405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg
485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590

Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu

```
        835                 840                 845

Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
    850                 855                 860

Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu
865                 870                 875                 880

Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
                885                 890                 895
Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
            900                 905                 910

Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
        915                 920                 925

Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
    930                 935                 940

Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
945                 950                 955                 960

Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
                965                 970                 975

Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
            980                 985                 990

Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
        995                 1000                1005

Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
    1010                1015                1020

Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Gly Ala Ser
1025                1030                1035                1040

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
                1045                1050                1055
```

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Arg Gly
1170 1175 1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185 1190 1195 1200
Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
1205 1210 1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly
1220 1225 1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
1235 1240 1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
1250 1255 1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265 1270 1275 1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
1285 1290 1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
1300 1305 1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
1315 1320 1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
1330 1335 1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345 1350 1355 1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
1365 1370 1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
1380 1385 1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
1395 1400 1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
1410 1415 1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425 1430 1435 1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
1445 1450 1455

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500

Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
1700 1705 1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
1715 1720 1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
1730 1735 1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745 1750 1755 1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
1765 1770 1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
1780 1785 1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
1795 1800 1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
1810 1815 1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825 1830 1835 1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
1845 1850 1855

Gln Ile Pro His Ser His Tyr
1860

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT 900

ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA 960

GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA 1020

AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT 1080

GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG 1140

ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT 1200

CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA 1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG 1320

GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG 1380

AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA 1440

TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT 1500

TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC 1560

TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA 1620

AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG 1680

CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC 1740

AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT 1800

CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA 1860

AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC 1920

ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC 1980

ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA 2040

TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA 2100

GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA 2160

GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG CGATACTTTC CCAGAGCTGA 2220

AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT 2280

TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT 2340

CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG 2400

AAAGATCTGT AGAGAGTAGC AGTATTTCAT TGGTACCTGG TACTGATTAT GGCACTCAGG 2460

AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520

GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580

ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640

GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700

TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760

AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820

TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880

AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940

ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000

ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060

CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120

AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180

GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GAAGCCAGCT 3240

CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300

TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360

ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420

GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480

ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540

ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600

AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660

TCCAGAAAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720

GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780

AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840

CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900

TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960

AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020

GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080

CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140

TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200

TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260

ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320

AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380

ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCTTCT GCCCTTGAGG 4440

ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500

GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620

GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680

ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740

AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800

CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860

AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGAGTCC AGCTGCTGCT CATACTACTG 4980

ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040

CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100

AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160

CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220

TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280

AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340

TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400

TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460

AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520

GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580

TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1863 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1 5 10 15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
20 25 30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
35 40 45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
50 55 60

Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65 70 75 80

Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
85 90 95

Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
100 105 110

Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
115 120 125

Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
130 135 140

Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145 150 155 160

Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
165 170 175

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
180 185 190

Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
195 200 205

Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
210 215 220

Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225 230 235 240

Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
245 250 255

His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
260 265 270

Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
275 280 285

Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290 295 300

Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305 310 315 320

Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325 330 335

Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu
340 345 350

Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu
355 360 365

Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu
370 375 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu
405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg

485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590

Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
835 840 845

Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
850 855 860

Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu
865 870 875 880

Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
885 890 895

Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
900 905 910

Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
915 920 925

Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
930 935 940

Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
945 950 955 960

Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
965 970 975

Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
980 985 990

Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
995 1000 1005

Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
1010 1015 1020

Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Glu Ala Ser
1025 1030 1035 1040

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
1045 1050 1055

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Lys Gly
1170 1175 1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185 1190 1195 1200
Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
1205 1210 1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly
1220 1225 1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
1235 1240 1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
1250 1255 1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265 1270 1275 1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
1285 1290 1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
1300 1305 1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
1315 1320 1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
1330 1335 1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345 1350 1355 1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
1365 1370 1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
1380 1385 1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
1395 1400 1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
1410 1415 1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425 1430 1435 1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
1445 1450 1455

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500

Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala

1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Ser Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
1700 1705 1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
1715 1720 1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
1730 1735 1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745 1750 1755 1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
1765 1770 1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
1780 1785 1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
1795 1800 1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
1810 1815 1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825 1830 1835 1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
1845 1850 1855

Gln Ile Pro His Ser His Tyr
1860

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT 900

ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA 960

GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA 1020

AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT 1080

GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG 1140

ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT 1200

CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA 1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG 1320

GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG 1380

AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA 1440

TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT 1500

TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC 1560

TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA 1620

AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG 1680

CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC 1740

AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT 1800

CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA 1860

AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC 1920

ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC 1980

ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA 2040

TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA 2100

GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA 2160

GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA 2220

AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT 2280

TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT 2340

CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG 2400

AAAGATCTGT AGAGAGTAGC AGTATTTCAC TGGTACCTGG TACTGATTAT GGCACTCAGG 2460

AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520

GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580

ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640

GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700

TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760

AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820

TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880

AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940

ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000

ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060

CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120

AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180

GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GGAGCCAGCT 3240

CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300

TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360

ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420

GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480

ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540

ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600

AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660

TCCAGAGAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720

GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780

AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840

CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900

TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960

AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020

GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080

CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140

TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200

TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260

ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320

AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380

ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCTTCT GCCCTTGAGG 4440

ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500

GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620

GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680

ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740

AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800

CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860

AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG 4980

ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040

CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100

AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160

CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220

TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280

AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340

TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400

TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460

AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520

GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580

TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1863 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1               5                  10                  15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
            20                  25                  30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
        35                  40                  45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
    50                  55                  60

Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65                  70                  75                  80

Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
                85                  90                  95

Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
            100                 105                 110

Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
        115                 120                 125

Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
    130                 135                 140
```

Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145 150 155 160

Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
165 170 175

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
180 185 190

Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
195 200 205

Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
210 215 220

Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225 230 235 240

Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
245 250 255

His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
260 265 270

Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
275 280 285

Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290 295 300

Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305 310 315 320

Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325 330 335

Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu
340 345 350

Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu
355 360 365

Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu
370 375 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu
405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg
485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys

565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590

Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
835 840 845

Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
850 855 860

Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu
865 870 875 880

Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
885 890 895

Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
900 905 910

Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
915 920 925

Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
930 935 940

Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
945 950 955 960

Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
965 970 975

Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
980 985 990

Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
995 1000 1005

Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
1010 1015 1020

Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Gly Ala Ser
1025 1030 1035 1040

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
1045 1050 1055

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Arg Gly
1170 1175 1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185 1190 1195 1200

Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
1205 1210 1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly

1220 1225 1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
1235 1240 1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
1250 1255 1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265 1270 1275 1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
1285 1290 1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
1300 1305 1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
1315 1320 1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
1330 1335 1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345 1350 1355 1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
1365 1370 1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
1380 1385 1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
1395 1400 1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
1410 1415 1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425 1430 1435 1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
1445 1450 1455

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500

Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

```
Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
    1650                1655                1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665                1670                1675                1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
                1685                1690                1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
            1700                1705                1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
        1715                1720                1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
    1730                1735                1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745                1750                1755                1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
                1765                1770                1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
            1780                1785                1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
        1795                1800                1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
    1810                1815                1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825                1830                1835                1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
                1845                1850                1855

Gln Ile Pro His Ser His Tyr
            1860
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGTTGTCA TTTTATAAAC CTTT 24

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 2R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TGTCTTTTCT TCCCTAGTAT GT 22

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 3F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCCTGACACA GCAGACATTT A 21

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 3R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGGATTTTC GTTCTCACTT A 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 5F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTCTTAAGGG CAGTTGTGAG 20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 5R-M13* primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCTACTGT GGTTGCTTCC 20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 6/7F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTTATTTTAG TGTCCTTAAA AGG 23

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 6R

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTTCATGGAC AGCACTTGAG TG 22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 7F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CACAACAAAG AGCATACATA GGG 23

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 6/7R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGGGTTCAC TCTGTAGAAG 20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 8F1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCTCTTCAG GAGGAAAAGC A 21

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 8R1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCTGCCTACC ACAAATACAA A 21

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 9F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCACAGTAGA TGCTCAGTAA ATA 23

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 9R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TAGGAAAATA CCAGCTTCAT AGA 23

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 10F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGGTCAGCTT TCTGTAATCG 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 10R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTATCTACCC ACTCTCTTCT TCAG 24

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11AF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCACCTCCAA GGTGTATCA 19

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11AR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGTTATGTTG GCTCCTTGCT 20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11BF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CACTAAAGAC AGAATGAATC TA 22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11BR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GAAGAACCAG AATATTCATC TA 22

(2) INFORMATION FOR SEQ ID NO:27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11CF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGATGGGGAG TCTGAATCAA 20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11CR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TCTGCTTTCT TGATAAAATC CT 22

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11DF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AGCGTCCCCT CACAAATAAA 20

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11DR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTCAAGCGCAT GAATATGCCT 20

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11EF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTATAAGCAA TATGGAACTC GA 22

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11ER primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TTAAGTTCACT GGTATTTGAA CA 23

(2) INFORMATION FOR SEQ ID NO:33:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11FF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GACAGCGATA CTTTCCCAGA 20

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11FR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TGGAACAACC ATGAATTAGT C 21

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11GF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGAAGTTAGC ACTCTAGGGA 20

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11GR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GCAGTGATAT TAACTGTCTG TA 22

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11HF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TGGGTCCTTA AAGAAACAAA GT 22

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11HR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TCAGGTGACA TTGAATCTTC C 21

(2) INFORMATION FOR SEQ ID NO:39:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11IF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCACTTTTTC CCATCAAGTC A 21

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11IR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TCAGGATGCT TACAATTACT TC 22

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11JF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CAAAATTGAA TGCTATGCTT AGA 23

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11JR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCGGTAACCC TGAGCCAAAT 20

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11KF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCAAAAGCGT CCAGAAAGGA 20

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11KR-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TATTTGCAGT CAAGTCTTCC AA 22

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11LF-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTAATATTGG CAAAGGCATC T 21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 11LR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TAAAATGTGC TCCCCAAAAG CA 22

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 12F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTCCTGCCAA TGAGAAGAAA 20

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 12R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGTCAGCAAA CCTAAGAATG T 21

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AATGGAAAGC TTCTCAAAGT A 21

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATGTTGGAGC TAGGTCCTTA C 21

(2) INFORMATION FOR SEQ ID NO:51:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 14F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CTAACCTGAA TTATCACTAT CA 22

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 14R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTGTATAAAT GCCTGTATGC A 21

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 15F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TGGCTGCCCA GGAAGTATG 19

(2) INFORMATION FOR SEQ ID NO:54:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 15R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AACCAGAATA TCTTTATGTA GGA 23

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 16F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

AATTCTTAAC AGAGACCAGA AC 22

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 16R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GAAAACTCTTT CCAGAATGTT GT 22

(2) INFORMATION FOR SEQ ID NO:57:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 17F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GTGTAGAACG TGCAGGATTG 20

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 17R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCGCCTCATG TGGTTTTA 18

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 18F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GGCTCTTTAG CTTCTTAGGA C 21

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 18R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GGAGACCATTT TCCCAGCATC 20

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 19F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CTGTCATTCT TCCTGTGCTC 20

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 19R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CATTGTTAAG GAAAGTGGTG C 21

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 20F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

ATATGACGTG TCTGCTCCAC 20

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 20R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GGGAATCCAA ATTACACAGC 20

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 21F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

AAGCTCTTCC TTTTTGAAAG TC 22

(2) INFORMATION FOR SEQ ID NO:66:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 21R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GTAGAGAAAT AGAATAGCCT CT 22

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 22F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

TCCCATTGAG AGGTCTTGCT 20

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 22R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GAGAAGACTT CTGAGGCTAC 20

(2) INFORMATION FOR SEQ ID NO:69:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 23F-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TGAAGTGACA GTTCCAGTAG T 21

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 23R-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CATTTTAGCC ATTCATTCAA CAA 23

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 24F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ATGAATTGAC ACTAATCTCT GC 22

(2) INFORMATION FOR SEQ ID NO:72:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 24R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GTAGCCAGGA CAGTAGAAGG A 21

WE CLAIM:

1. An isolated consensus DNA sequence of the BRCA1 coding sequence as set forth in SEQ. ID. NO.: 1.

2. A consensus protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 2.

3. An isolated coding sequence of the BRCA1 gene as set forth in SEQ. ID. NO.: 3.

4. A protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 4 .

5. An isolated coding sequence of the BRCA1 gene as set forth in SEQ. ID. NO.: 5.

6. A protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 6.

7. A BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer which comprises an alternative pair of codons, AGC and AGT, which occur at position 2201 at frequencies of about 35-45%, and from about 55-65%, respectively.

8. A BRCA1 gene according to Claim 7 wherein AGC occurs at a frequency of about 40%.

9. A set of at least two alternative codon pairs which occur at polymorphic positions in a BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer, wherein codon pairs are selected from the group consisting of:

- AGC and AGT at position 2201;
- TTG and CTG at position 2430;

- CCG and CTG at position 2731;
- GAA and GGA at position 3232;
- AAA and AGA at position 3667;
- TCT and TCC at position 4427; and
- AGT and GGT at position 4956.

10. A set of at least two alternative codon pairs according to claim 9, wherein the codon pairs occur in the following frequencies, respectively, in a population of individuals free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

11 A set according to Claim 10 which is at least three codon pairs.

12 A set according to Claim 10 which is at least four codon pairs.

13. A set according to Claim 10 which is at least five codon pairs.

14. A set according to Claim 10 which is at least six codon pairs.

15 A set according to Claim 10 which is at least seven codon pairs.

16. A method of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with disease, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence, SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5;

(e) determining the presence or absence of each of the following polymorphic variation in said individual's BRCA1 coding sequence:

- AGC and AGT at position 2201,
- TTG and CTG at position 2430,
- CCG and CTG at position 2731,
- GAA and GGA at position 3232,
- AAA and AGA at position 3667,
- TCT and TCC at position 4427, and
- AGT and GGT at position 4956;

(f) determining any sequence differences between said individual's BRCA1 coding sequences and SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5 wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427, and 4956, is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

17. A method of claim 16 wherein, codon variations occur at the following frequencies, respectively, in a caucasian population of individuals free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

18. A method according to claim 16 wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label a bioluminescent label, a chemiluminescent label, or an enzyme label.

19. A method of detecting a increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence, SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5;

(e) determining any sequence differences between said individual's BRCA1 coding sequences and SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5; to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a base change which is not any one of the following:

- AGC and AGT at position 2201,
- TTG and CTG at position 2430,
- CCG and CTG at position 2731,

- GAA and GGA at position 3232,
- AAA and AGA at position 3667,
- TCT and TCC at position 4427, and
- AGT and GGT at position 4956 is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

20. A method of claim 19 wherein, codon variations occur at the following frequencies, respectively, in a population free of disease:

- at position 2201, AGC and AGT occur at frequencies from about 40%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

21. A method according to claim 19 wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label a bioluminescent label, a chemiluminescent label, or an enzyme label.

22. A set of codon pairs, which occur at polymorphic positions in a BRCA1 gene with a BRCA1 coding sequence according to Claim 1, wherein said set of codon pairs is:

- AGC and AGT at position 2201;
- TTG and CTG at position 2430;
- CCG and CTG at position 2731;
- GAA and GGA at position 3232;

- AAA and AGA at position 3667;
- TCT and TCC at position 4427; and
- AGT and GGT at position 4956.

23. A set of at least two alternative codon pairs according to claim 22 wherein set of at least two alternative codon pairs occur at the following frequencies:

- at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and
- at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

24. A BRCA1 coding sequence according to claim 1 wherein the codon pairs occur at the following frequencies:

- at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;
- at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;
- at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;
- at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from

about 45-55%, respectively; and
at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

25. A method of determining the consensus genomic sequence or consensus coding sequence for a target gene, comprising:

a) screening a number of individuals in a population for a family history which indicates inheritance of normal alleles for a target gene;

b) isolating at least one allele of the target gene from individuals found to have a family history which indicates inheritance of normal alleles for a target gene;

c) sequencing each allele;

d) comparing the nucleic acid sequence of the genomic sequence or of the coding sequence of each allele of the target gene to determine similarities and differences in the nucleic acid sequence; and

e) determining which allele of the target gene occurs with the greatest frequency.

26. A method of performing gene therapy, comprising:

a) transfecting cancer cell *in vivo* with an effective amount of a vector transformed with a BRCA1 coding sequences of SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, or SEQ. ID. NO.: 5;

b) allowing the cells to take up the vector, and

c) measuring a reduction in tumor growth.

27. A method of performing protein therapy, comprising:

a) injecting into a patient, an effective amount of BRCA1 tumor growth inhibiting protein of SEQ. ID. NO.: 2, SEQ. ID. NO.: 4, or SEQ. ID. NO.: 6;

b) allowing the cells to take up the protein, and

c) measuring a reduction in tumor growth.

ABSTRACT OF THE DISCLOSURE

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

~~isolated coding sequences and to the protein sequences they code for.~~ This invention is directed to three coding sequences of the BRCA1 gene. The three coding sequences, BRCA1$^{(omi1)}$, BRCA1$^{(omi2)}$, and BRCA1$^{(omi3)}$ and their frequencies of occurrence are provided together with the protein sequences they code for. Another aspect of this invention is a method of determining the consensus sequence for any gene. Another aspect of the invention is a method of identifying an individual having an increased genetic susceptibility to breast or ovarian cancer because they have inherited a causative mutation in their BRCA1 gene. This invention is also related to a method of performing gene therapy with any of the isolated BRCA1 coding sequences.

I hereby appoint Glenn E. Karta, Registration No. 30,649, and R. Thomas Gallegos, Registration No. 32,692, my attorneys with full power of substitution and appointment of associate attorneys, to prosecute this application and to transact all business in the United States Patent and Trademark Office connected therewith.

Please address all correspondence to:

R. Thomas Gallegos
Codon Pharmaceuticals, Inc.
200 Perry Parkway
Gaithersburg, Maryland 20877
(301) 527-2051

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor: PATRICIA D. MURPHY

Inventor's signature: Patricia D Murphy

Date: 6/6/97
Residence: Slingerland, New York 12159 U.S.A.
Citizenship: United States of America
Post Office Address: 16 Stockbridge Road, Slingerlands, New York 12159

Full name of second joint inventor, if any: ANTONETTE C. ALLEN

Inventor's signature: Antonette C Allen

Date: 6/26/97
Residence: Severn, Maryland 21144 U.S.A.
Citizenship: United States of America
Post Office Address: 1405 Hawaii Avenue, Severn, Maryland 21144

Full name of third joint inventor, if any: CHRISTOPHER P. ALVARES

Inventor's signature: Christopher Alvares

Date: 6/17/97
Residence: Potomac, Maryland 20854 U.S.A.
Citizenship: United States of America

Post Office Address: 7547 Heatherton Lane, Potomac, Maryland 20854

Full name of fourth joint inventor, if any: BRENDA S. CRITZ

Inventor's signature: Brenda S Critz

Date: 6/18/97
Residence: Frederick, Maryland 21702 U.S.A.
Citizenship: United States of America
Post Office Address: 1602 Andover Lane, Frederick, Maryland 21702

Full name of fifth joint inventor, if any: SHERI L. OLSON

Inventor's signature: Sheri L Olson

Date: 6/18/97
Residence: Arlington, Virginia 22206 U.S.A.
Citizenship: United States of America
Post Office Address: 3000 S. Randolph Street, Apt. 144, Arlington, Virginia 22206

Full name of sixth joint inventor, if any: DENISE B. SCHELTER

Inventor's signature: Denise B Schelter

Date: 7/8/97
Residence: Silver Spring, Maryland 20902 U.S.A.
Citizenship: United States of America
Post Office Address: 2813 Parker Court, Silver Spring, Maryland 20902

Full name of seventh joint inventor, if any: BIN ZENG

Inventor's signature: Bin Zeng

Date: 6/25/97
Residence: Rockville, Maryland 20854 U.S.A.
Citizenship: Canada
Post Office Address: 121 Talbott Street, Rockville, Maryland 20852

VERIFIED STATEMENT CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) & 1.27 (c))--SMALL BUSINESS CONCERN	Docket No.: PA-0054 CIP

Applicant or Patentee: P. D. MURPHY, et al.
Serial or Patent No.:
Filed or Issued: February 12, 1997
Title: CODING SEQUENCES OF THE HUMAN BRCA1 GENE

I hereby declare that I am

___the owner of the small business concern identified below:

_X_ an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF SMALL BUSINESS CONCERN: ONCORMED, INC.
ADDRESS OF SMALL BUSINESS CONCERN: 205 PERRY PARKWAY
GAITHERSBURG, MD 20877

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention described in:

_X_ the specification filed herewith with title as listed above.
___ the application identified above.
___ the patent identified above.

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights in the invention must file separate verified statements averring to their status as small entities, and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern or organization having any rights in the invention is listed below:
_X_ no such person, concern or organization exists.
___ each such person, concern or organization is listed below.

Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING: DOUG DOLGINOW
TITLE OF PERSON IF OTHER THAN OWNER: President
ADDRESS OF PERSON SIGNING: 205 PERRY PARKWAY
GAITHERSBURG, MD 20877

SIGNATURE D. Dolginow DATE 2/12/97

DECLARATION AND POWER OF ATTORNEY	Docket: PA-0054 CIP

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled .

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

the specification of which

__X__ is attached hereto.
_____ was filed on as Application Serial No. and was amended on.

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application of which priority is claimed:

PRIOR FOREIGN APPLICATION(S)

			Priority claimed	
(Number)	(Day/month/year filed)	(Country)	Yes	No
(Number)	(Day/month/year filed)	(Country)	Yes	No
(Number)	(Day/month/year filed)	(Country)	Yes	No

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

08/598,591	February 12, 1996	Pending
(Application Serial No.)	(Filing Date)	(Status)
(Application Serial No.)	(Filing Date)	(Status)

2

I hereby appoint Glenn E. Karta, Registration No. 30,649, and R. Thomas Gallegos, Registration No. 32,692, my attorneys with full power of substitution and appointment of associate attorneys, to prosecute this application and to transact all business in the United States Patent and Trademark Office connected therewith.

Please address all correspondence to:

R. Thomas Gallegos
Codon Pharmaceuticals, Inc.
200 Perry Parkway
Gaithersburg, MD 20877
(301) 527-2051

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00 Full name of sole or first inventor: PATRICIA D. MURPHY et al

Inventor's signature: _____

Date: _____
Residence: Slingerland, New York 12159 U.S.A. NY
Citizenship: United States of America
Post Office Address: 16 Stockbridge Road, Slingerlands, New York 12159

2-00 Full name of second joint inventor, if any: ANTONETTE C. ALLEN

Inventor's signature: _____

Date: _____
Residence: Severn, Maryland 21144 U.S.A. MD
Citizenship: United States of America
Post Office Address: 1405 Hawaii Avenue, Maryland 21144

3-00 Full name of third joint inventor, if any: CHRISTOPHER P. ALVARES

Inventor's signature: _____

Date: _____
Residence: Potomac, Maryland 20854 U.S.A. MD
Citizenship: United States of America
Post Office Address: 7547 Heatherton Lane, Potomac, Maryland 20854

4-00 Full name of fourth joint inventor, if any: ~~BRENDA~~ S. CRITZ

Inventor's signature: _____

Date: _____
Residence: ~~Frederick~~, Maryland 21702 U.S.A. MD
Citizenship: United States of America
Post Office Address: 1602 Andover Lane, Frederick, Maryland 21702

5-00 Full name of fifth joint inventor, if any: SHERI J. OLSON

Inventor's signature: _____

Date: _____
Residence: Arlington, Virginia 22206 U.S.A. VA
Citizenship: United States of America
Post Office Address: 3000 S. Randolph Street, Apt. 144, Arlington, Virginia 22206

6-00 Full name of sixth joint inventor, if any: DENISE P. SCHELTER

Inventor's signature: _____

Date: _____
Residence: Silver Spring, Maryland 20902 U.S.A.
Citizenship: United States of America
Post Office Address: 2813 Parker Court, Maryland 20866

7-00 Full name of seventh joint inventor, if any: ~~BIN~~ ZENG

Inventor's signature: _____

Date: _____
Residence: Rockville, Maryland 20854 U.S. A. MD
Citizenship: Canada
Post Office Address: 121 Talbott Street, Rockville, Maryland 20852

FIGURE 1

File No.: PA-0054 CIP PATENT

CERTIFICATE OF MAILING UNDER 37 CFR 1.10

EXPRESS MAIL mailing label No.: EH788380659US
Date of Deposit: February 12, 1997

I hereby certify that the hereto attached paper(s), namely:

- New Patent Application Transmittal (in duplicate)
- Small Entity Status
- Declaration and Power of Attorney
- Patent Application (127 pages)

in connection with the following patent application:

Title: CODING SEQUENCES OF THE HUMAN BRCA1 GENE
Serial No.:
Applicants: PATRICIA D. MURPHY, ANTONETTE C. ALLEN, CHRISTOPHER P. ALVARES, BRENDA S. CRITZ, SHERI J. OLSON, DENISE B. SCHELTER, and BIN ZENG
Filing date:

is/are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service, on the date indicated above, and is addressed to the Commissioner of Patents and Trademarks, Box Patent Application, Washington, D.C. 20231.

Signature of person mailing paper(s):

R. Thomas Gallegos
Reg. No. 32,692

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NUMBER	FILING/RECEIPT DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NO./TITLE
08/798,691	02/12/97	MURPHY	PA-0054CIP

0222/0328

R THOMAS GALLEGOS
CODON PHARMACEUTICALS INC
200 PERRY PARKWAY
GAITHERSBURG MD 20877

NOT ASSIGNED

1806

DATE MAILED: 03/28/97

NOTICE TO FILE MISSING PARTS OF APPLICATION
Filing Date Granted

An Application Number and Filing Date have been assigned to this application. However, the items indicated below are missing. The required items and fees identified below must be timely submitted ALONG WITH THE PAYMENT OF A SURCHARGE for items 1 and 3-6 only of $ 130.00 for a ☑ large entity ☐ small entity in compliance with 37 CFR 1.27. The surcharge is set forth in 37 CFR 1.16(e). Applicant is given TWO MONTHS FROM THE DATE OF THIS NOTICE within which to file all required items and pay any fees required above to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

If all required items on this form are filed within the period set above, the total amount owed by applicant as a ☐ large entity ☐ small entity (verified statement filed), is $ 130.00.

☐ 1. The statutory basic filing fee is:
- ☐ missing.
- ☐ insufficient.

Applicant must submit $_____ to complete the basic filing fee and/or file a verified small entity statement claiming such status (37 CFR 1.27).

☐ 2. Additional claim fees of $_____, including any multiple dependent claim fees, are required.
Applicant must either submit the additional claim fees or cancel additional claims for which fees are due.

☐ 3. The oath or declaration:
- ☐ is missing.
- ☐ does not cover the newly submitted items.
- ☐ does not identify the application to which it applies.
- ☐ does not include the city and state or foreign country of applicant's residence.

An oath or declaration in compliance with 37 CFR 1. 63, including residence information and identifying the application by the above Application Number and Filing Date is required.

☐ 4. The signature(s) to the oath or declaration is/are:
- ☑ missing.
- ☐ by a person other than inventor or person qualified under 37 CFR 1.42, 1.43, or 1.47.

A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.

☐ 5. The signature of the following joint inventor(s) is missing from the oath or declaration:

_____

An oath or declaration listing the names of all inventors and signed by the omitted inventor(s), identifying this application by the above Application Number and Filing Date, is required.

☐ 6. A $_____ processing fee is required since your check was returned without payment (37 CFR 1.21(m)).

☐ 7. Your filing receipt was mailed in error because your check was returned without payment.

☐ 8. The application does not comply with the Sequence Rules.
See attached "Notice to Comply with Sequence Rules 37 CFR 1.821-1.825."

☐ 9. OTHER:

Direct the response and any questions about this notice to "Attention: Box Missing Parts."

A copy of this notice MUST be returned with the response.

Annette Rivers
Customer Service Center
Initial Patent Examination Division (703) 308-1202

File No. PA-0054 CIP PATENT

CERTIFICATE OF MAILING UNDER 37 CFR 1.8

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to: Commissioner of Patents and Trademarks, Box Missing Parts, Washington, D.C. 20231 on July 10, 1997.

R Thomas Gallegos, Reg. No. 32,692
R. Thomas Gallegos
Date of Signature: July 10, 1997

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Patricia D. Murphy, *et al.*

Serial No.: 08/798,691

Filed: February 12, 1997

Title: CODING SEQUENCES OF THE HUMAN BRCA1 GENE

Group Art Unit:

Examiner:

SUBMISSION OF MISSING PARTS

Hon. Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

In response to Form PTO-1533 mailed March 28, 1997 (copy attached), applicants submit herewith a declaration properly signed by joint inventor(s) Patricia D. Murphy, et al. The Commissioner is hereby requested to charge the $65 surcharge, as well as any other fees which may be occasioned by this paper, to Deposit Account 15-0609. This paper is enclosed in duplicate.

Respectfully submitted,

R Thomas Gallegos

R. Thomas Gallegos
Attorney for Applicants
Reg. No. 32,692
Codon Pharmaceuticals, Inc.
200 Perry Parkway
Gaithersburg, MD 20877
(301) 527-2051

A.N

File No.: PA-0054 CIP PATENT

CERTIFICATE OF MAILING UNDER 37 CFR 1.8

I hereby certify that this correspondence is being deposited with the United States Postal Service as First Class Mail in an envelope addressed to: Commissioner of Patents and Trademarks, Box Missing Parts, Washington, D.C. 20231 on July 10, 1997.

R Thomas Gallegos, Reg. No. 32,692
R. Thomas Gallegos
Date of Signature: July 10, 1997

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: P. D. MURPHY, *et al.*

Serial No.: 08/798,691

Filed: FEBRUARY 12, 1997

Title: CODING SEQUENCES OF THE HUMAN BRCA1 GENE

Group Art Unit:

Examiner:

PETITION FOR EXTENSION OF TIME (37 CFR 1.136(a))

Hon. Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

Pursuant to 37 CFR 1.136(a), applicants hereby petition for a two month extension of time within which to response to outstanding Notice of Missing Parts of Application in the above-identified application. The outstanding Action was mailed on March 28, 1997; hence applicant seeks an extension to and including July 28, 1997. The Commissioner is hereby requested to charge the appropriate fee of $195 (small entity) to Deposit Account 15-0609. The Commissioner is also authorized to charge

08/01/1997 DBEACH 00000012 DA#:150609 08798691
02 FC:216 195.00 CH

any additional fee which may be occasioned by this submission, or to credit any overpayment, to Deposit Account 15-0609. This paper is enclosed in duplicate.

Respectfully submitted,

R. Thomas Gallegos
Attorney for Applicants
Reg. No. 32,692
Codon Pharmaceuticals, Inc.
200 Perry Parkway
Gaithersburg, MD 20877
(301) 527-2051

VERIFIED STATEMENT CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) & 1.27 (c))--SMALL BUSINESS CONCERN	Docket No.: PA-0054 CIP

Applicant or Patentee: P. D. MURPHY, et al.
Serial or Patent No.: 08/798,691
Filed or Issued: February 12, 1997
Title: CODING SEQUENCES OF THE HUMAN BRCA1 GENE

I hereby declare that I am

___the owner of the small business concern identified below:

_X_ an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF SMALL BUSINESS CONCERN: ONCORMED, INC.
ADDRESS OF SMALL BUSINESS CONCERN: 205 PERRY PARKWAY
GAITHERSBURG, MD 20877

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention described in:

___ the specification filed herewith with title as listed above.
_X_ the application identified above.
___ the patent identified above.

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights in the invention must file separate verified statements averring to their status as small entities, and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern or organization having any rights in the invention is listed below:

_X_ no such person, concern or organization exists.
___ each such person, concern or organization is listed below.

Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING: DOUG DOLGINOW
TITLE OF PERSON IF OTHER THAN OWNER: President
ADDRESS OF PERSON SIGNING: 205 PERRY PARKWAY
GAITHERSBURG, MD 20877

SIGNATURE D. Dolginow DATE 6/9/97

DECLARATION AND POWER OF ATTORNEY	Docket: PA-0054 CIP

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

the specification of which

____ is attached hereto.
__X__ was filed on February 12, 1997 as Application Serial No. 08/798,691 and was amended on.

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application of which priority is claimed:

PRIOR FOREIGN APPLICATION(S)

			Priority claimed	
(Number)	(Day/month/year filed)	(Country)	Yes	No
(Number)	(Day/month/year filed)	(Country)	Yes	No
(Number)	(Day/month/year filed)	(Country)	Yes	No

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

08/598,591	February 12, 1996	Pending
(Application Serial No.)	(Filing Date)	(Status)
(Application Serial No.)	(Filing Date)	(Status)

A N

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NUMBER	FILING/RECEIPT DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NO./TITLE
08/798,691	02/12/97	MURPHY	PA-0054CIP

0222/0328

R THOMAS GALLEGOS
CODON PHARMACEUTICALS INC
200 PERRY PARKWAY
GAITHERSBURG MD 20877

NOT ASSIGNED

1806

DATE MAILED: 03/28/97

NOTICE TO FILE MISSING PARTS OF APPLICATION
Filing Date Granted

An Application Number and Filing Date have been assigned to this application. However, the items indicated below are missing. The required items and fees identified below must be timely submitted ALONG WITH THE PAYMENT OF A SURCHARGE for items 1 and 3-6 only of $ 130.00 for a ☑ large entity ☐ small entity in compliance with 37 CFR 1.27. The surcharge is set forth in 37 CFR 1.16(e). Applicant is given TWO MONTHS FROM THE DATE OF THIS NOTICE within which to file all required items and pay any fees required above to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

If all required items on this form are filed within the period set above, the total amount owed by applicant as a ☑ large entity ☐ small entity (verified statement filed), is $ 130.00.

☐ 1. The statutory basic filing fee is:
- ☐ missing.
- ☐ insufficient.

Applicant must submit $ _____ to complete the basic filing fee and/or file a verified small entity statement claiming such status (37 CFR 1.27).

☐ 2. Additional claim fees of $ _____, including any multiple dependent claim fees, are required.
Applicant must either submit the additional claim fees or cancel additional claims for which fees are due.

☐ 3. The oath or declaration:
- ☐ is missing.
- ☑ does not cover the newly submitted items.
- ☐ does not identify the application to which it applies.
- ☐ does not include the city and state or foreign country of applicant's residence.

An oath or declaration in compliance with 37 CFR 1. 63, including residence information and identifying the application by the above Application Number and Filing Date is required.

☑ 4. The signature(s) to the oath or declaration is/are:
- ☑ missing.
- ☐ by a person other than inventor or person qualified under 37 CFR 1.42, 1.43, or 1.47.

A properly signed oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.

☐ 5. The signature of the following joint inventor(s) is missing from the oath or declaration:

_____

An oath or declaration listing the names of all inventors and signed by the omitted inventor(s), identifying the application by the above Application Number and Filing Date, is required.

☐ 6. A $ _____ processing fee is required since your check was returned without payment (37 CFR 1.21(m)).

☐ 7. Your filing receipt was mailed in error because your check was returned without payment.

☐ 8. The application does not comply with the Sequence Rules.
See attached "Notice to Comply with Sequence Rules 37 CFR 1.821-1.825."

☐ 9. OTHER.

Direct the response and any questions about this notice to "Attention: Box Missing Parts."

A copy of this notice MUST be returned with the response.

Annette Rives

Customer Service Center
Initial Patent Examination Division (703) 308-1202

FORM PTO-1533 (REV.7-96) **PART 2 - COPY TO BE RETURNED WITH RESPONSE**

3/a
KDAVIS
7-23-97

PATENT
Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Murphy, Patricia D., et al.	
U.S. Serial No.: 08/798,691	Group Art Unit: 1807
Filing Date: February 12, 1997	Examiner: Rees, D.
For: **CODING SEQUENCES OF THE HUMAN BRCA1 GENE**	Atty. Docket No.: 05371.0014.999

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Prior to the commencement of the Examination of the above-identified application, Applicants herewith submit the following Amendment to the Specification and Remarks.

AMENDMENT

IN THE SPECIFICATION: 43-119

Please delete pages ~~45-121~~ and substitute new pages ~~45-101~~ 43-99 in lieu thereof.

TDavis 12/6/97

REMARKS

43-119

Applicants have amended the Specification to delete pages ~~45-121~~ and substitute therefore new pages ~~45-101~~ 43-99 corresponding to the Sequence Listing for the above-identified application. The replacement pages consist of the corrected version of the original Sequence Listing in which the title of the application had

TDavis 12/6/97

TDavis 12/6/97

been incorrectly stated and in which nucleotides had been incorrectly typed in SEQ. ID. NO. 1. No new matter has been added by this Amendment.

Respectfully submitted,

Suye Sirai for 32,692 REG. NO. 39,145

Dated: August 14, 1997

R. Thomas Gallegos (Reg. No.)

ONCORMED, INC.
200 Perry Parkway
Gaithersburg, Maryland 20877
(301) 527-2051

Enclosures

PATENT
Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Murphy, Patricia D., et al.

Serial No.: 08/798,691

Filing Date: February 12, 1997

For: **CODING SEQUENCES OF THE HUMAN BRCA1 GENE**

Group Art Unit: 1807

Examiner: Rees, D.

REQUIREMENTS FOR PATENT APPLICATION CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURE

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In response to the Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosure under 37 C.F.R. § 1.821 *et seq.*, in connection with the above-identified application, Applicants submit herewith Sequence Listing in paper and computer readable form.

I hereby state that the content of the paper and computer readable copies of the Sequence Listing, submitted in accordance with 37 C.F.R. § 1.821 (c) and (e), respectively, are the same.

I hereby state that the submission, filed in accordance with 37 C.F.R. § 1.821 (g), herein does not include new matter.

Applicants believe that no fee is due with the submission of the Sequence Listing in Response to the Notice to Comply. However, if any fee is required, please charge the required fee to Howrey & Simon Deposit Account No. 08-3038. A copy of this sheet is enclosed.

Respectfully submitted,

for REG. NO. 39,445
32,692

Dated: August 14, 1997

R. Thomas Gallegos (Reg. No.)

ONCORMED, INC.
200 Perry Parkway
Gaithersburg, Maryland 20877
(301) 527-2051

Enclosures

PATENT
Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Murphy, Patricia D., *et al.*

U.S. Serial No.: 08/798,691

Filing Date: February 12, 1997

For: **CODING SEQUENCES OF THE HUMAN BRCA1 GENE**

Group Art Unit: 1807

Examiner: Rees, D.

TRANSMITTAL LETTER

Assistant Commissioner for Patents
Washington, DC 20231

Sir:

The following are enclosed for consideration in the above-identified application:

		FEE
[ ]	Response to Final Office Action	$
[ ]	Response to Notice of Incomplete Application	$
[ ]	Declaration: [ ] Original; [ ] Supplemental	$
[ ]	Submission of Formal Drawings	$
[ ]	Formal Drawings: ___ Sheets ___ Figures	$
[ ]	Information Disclosure Statement and Form 1449 and ____ References	$
[X]	**Preliminary Amendment**	**$**
[ ]	Request for Extension of Time for ___ month	$
[ ]	Issue Fee: [ ] Part B - Issue Fee Transmittal [ ] Part C - Charge to Deposit Account	$
[X]	**Requirements for Patent Application Containing Nucleotide Sequence And/Or Amino Acid Sequence Disclosure**	**$**
[X]	**Sequence Diskette (1)** and Sequence Listing (pages 45-101)	$
[ ]	Request for Oral Hearing	$
[ ]	Filing Fees	$
TOTAL FEES BEING SUBMITTED		$ -0-

Respectfully submitted,

REG. NO.
for 39,445

Date: August 14, 1997

32,692
R. Thomas Gallegos (Reg. No.)

ONCORMED, INC.
200 Perry Parkway
Gaithersburg, Maryland 20877
(301) 527-2051

T36CX

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Schelter, Denise B.
Zeng, Bin

(ii) TITLE OF INVENTION: Coding Sequences of the Human BRCA1 Gene

(iii) NUMBER OF SEQUENCES: 72

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ONCORMED
(B) STREET: 200 Perry Parkway
(C) CITY: Gaithersberg
(D) STATE: MD
(E) COUNTRY: USA
(F) ZIP: 20877

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/798,691
(B) FILING DATE: 12-Feb-97
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Thomas Gallegos
(B) REGISTRATION NUMBER: 32,692
(C) REFERENCE/DOCKET NUMBER: PA-0054CIP

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 301-527-2051
(B) TELEFAX: 301-208-6997

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT 900

ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA 960

GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA 1020

AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT 1080

GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG 1140

ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT 1200

CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA 1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG 1320

GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG 1380

AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA 1440

TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT 1500

TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC 1560

TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA 1620

AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG 1680

CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC 1740

AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT 1800

CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA 1860

AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC 1920

ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC 1980

ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA 2040

TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA 2100

GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA 2160

GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA 2220

AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT 2280

TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT 2340

CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG 2400

AAAGATCTGT AGAGAGTAGC AGTATTTCAC TGGTACCTGG TACTGATTAT GGCACTCAGG 2460

AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520

GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580

ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640

GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700

TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760

AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820

TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880

AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940

ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000

ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060

CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120

AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180

GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GGAGCCAGCT 3240

CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300

TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360

ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420

GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480

ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540

ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600

AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660

TCCAGAGAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720

GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780

AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840

CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900

TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960

AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020

GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080

CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140

TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200

TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260

ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320

AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380

ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCCTCT GCCCTTGAGG 4440

ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500

GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620

GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680

ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740

AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800

CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860

AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG 4980

ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040

CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100

AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160

CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220

TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280

AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340

TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400

TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460

AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520

GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580

TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1863 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1               5                  10                  15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
            20                  25                  30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
        35                  40                  45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
    50                  55                  60

Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65                  70                  75                  80

Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
                85                  90                  95

Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
            100                 105                 110

Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
        115                 120                 125

Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
    130                 135                 140

Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145                 150                 155                 160

Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
                165                 170                 175

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
            180                 185                 190
```

Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
195 200 205

Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
210 215 220

Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225 230 235 240

Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
245 250 255

His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
260 265 270

Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
275 280 285

Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290 295 300

Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305 310 315 320

Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325 330 335

Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu
340 345 350

Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu
355 360 365

Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu
370 375 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu
405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg
485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590

Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser

770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
835 840 845

Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
850 855 860

Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu
865 870 875 880

Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
885 890 895
Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
900 905 910

Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
915 920 925

Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
930 935 940

Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
945 950 955 960

Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
965 970 975

Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
980 985 990

Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
995 1000 1005

Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
1010 1015 1020

Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Gly Ala Ser
1025 1030 1035 1040

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
1045 1050 1055

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Arg Gly
1170 1175 1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185 1190 1195 1200
Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
1205 1210 1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly
1220 1225 1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
1235 1240 1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
1250 1255 1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265 1270 1275 1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
1285 1290 1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
1300 1305 1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
1315 1320 1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
1330 1335 1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345 1350 1355 1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
1365 1370 1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
1380 1385 1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
1395 1400 1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
1410 1415 1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425 1430 1435 1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
1445 1450 1455

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500
Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu

1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
1700 1705 1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
1715 1720 1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
1730 1735 1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745 1750 1755 1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
1765 1770 1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
1780 1785 1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
1795 1800 1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
1810 1815 1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825 1830 1835 1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
1845 1850 1855

Gln Ile Pro His Ser His Tyr
1860

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT 900

ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA 960

GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA 1020

AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT 1080

GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG 1140

ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT 1200

CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA 1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG 1320

GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG 1380

AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA 1440

TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT 1500

TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC 1560

TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA 1620

AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG 1680

CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC 1740

AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT 1800

CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA 1860

AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC 1920

ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC 1980

ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA 2040

TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA 2100

GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA 2160

GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG CGATACTTTC CCAGAGCTGA 2220

AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT 2280

TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT 2340

CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG 2400

AAAGATCTGT AGAGAGTAGC AGTATTTCAT TGGTACCTGG TACTGATTAT GGCACTCAGG 2460

AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520

GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580

ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640

GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700

TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760

AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820

TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880

AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940

CH61

5856

ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000

ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060

CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120

AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180

GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GAAGCCAGCT 3240

CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300

TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360

ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420

GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480

ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540

ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600

AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660

TCCAGAAAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720

GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780

AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840

CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900

TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960

AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020

GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080

CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140

TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200

TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260

ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320

AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380

ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCTTCT GCCCTTGAGG 4440

ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500

GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620

GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680

ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740

AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800

CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860

AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGAGTCC AGCTGCTGCT CATACTACTG 4980

ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040

CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100

AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160

CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220

TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280

AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340

TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400

TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460

AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520

GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580

TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1863 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1               5                  10                  15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
            20                  25                  30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
        35                  40                  45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
    50                  55                  60

Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65                  70                  75                  80

Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
                85                  90                  95

Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
            100                 105                 110

Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
        115                 120                 125

Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
    130                 135                 140

Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145                 150                 155                 160

Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
                165                 170                 175

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
            180                 185                 190

Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
        195                 200                 205

Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
    210                 215                 220

Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225                 230                 235                 240
```


61 59

Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
245 250 255

His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
260 265 270

Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
275 280 285

Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290 295 300

Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305 310 315 320

Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325 330 335

Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu
340 345 350

Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu
355 360 365

Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu
370 375 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu
405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg
485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590
Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro

54

8361

820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
835 840 845

Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
850 855 860

Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu
865 870 875 880

Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
885 890 895

Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
900 905 910

Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
915 920 925

Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
930 935 940

Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
945 950 955 960

Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
965 970 975

Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
980 985 990

Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
995 1000 1005

Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
1010 1015 1020

Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Glu Ala Ser
1025 1030 1035 1040

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
1045 1050 1055

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Lys Gly
1170 1175 1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185 1190 1195 1200
Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
1205 1210 1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly
1220 1225 1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
1235 1240 1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
1250 1255 1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265 1270 1275 1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
1285 1290 1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
1300 1305 1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
1315 1320 1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
1330 1335 1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345 1350 1355 1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
1365 1370 1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
1380 1385 1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
1395 1400 1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
1410 1415 1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425 1430 1435 1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
1445 1450 1455

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500
Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Ser Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp

```
            1700                1705                1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
        1715                1720                1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
    1730                1735                1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745                1750                1755                1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
                1765                1770                1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
            1780                1785                1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
        1795                1800                1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
    1810                1815                1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825                1830                1835                1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
                1845                1850                1855

Gln Ile Pro His Ser His Tyr
            1860
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT 900

ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA 960

GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA 1020

AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT 1080

GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG 1140

ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT 1200

CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA 1260

AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG 1320

GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG 1380

AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA 1440

TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT 1500

TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC 1560

TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA 1620

AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG 1680

CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC 1740

AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT 1800

CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA 1860

AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC 1920

ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC 1980

ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA 2040

TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA 2100

GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA 2160

GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA 2220

AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT 2280

TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT 2340

CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG 2400

AAAGATCTGT AGAGAGTAGC AGTATTTCAC TGGTACCTGG TACTGATTAT GGCACTCAGG 2460

AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520

GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580

ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640

GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700

TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760

AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820

TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880

AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940

ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000

ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060

CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120

AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180

GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GGAGCCAGCT 3240

CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300

TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360

ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420

GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480

ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540

ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600

AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660

TCCAGAGAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720

GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780

AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840

CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900

TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960

AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020

GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080

CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140

TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200

TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260

ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320

AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380

ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCTTCT GCCCTTGAGG 4440

ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500

GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560

CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620

GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680

ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740

AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800

CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860

AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920

AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG 4980

ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040

CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100

AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160

CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220

TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280

AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340

TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400

TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460

AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520

GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580

TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

GTGTAGCACT CTACCAGTGC CAGGAGCTGG ACACCTACCT GATACCCCAG ATCCCCCACA 5700

GCCACTACTG A 5711

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1863 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn
1               5                  10                  15

Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys
            20                  25                  30

Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met
        35                  40                  45

Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys
    50                  55                  60

Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser
65                  70                  75                  80

Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp
                85                  90                  95

Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn
            100                 105                 110

Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met
        115                 120                 125

Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn
    130                 135                 140

Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly
145                 150                 155                 160

Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr
                165                 170                 175

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn
            180                 185                 190

Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr
        195                 200                 205

Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala
    210                 215                 220

Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln
225                 230                 235                 240

Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg
                245                 250                 255

His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu
            260                 265                 270

Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser
        275                 280                 285
```

Ser Leu Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe
290 295 300

Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg
305 310 315 320

Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr
325 330 335

Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu
340 345 350

Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu
355 360 365

Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu
370 375 380

Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp
385 390 395 400

Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu
405 410 415

Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu
420 425 430

Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His
435 440 445

Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr
450 455 460

Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn
465 470 475 480

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg
485 490 495

Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu
500 505 510

His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr
515 520 525

Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln
530 535 540

Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp
545 550 555 560

Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys
565 570 575

Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser
580 585 590
Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys
595 600 605

Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu
610 615 620

Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln
625 630 635 640

Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Lys Tyr Asn
645 650 655

Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys
660 665 670

Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr
675 680 685

Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn
690 695 700

Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu
705 710 715 720

Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu
725 730 735

Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu
740 745 750

Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser
755 760 765

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser
770 775 780

Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys
785 790 795 800

Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His
805 810 815

Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro
820 825 830

Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu
835 840 845

Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser
850 855 860

Lys Arg Gln Ser Phe Ala Leu Phe Ser Asn Pro Gly Asn Ala Glu Glu

865 870 875 880

Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser
885 890 895
Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys
900 905 910

Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly
915 920 925

Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys
930 935 940

Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly
945 950 955 960

Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn
965 970 975

Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr
980 985 990

Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met
995 1000 1005

Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser
1010 1015 1020

Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Gly Ala Ser
1025 1030 1035 1040

Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser
1045 1050 1055

Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu
1060 1065 1070

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val
1075 1080 1085

Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys
1090 1095 1100

His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val
1105 1110 1115 1120

Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro
1125 1130 1135

Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp
1140 1145 1150

Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn
1155 1160 1165

```
Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Arg Gly
    1170                1175                1180

Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln
1185                1190                1195                1200
Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu
                1205                1210                1215

Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly
            1220                1225                1230

Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala
        1235                1240                1245

Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys
    1250                1255                1260

Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser
1265                1270                1275                1280

Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe
                1285                1290                1295

Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr
            1300                1305                1310

Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser
        1315                1320                1325

Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp
    1330                1335                1340

Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser
1345                1350                1355                1360

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr
                1365                1370                1375

Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu
            1380                1385                1390

Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln
        1395                1400                1405

Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln
    1410                1415                1420

Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu
1425                1430                1435                1440

Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr
                1445                1450                1455
```

Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu
1460 1465 1470

Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn
1475 1480 1485

Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu
1490 1495 1500
Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg
1505 1510 1515 1520

Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu
1525 1530 1535

Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr
1540 1545 1550

Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile
1555 1560 1565

Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala
1570 1575 1580

Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu
1585 1590 1595 1600

Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Gly Pro Ala Ala
1605 1610 1615

Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val
1620 1625 1630

Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys
1635 1640 1645

Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu
1650 1655 1660

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile
1665 1670 1675 1680

Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val
1685 1690 1695

Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp
1700 1705 1710

Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met
1715 1720 1725

Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg
1730 1735 1740

Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile
1745 1750 1755 1760

Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro
1765 1770 1775

Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val
1780 1785 1790

Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val
1795 1800 1805
Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile
1810 1815 1820

Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp
1825 1830 1835 1840

Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro
1845 1850 1855

Gln Ile Pro His Ser His Tyr
1860

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGTTGTCA TTTTATAAAC CTTT 24

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 2R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TGTCTTTTCT TCCCTAGTAT GT 22

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 3F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TCCTGACACA GCAGACATTT A 21

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 3R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGGATTTTC GTTCTCACTT A 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 5F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTCTTAAGGG CAGTTGTGAG 20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 5R-M13* primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCTACTGT GGTTGCTTCC 20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 6/7F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTTATTTTAG TGTCCTTAAA AGG 23

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 6R

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTTCATGGAC AGCACTTGAG TG 22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 7F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CACAACAAAG AGCATACATA GGG 23

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 6/7R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGGGTTCAC TCTGTAGAAG 20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 8F1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCTCTTCAG GAGGAAAAGC A 21

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 8R1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCTGCCTACC ACAAATACAA A 21

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 9F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCACAGTAGA TGCTCAGTAA ATA 23

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 9R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TAGGAAAATA CCAGCTTCAT AGA 23

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 10F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGGTCAGCTT TCTGTAATCG 20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 10R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTATCTACCC ACTCTCTTCT TCAG 24

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11AF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCACCTCCAA GGTGTATCA 19

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11AR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGTTATGTTG GCTCCTTGCT 20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11BF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CACTAAAGAC AGAATGAATC TA 22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11BR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GAAGAACCAG AATATTCATC TA 22

(2) INFORMATION FOR SEQ ID NO:27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11CF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGATGGGGAG TCTGAATCAA 20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11CR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TCTGCTTTCT TGATAAAATC CT 22

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11DF1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

AGCGTCCCCT CACAAATAAA 20

(2) INFORMATION FOR SEQ ID NO:30:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11DR1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TCAAGCGCAT GAATATGCCT 20

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11EF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTATAAGCAA TATGGAACTC GA 22

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11ER primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TTAAGTTCACT GGTATTTGAA CA 23

(2) INFORMATION FOR SEQ ID NO:33:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11FF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GACAGCGATA CTTTCCCAGA 20

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11FR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TGGAACAACC ATGAATTAGT C 21

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11GF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GGAAGTTAGC ACTCTAGGGA 20

(2) INFORMATION FOR SEQ ID NO:36:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11GR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GCAGTGATAT TAACTGTCTG TA 22

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11HF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TGGGTCCTTA AAGAAACAAA GT 22

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11HR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TCAGGTGACA TTGAATCTTC C 21

(2) INFORMATION FOR SEQ ID NO:39:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11IF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCACTTTTTC CCATCAAGTC A 21

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11IR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TCAGGATGCT TACAATTACT TC 22

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11JF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CAAAATTGAA TGCTATGCTT AGA 23

(2) INFORMATION FOR SEQ ID NO:42:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11JR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TCGGTAACCC TGAGCCAAAT 20

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11KF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCAAAAGCGT CCAGAAAGGA 20

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11KR-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TATTTGCAGT CAAGTCTTCC AA 22

(2) INFORMATION FOR SEQ ID NO:45:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11LF-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTAATATTGG CAAAGGCATC T 21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 11LR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TAAAATGTGC TCCCCAAAAG CA 22

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 12F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTCCTGCCAA TGAGAAGAAA 20

(2) INFORMATION FOR SEQ ID NO:48:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 12R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGTCAGCAAA CCTAAGAATG T 21

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AATGGAAAGC TTCTCAAAGT A 21

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATGTTGGAGC TAGGTCCTTA C 21

(2) INFORMATION FOR SEQ ID NO:51:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 14F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CTAACCTGAA TTATCACTAT CA 22

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 14R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTGTATAAAT GCCTGTATGC A 21

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 15F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TGGCTGCCCA GGAAGTATG 19

(2) INFORMATION FOR SEQ ID NO:54:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 15R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AACCAGAATA TCTTTATGTA GGA 23

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 16F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

AATTCTTAAC AGAGACCAGA AC 22

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 16R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AAAACTCTTT CCAGAATGTT GT 22

(2) INFORMATION FOR SEQ ID NO:57:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 17F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GTGTAGAACG TGCAGGATTG 20

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 17R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCGCCTCATG TGGTTTTA 18

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 18F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GGCTCTTTAG CTTCTTAGGA C 21

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 18R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GAGACCATTT TCCCAGCATC 20

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 19F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CTGTCATTCT TCCTGTGCTC 20

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 19R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CATTGTTAAG GAAAGTGGTG C 21

(2) INFORMATION FOR SEQ ID NO:63:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 20F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

ATATGACGTG TCTGCTCCAC 20

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 20R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GGGAATCCAA ATTACACAGC 20

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 21F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

AAGCTCTTCC TTTTTGAAAG TC 22

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 21R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GTAGAGAAAT AGAATAGCCT CT 22

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 22F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

TCCCATTGAG AGGTCTTGCT 20

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 22R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GAGAAGACTT CTGAGGCTAC 20

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 23F-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TGAAGTGACA GTTCCAGTAG T 21

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 23R-1 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CATTTTAGCC ATTCATTCAA CAA 23

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(B) STRAIN: 24F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ATGAATTGAC ACTAATCTCT GC 22

(2) INFORMATION FOR SEQ ID NO:72:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(B) STRAIN: 24R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GTAGCCAGGA CAGTAGAAGG A 21

This Raw Listing contains the General Information Section and up to the first 5 pages.

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Schelter, Denise B.
Zeng, Bin

(ii) TITLE OF INVENTION: Coding Sequences of the Human BRCA1 Gene

(iii) NUMBER OF SEQUENCES: 72

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ONCORMED
(B) STREET: 200 Perry Parkway
(C) CITY: Gaithersberg
(D) STATE: MD
(E) COUNTRY: USA
(F) ZIP: 20877

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/798,691
(B) FILING DATE: 12-Feb-97
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Thomas Gallegos
(B) REGISTRATION NUMBER: 32,692
(C) REFERENCE/DOCKET NUMBER: PA-0054CIP

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 301-527-2051
(B) TELEFAX: 301-208-6997

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: Not Relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

```
100  CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT    900
101
102  ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA    960
103
104  GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA   1020
105
106  AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT   1080
107
108  GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG   1140
109
110  ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT   1200
111
112  CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA   1260
113
114  AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG   1320
115
116  GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG   1380
117
118  AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA   1440
119
120  TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT   1500
121
122  TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC   1560
123
124  TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA   1620
125
126  AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG   1680
127
128  CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC   1740
129
130  AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT   1800
131
132  CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA   1860
133
134  AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC   1920
135
136  ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC   1980
137
138  ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA   2040
139
140  TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA   2100
141
142  GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA   2160
143
144  GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA   2220
145
146  AGGTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT   2280
147
148  TTGTCAATCC TAGCCTTCCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT   2340
149
150  CTAATAATGC TGAAGACCCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG   2400
151
152  AAAGATCTGT AGAGAGTAGC AGTATTTCAC TGGTACCTGG TACTGATTAT GGCACTCAGG   2460
```

153
154 AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520
155
156 GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580
157
158 ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640
159
160 GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700
161
162 TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760
163
164 AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820
165
166 TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880
167
168 AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940
169
170 ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000
171
172 ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060
173
174 CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120
175
176 AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180
177
178 GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GGAGCCAGCT 3240
179
180 CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300
181
182 TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360
183
184 ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420
185
186 GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480
187
188 ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540
189
190 ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600
191
192 AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660
193
194 TCCAGAGAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720
195
196 GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780
197
198 AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840
199
200 CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900
201
202 TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960
203
204 AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020
205

206 GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT 4080
207
208 CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG 4140
209
210 TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA 4200
211
212 TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG 4260
213
214 ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC 4320
215
216 AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC 4380
217
218 ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCCTCT GCCCTTGAGG 4440
219
220 ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA 4500
221
222 GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG 4560
223
224 CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT 4620
225
226 GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA 4680
227
228 ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG 4740
229
230 AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA 4800
231
232 CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG 4860
233
234 AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA 4920
235
236 AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG 4980
237
238 ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG 5040
239
240 CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG 5100
241
242 AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA 5160
243
244 CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC 5220
245
246 TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC 5280
247
248 AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG 5340
249
250 TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT 5400
251
252 TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG 5460
253
254 AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG 5520
255
256 GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT 5580
257
258 TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA 5640

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION ***US/08/798,691A***

INPUT SET: S20528.raw

Line	Error	Original Text

HOWREY & SIMON

* COPY *

Attorneys at Law
1299 Pennsylvania Ave., N.W.
Washington, D.C. 20004-2402
(202) 783-0800
FAX (202) 383-6610

In Los Angeles
(213) 892-1800

September 4, 1997

HAND DELIVERY

Assistant Commissioner for Patents
1911 South Clark Street
Crystal Mall One
7th Floor
Arlington, Virginia 22202

Attention: Pam Davis

Re: PCT Patent Application Serial No.: PCT/US97/03038
U.S. Patent Application Serial No.: 08/798,691
Entitled: Coding Sequences of the Human BRCA1 Gene
Inventor(s): Patricia D. Murphy et al.
Our References: 05371.0014.228, 05371.0014.999
OncorMed Ref. No.: PA-0045-CIP PCT

Dear Ms. Davis:

Pursuant to our telephone conversation of today's date, enclosed please find duplicate copies of the documents relating to the filing of the Sequence Listing for the above-captioned application, i.e., Transmittal Letter, Requirements for Patent Application Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosure, Preliminary Amendment, Sequence Diskette (1), and Sequence Listing (pages 45-101). The original documents were filed with the PTO on August 14, 1997.

If you have any questions or need further information, please contact me at (202) 383-7073.

Very truly yours,

Elliot C. Mendelson

Enclosures

1806 #5 Penny B. 10/8/97

Express Mail No. _____

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of:	Murphy, Patricia D., *et al.*		
Serial No.:	08/798,691	Art Unit:	1397
Filed:	February 12, 1997	Examiner:	N/A
For:	**CODING SEQUENCES OF THE HUMAN BRCA1 GENE**	Attorney Docket No.:	**05371.0014.999**

POWER OF ATTORNEY

Asst. Commissioner of Patents
Washington, D.C. 20231

Dear Sir:

I hereby appoint the following attorneys to prosecute the above-identified patent application and to transact all business in the Patent and Trademark Office connected therewith.

Attorney	Registration No.
Jeffrey I. Auerbach	32,680
James F. Davis	21,072
Joel M. Freed	25,101
Alan M. Grimaldi	26,599
Albert P. Halluin	25,227
Richard H. Kjeldgaard	30,186
Joseph P. Lavelle	31,036
Stephen J. Rosenman	29,209
Melvin L. Barnes, Jr.	38,375
Michael J. Bell	39,604
Mark R. Buscher	35,006
Celine T. Callahan	34,301
Cono A. Carrano	39,623
Thomas M. Dunham	39,965
Michael N. Haynes	40,014
David R. Marsh	P-41,408
Joseph A. Micallef	39,772
Karen L. Nicastro	35,968
Timothy L. Scott	37,931

Serge Sira	39,445
Anthony D. Miller	34,394
Russell O. Paige	P-40,758
Stephen J. Pentlicki	40,125

Direct all telephone calls to Albert P. Halluin at (650) 463-8109.

Address all correspondence to: Albert P. Halluin
Howery & Simon
1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004

Respectfully submitted,

Dated: 26 Sept '97

By: R Thomas Gallegos
Thomas Gallegos Reg. No. 32,692

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

SERIAL NUMBER	FILING DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKETT NO.

EXAMINER	
ART UNIT	PAPER NUMBER
	6

DATE MAILED:

EXAMINER INTERVIEW SUMMARY RECORD

All participants (applicant, applicant's representative, PTO personnel):

(1) MINH-TAM DAVIS (3) Albert P. Halluin

(2) Tom Gallegos (4)

Date of interview Oct 17/97

Type: ☐ Telephonic ☐ Personal (copy is given to ☐ applicant ☒ applicant's representative).

Exhibit shown or demonstration conducted: ☐ Yes ☐ No. If yes, brief description:

Agreement ☐ was reached with respect to some or all of the claims in question. ☒ was not reached.

Claims discussed: all claims

Identification of prior art discussed:

Description of the general nature of what was agreed to if an agreement was reached, or any other comments:
1/ Differences among Reg TD nº 1, 3 and 5; 2/ deletion ~~of some~~ claims for divisional applications, and 3/ changes into Markush languages of cls 1, 6, 18 4/ will submit terminal disclaimer.

(A fuller description, if necessary, and a copy of the amendments, if available, which the examiner agreed would render the claims allowable must be attached. Also, where no copy of the amendments which would render the claims allowable is available, a summary thereof must be attached.)

☐ 1. It is not necessary for applicant to provide a separate record of the substance of the interview.

Unless the paragraph below has been checked to indicate to the contrary, A FORMAL WRITTEN RESPONSE TO THE LAST OFFICE ACTION IS NOT WAIVED AND MUST INCLUDE THE SUBSTANCE OF THE INTERVIEW (e.g., items 1-7 on the reverse side of this form). If a response to the last Office action has already been filed, then applicant is given one month from this interview date to provide a statement of the substance of the interview.

☐ 2. Since the examiner's interview summary above (including any attachments) reflects a complete response to each of the objections, rejections and requirements that may be present in the last Office action, and since the claims are now allowable, this completed form is considered to fulfill the response requirements of the last Office action. Applicant is not relieved from providing a separate record of the substance of the interview unless box 1 above is also checked.

Minh Tam B. Davis
Examiner's Signature

Patent
Attorney's Docket No. 05371.0014.999

#7
B. Henry 10/22/97

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of	)	
	)	
MURPHY, et al.	)	
	)	
Application No.: 08/798,691	)	Group Art Unit: 1809
	)	
Filed: February 12, 1997	)	Examiner: Rees, D.
	)	
For: CODING SEQUENCES OF THE HUMAN BRCA1 GENE	)	
	)	

INFORMATION DISCLOSURE STATEMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In accordance with the duty of disclosure as set forth in 37 C.F.R. §1.56, Applicants hereby submit the following information in conformance with 37 C.F.R. §§ 1.97 and 1.98. Pursuant to 37 C.F.R. § 1.98, a copy of each of the documents cited is enclosed.

U.S. PATENT DOCUMENTS

Caruthers, et al., U.S. Patent No. 4,458,066 - issued July 3, 1984

FOREIGN PATENT DOCUMENTS

European Patent Application No. 0705903 - October, 1996

I, the undersigned, hereby certify that the foreign patent listed above on this information disclosure statement was cited in a communication from the International Search Authority in a counterpart PCT application, which communication was received not more than three months prior to the filing of this statement.

OTHER DOCUMENTS

1. Arteaga, et al., "Tissue-Targeted Antisense c-*fos* Retroviral Vector Inhibits

Established Breast Cancer Xenografts in Nude Mice", *Cancer Research* 56:1098-1103 (1996);

2. Beaucage, et al., "Deoxynucleoside Phosphoramidites-A New Class of Key Intermediates for Deoxypolynucleotide Synthesis", *Tetrahedron Letters* 22(20): 1859-1662 (1991);
3. Beaudet, et al., "A Suggested Nomenclature for Designating Mutations", *Human Mutation* 2:245-248 (1993);
4. Conner, et al., "Detection of Sickle Cell β^s-globin Allele by Hybridization with Synthetic Oligonucleotides", *Proc. Natl. Acad. Sci.* 80:278-282 (1983);
5. Easton, et al., "Breast and Ovarian Cancer Incidence in BRCA1-Mutation Carriers", *American Journal of Human Genetics* 56:265-271 (1995);
6. Friedman, et al., "Confirmation of BRCA1 by Analysis of Germline Mutations Linked to Breast and Ovarian Cancer in Ten Families", *Nature Genetics* 8:399-404 (1994);
7. Friend, et al., Breast Cancer Information on the Web *Nature Genetics* 11:238 (1995);
8. Holt, et al., "Growth Retardation and Tumour Inhibition by BRCA1", *Nature Genetics* 12:298-302 (1996);
9. Holt, et al., Gene Therapy Protocol ORDA #: 9603-149 ORDA approved Protocol for BRCA1 Gene Therapy;
10. Hoskins, et al., "Assessment and Counseling for Women With a Family History of Breast Cancer", *JAMA*, 273(7):577-585 (1995);
11. Jensen, et al., "BRCA1 is Secreted and Exhibits Properties of a Granin", *Nature Genetics* 12:303-308 (1996);
12. Landgren, et al., "A Ligase-Mediated Gene Detection Technique", *Science* 241:1007-1021 (1988);
13. Landgren, et al., "DNA Diagnostics-Molecular Techniques and Automation", *Science* 242:229-237 (1988);
14. Miki, Y, et al., "A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1", *Science* 266:66-71 (1994);
15. Rowell, S., et al., "Inherited Predisposition to Breast and Ovarian Cancer", *American Journal of Human Genetics* 55:861-865 (1994);
16. Sanger, et al., "Cloning in Single-Stranded Bacteriophase as an Aid to Rapid DNA Sequencing", *J. Mol. Bio.*, 143:161-178 (1980);

17. Steeg, "Granin Expectations in Breast Cancer?", *Nature Genetics* 12:223-225 (1996); and

18. Thompson, et al., "Decreased Expression of BRCA1 Accelerates Growth and is Often Present During Sporadic Breast Cancer Progression", *Nature Genetics* 9:444-450 (1995).

The following references are not available at this time and will be provided at a later date:

1. Maniatis, et al., *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor, NY pp. 280-281 (1982);

2. PCR. A Practical Approach, ILR Press, Eds. M.J. McPherson, P. Quirke, and G. R. Taylor (1992); and

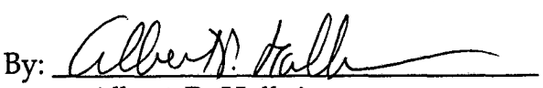
3. Saiki, et al., *Bio/Technology* 3:1008-1012 (1985).

The documents are being submitted within 3 months of the filing or entry of the national stage of this application or before the first Office Action on the merits, whichever is later, therefore no fee or certification is required under 37 C.F.R. § 1.97(b).

To assist the Examiner, the listed on the attached form PTO-1449. It is respectfully requested that an Examiner initialled copy of this form be returned to the undersigned.

Respectfully submitted,

HOWREY & SIMON

By: 

Albert P. Halluin
Registration No. 25,227

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109
Date: October 15, 1997

SHEET 1 OF 1

SUPPLEMENTAL INFORMATION DISCLOSURE CITATION PTO-1449	ATTORNEY'S DKT NO. 05371.0014.999	APPLICATION NO. 08/798,691
	APPLICANT MURPHY, et al.	
	FILING DATE February 12, 1997	GROUP 1809

U.S. PATENT DOCUMENTS

EXAMINER'S INITIALS	PATENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE
TDavis	5,693,473	Dec. 2, 1997	Shattuck-Eidens et al.	435	6	June 7, 1995

FOREIGN PATENT DOCUMENTS

EXAMINER'S INITIALS	PATENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translation	
						Yes	No

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

TDavis	Holt, et al., Gene Therapy Protocol ORDA #: 9603-149 ORDA approved Protocol for BRCA1 Gene Therapy, pp. 1003-1030 (1995) ~~(2 pages)~~.

EXAMINER	MINH-TAM DAVIS	DATE CONSIDERED	12/10/97

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

Patent
Attorney's Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of	)	
	)	
MURPHY, et al.	)	
	)	
Application No.: 08/798,691	)	Group Art Unit: 1809
	)	
Filed: February 12, 1997	)	Examiner: Rees, D.
	)	
For: CODING SEQUENCES OF THE HUMAN BRCA1 GENE	)	
	)	

INFORMATION DISCLOSURE STATEMENT TRANSMITTAL LETTER

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Enclosed is an Information Disclosure Statement and accompanying form PTO-1449 for the above-identified patent application.

[X] No additional fee for submission of an IDS is required.

[] The fee of $230.00 as set forth in 37 C.F.R. § 1.17(p) is also enclosed.

[] A certification under 37 C.F.R. § 1.97(e) is also enclosed.

[] A certification under 37 C.F.R. § 1.97(e), a petition requesting consideration of the information disclosure statement, and the petition fee of $130.00 as set forth in 37 C.F.R. § 1.17(i) are also enclosed.

[] Charge $_____ to Deposit Account No. 08-3038 for the fee due.

[] A check in the amount of $_____ is enclosed for the fee due.

Information Disclosure Statement Transmittal Letter
Application Serial No. 08/798,691
Attorney's Docket No. 05371.0014.999

The Commissioner is hereby authorized to charge any appropriate fees under 37 C.F.R. §§1.16, 1.17 and 1.21 that may be required by this paper, and to credit any overpayment, to Deposit Account No. 08-3038. This paper is submitted in duplicate.

Respectfully submitted,

HOWREY & SIMON

By: _____
Albert P. Halluin
Registration No. 25,227

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109

Date: October 15, 1997

5 9 8 5 9

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of: Murphy, *et al.*

Serial No.: 08/798,691 — Group Art Unit: 1807

Filed: February 12, 1997 — Examiner: Rees, D.

For: **CODING SEQUENCES OF THE HUMAN BRCA1 GENE** — Attorney Docket No.: 05371-0014-999

TERMINAL DISCLAIMER

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Petitioner, ONCORMED, INC., having a place of business at 205 Perry Parkway, Gaithersburg, MD 20877, is the assignee of record of the entire interest in U.S. Patent Application No. 08/798,691, filed February 12, 1997, the assignment of which is recorded at Reel 8600, Frame 0647, on July 14, 1997.

ONCORMED, INC. is also the assignee of record of the entire interest in U.S. Patent No. 5,654,155, which issued August 5, 1997, the assignment of which was recorded at Reel 8118, Frame 0910, on June 6, 1996.

PETITIONER hereby disclaims except as provided below the terminal part of the statutory term of any patent granted on U.S. Patent Application Serial No. 08/798,691 which would extend beyond the expiration date on the full statutory term defined in 35 U.S.C. 154 to 156 and 173 as shortened by any terminal disclaimer filed prior to the grant of U.S. Patent No. 5,654,155, and hereby agrees that any patent granted on U.S. Patent Application Serial No. 08/798,691 shall be enforceable only for and during such period that the legal title to said patent shall be the same as the legal title to U.S. Patent No. 5,654,155, this agreement to run with any patent

granted on the above-identified application and to be binding upon the grantee, its successors or assignees.

In making the above disclaimer, petitioner does not disclaim the terminal part of any patent granted on the instant application that would extend to the expiration date of the full statutory term as defined in 35 U.S.C. 154 to 156 and 173, as presently shortened by any terminal disclaimer, of U.S. Patent No. 5,654,155 in the event that it later: expires for failure to pay a maintenance fee, is held unenforceable, is found invalid by a court of competent jurisdiction, is statutorily disclaimed in whole or terminally disclaimed under 37 C.F.R. § 1.321, has all claims canceled by a reexamination certificate, is reissued, or is otherwise terminated prior to the expiration of its full statutory term as presently shortened by any terminal disclaimer filed prior to the grant of the patent.

The undersigned has reviewed all the evidentiary documents accompanying or referred to in the instant Terminal Disclaimer and it is certified to the best of the undersigned's knowledge and belief, title is in the assignee identified above.

The undersigned (whose title is supplied below) is empowered to act on behalf of the assignee.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

The $55 fee for this Terminal Disclaimer, as set forth in 37 C.F.R. § 1.20(d), is hereby proffered, and the Commissioner is authorized to charge this fee and any insufficiency, or to credit any overpayment of fees in connection with this communication to Howrey & Simon Deposit Account No. 08-3038. A duplicate copy of this paper is enclosed.

Respectfully submitted,

Date 17 October 1997

Albert P. Halluin, Reg. No. 25,227
Attorney of Record

HOWREY & SIMON
1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
Tel. (650)463-8100

GP1809
#9 11.28.97
TGray

EXPRESS MAIL NO. EM555261432US

Patent
Attorney's Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of

MURPHY, et al.

Application No.: 08/798,691 — Group Art Unit: 1800

Filed: February 12, 1997 — Examiner: Minh-Tam Davis

For: CODING SEQUENCES OF THE HUMAN BRCA1 GENE

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT TRANSMITTAL LETTER

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Enclosed is an Information Disclosure Statement and accompanying form PTO-1449 for the above-identified patent application.

[X] No additional fee for submission of an IDS is required.

[ ] The fee of $230.00 as set forth in 37 C.F.R. § 1.17(p) is also enclosed.

[ ] A certification under 37 C.F.R. § 1.97(e) is also enclosed.

[ ] A certification under 37 C.F.R. § 1.97(e), a petition requesting consideration of the information disclosure statement, and the petition fee of $130.00 as set forth in 37 C.F.R. § 1.17(i) are also enclosed.

[ ] Charge $_____ to Deposit Account No. 08-3038 for the fee due.

[ ] A check in the amount of $_____ is enclosed for the fee due.

The Commissioner is hereby authorized to charge any appropriate fees under 37 C.F.R. §§1.16, 1.17 and 1.21 that may be required by this paper, and to credit any overpayment, to Deposit Account No. 08-3038. This paper is submitted in duplicate.

Respectfully submitted,

HOWREY & SIMON

By: Albert P. Halluin / J. David Smith
Albert P. Halluin
Registration No. 25,227 39,839

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109

Date: October 30, 1997

The Commissioner is hereby authorized to charge any appropriate fees under 37 C.F.R. §§1.16, 1.17 and 1.21 that may be required by this paper, and to credit any overpayment, to Deposit Account No. 08-3038. This paper is submitted in duplicate.

Respectfully submitted,

HOWREY & SIMON

By: Albert P. Halluin / J. David Smith 39,839
Albert P. Halluin
Registration No. 25,227

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109

Date: October 30, 1997

EXPRESS MAIL NO. EM555261432US

Patent
Attorney's Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

re Patent Application of	)	
	)	
MURPHY, et al.	)	
	)	
Application No.: 08/798,691	)	Group Art Unit: 1809
	)	
Filed: February 12, 1997	)	Examiner: Minh-Tam Davis
	)	
For: CODING SEQUENCES OF THE HUMAN BRCA1 GENE	)	
	)	

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In accordance with the duty of disclosure as set forth in 37 C.F.R. §1.56, Applicants hereby submit the following information in conformance with 37 C.F.R. §§ 1.97 and 1.98. Pursuant to 37 C.F.R. § 1.98, a copy of each of the documents cited is enclosed.

OTHER DOCUMENTS

1. Maniatis, et al., "Isolation of High-Molecular-Weight, Eukaryotic DNA from Cells Grown in Tissue Culture," *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor, NY pp. 280-281 (1982);
2. Copy of cover page and two pages of book - PCR. A Practical Approach, ILR Press, Eds. M.J. McPherson, P. Quirke, and G. R. Taylor (1992); and
3. Saiki, et al., "A Novel Method for the Detection of Polymorphic Restriction Sites by Cleavage of Oligonucleotide Probes: Application to Sickle-Cell Anemia," *Bio/Technology* 3:1008-1012 (1985).

The documents are being submitted within 3 months of the filing or entry of the national stage of this application or before the first Office Action on the merits, whichever is later, therefore no fee or certification is required under 37 C.F.R. § 1.97(b).

To assist the Examiner, the listed on the attached form PTO-1449. It is respectfully requested that an Examiner initialled copy of this form be returned to the undersigned.

Respectfully submitted,

HOWREY & SIMON

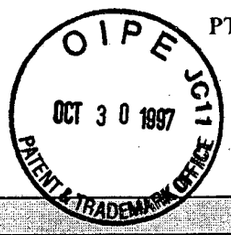
By: Albert P. Halluin /J. David Smith 39,839
Albert P. Halluin
Registration No. 25,227

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109

Date: October 30, 1997

SUPPLEMENTAL INFORMATION DISCLOSURE CITATION

PTO-1449



ATTORNEY'S DKT NO.	APPLICATION NO.
05371.0014.999	08/8798,691
APPLICANT: MURPHY, et al.	
FILING DATE: February 12, 1997	GROUP: 1809

U.S. PATENT DOCUMENTS

EXAMINER'S INITIALS	PATENT NO.	DATE	NAME	CLASS	SUBCLASS	FILING DATE

FOREIGN PATENT DOCUMENTS

EXAMINER'S INITIALS	PATENT NO.	DATE	COUNTRY	CLASS	SUBCLASS	Translation Yes	Translation No

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

MDavis	Maniatis, et al., "Isolation of High-Molecular-Weight, Eukaryotic DNA from Cells Grown in Tissue Culture," *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor, NY pp. 280-281 (1982);
	Copy of cover page and first two pages of book - PCR. A Practical Approach, ILR Press, Eds. M.J. McPherson, P. Quirke, and G. R. Taylor (1992); 3 pages.
	Saiki, et al., "A Novel Method for the Detection of Polymorphic Restriction Sites by Cleavage of Oligonucleotide Probes: Application to Sickle-Cell Anemia," *Bio/Technology* 3:1008-1012 (1985).

EXAMINER	DATE CONSIDERED
MINH-TAM DAVIS	12/3/97

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

Interview Summary

Application No.	Applicant(s)	
08/798,691	Murphy et al.	
Examiner	Group Art Unit	
Minh-Tam Davis	1806	

All participants (applicant, applicant's representative, PTO personnel):

(1) *Minh-Tam Davis* (3) _____

(2) *Albert Halluin, Tom Galligos* (4) _____

Date of Interview *December 4, 11, 12, 1997*

Type: ☒ Telephonic ☐ Personal (copy is given to ☐ applicant ☐ applicant's representative).

Exhibit shown or demonstration conducted: ☐ Yes ☒ No. If yes, brief description:

Agreement ☒ was reached. ☐ was not reached.

Claim(s) discussed: *all claims*

Identification of prior art discussed:
1) Shattuck-Eidens et al, 1997, US 5,693,473, 2) Friedman LS et al,1994, and 3) MPRSCH search data, Miki Y et al, 1994, concerning wild type nucleotides at the 7 claimed polymorphic sites.

Description of the general nature of what was agreed to if an agreement was reached, or any other comments:
1) Claims 3, 16-21, and 24 are allowable. 2) Claim 24 will have the additional language ....frequencies "in a Caucasian population of individual free of disease". 3) Claims 16, 17, 19, 20 will be corrected for formalities, and clarity. 4) In the abstract, the first sentence is incomplete, and will be corrected.

(A fuller description, if necessary, and a copy of the amendments, if available, which the examiner agreed would render the claims allowable must be attached. Also, where no copy of the amendents which would render the claims allowable is available, a summary thereof must be attached.)

1. ☒ It is not necessary for applicant to provide a separate record of the substance of the interview.

Unless the paragraph above has been checked to indicate to the contrary, A FORMAL WRITTEN RESPONSE TO THE LAST OFFICE ACTION IS NOT WAIVED AND MUST INCLUDE THE SUBSTANCE OF THE INTERVIEW. (See MPEP Section 713.04). If a response to the last Office action has already been filed, APPLICANT IS GIVEN ONE MONTH FROM THIS INTERVIEW DATE TO FILE A STATEMENT OF THE SUBSTANCE OF THE INTERVIEW.

2. ☒ Since the Examiner's interview summary above (including any attachments) reflects a complete response to each of the objections, rejections and requirements that may be present in the last Office action, and since the claims are now allowable, this completed form is considered to fulfill the response requirements of the last Office action. Applicant is not relieved from providing a separate record of the interview unless box 1 above is also checked.

Examiner Note: You must sign and stamp this form unless it is an attachment to a signed Office action.

Tam Davis

1806

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/798,691*

DATE: 09/08/97
TIME: 14:10:21

INPUT SET: S20151.raw

#10

This Raw Listing contains the General Information Section and up to the first 5 pages.

ENTERED

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Murphy, Patricia D.
Allen, Antonette C.
Alvares, Christopher P.
Critz, Brenda S.
Olson, Sheri J.
Schelter, Denise B.
Zeng, Bin

(ii) TITLE OF INVENTION: Coding Sequences of the Human BRCA1 Gene

(iii) NUMBER OF SEQUENCES: 72

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: ONCORMED
(B) STREET: 200 Perry Parkway
(C) CITY: Gaithersberg
(D) STATE: MD
(E) COUNTRY: USA
(F) ZIP: 20877

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/798,691
(B) FILING DATE: 12-Feb-97
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Thomas Gallegos
(B) REGISTRATION NUMBER: 32,692
(C) REFERENCE/DOCKET NUMBER: PA-0054CIP

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 301-527-2051
(B) TELEFAX: 301-208-6997

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5711 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: Not Relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: BRCA1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 17
(B) MAP POSITION: 17q21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCTCGCTGA GACTTCCTGG ACCCCGCACC AGGCTGTGGG GTTTCTCAGA TAACTGGGCC 60

CCTGCGCTCA GGAGGCCTTC ACCCTCTGCT CTGGGTAAAG TTCATTGGAA CAGAAAGAAA 120

TGGATTTATC TGCTCTTCGC GTTGAAGAAG TACAAAATGT CATTAATGCT ATGCAGAAAA 180

TCTTAGAGTG TCCCATCTGT CTGGAGTTGA TCAAGGAACC TGTCTCCACA AAGTGTGACC 240

ACATATTTTG CAAATTTTGC ATGCTGAAAC TTCTCAACCA GAAGAAAGGG CCTTCACAGT 300

GTCCTTTATG TAAGAATGAT ATAACCAAAA GGAGCCTACA AGAAAGTACG AGATTTAGTC 360

AACTTGTTGA AGAGCTATTG AAAATCATTT GTGCTTTTCA GCTTGACACA GGTTTGGAGT 420

ATGCAAACAG CTATAATTTT GCAAAAAAGG AAAATAACTC TCCTGAACAT CTAAAAGATG 480

AAGTTTCTAT CATCCAAAGT ATGGGCTACA GAAACCGTGC CAAAAGACTT CTACAGAGTG 540

AACCCGAAAA TCCTTCCTTG CAGGAAACCA GTCTCAGTGT CCAACTCTCT AACCTTGGAA 600

CTGTGAGAAC TCTGAGGACA AAGCAGCGGA TACAACCTCA AAAGACGTCT GTCTACATTG 660

AATTGGGATC TGATTCTTCT GAAGATACCG TTAATAAGGC AACTTATTGC AGTGTGGGAG 720

ATCAAGAATT GTTACAAATC ACCCCTCAAG GAACCAGGGA TGAAATCAGT TTGGATTCTG 780

CAAAAAAGGC TGCTTGTGAA TTTTCTGAGA CGGATGTAAC AAATACTGAA CATCATCAAC 840

```
100  CCAGTAATAA TGATTTGAAC ACCACTGAGA AGCGTGCAGC TGAGAGGCAT CCAGAAAAGT     900
101
102  ATCAGGGTAG TTCTGTTTCA AACTTGCATG TGGAGCCATG TGGCACAAAT ACTCATGCCA     960
103
104  GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA    1020
105
106  AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT    1080
107
108  GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG    1140
109
110  ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT    1200
111
112  CAGAGAATCC TAGAGATACT GAAGATGTTC CTTGGATAAC ACTAAATAGC AGCATTCAGA    1260
113
114  AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG    1320
115
116  GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTTCTAAAT GAGGTAGATG    1380
117
118  AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA    1440
119
120  TATGTAAAAG TGAAAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GACAAAATAT    1500
121
122  TTGGGAAAAC CTATCGGAAG AAGGCAAGCC TCCCCAACTT AAGCCATGTA ACTGAAAATC    1560
123
124  TAATTATAGG AGCATTTGTT ACTGAGCCAC AGATAATACA AGAGCGTCCC CTCACAAATA    1620
125
126  AATTAAAGCG TAAAAGGAGA CCTACATCAG GCCTTCATCC TGAGGATTTT ATCAAGAAAG    1680
127
128  CAGATTTGGC AGTTCAAAAG ACTCCTGAAA TGATAAATCA GGGAACTAAC CAAACGGAGC    1740
129
130  AGAATGGTCA AGTGATGAAT ATTACTAATA GTGGTCATGA GAATAAAACA AAAGGTGATT    1800
131
132  CTATTCAGAA TGAGAAAAAT CCTAACCCAA TAGAATCACT CGAAAAAGAA TCTGCTTTCA    1860
133
134  AAACGAAAGC TGAACCTATA AGCAGCAGTA TAAGCAATAT GGAACTCGAA TTAAATATCC    1920
135
136  ACAATTCAAA AGCACCTAAA AAGAATAGGC TGAGGAGGAA GTCTTCTACC AGGCATATTC    1980
137
138  ATGCGCTTGA ACTAGTAGTC AGTAGAAATC TAAGCCCACC TAATTGTACT GAATTGCAAA    2040
139
140  TTGATAGTTG TTCTAGCAGT GAAGAGATAA AGAAAAAAAA GTACAACCAA ATGCCAGTCA    2100
141
142  GGCACAGCAG AAACCTACAA CTCATGGAAG GTAAAGAACC TGCAACTGGA GCCAAGAAGA    2160
143
144  GTAACAAGCC AAATGAACAG ACAAGTAAAA GACATGACAG TGATACTTTC CCAGAGCTGA    2220
145
146  AGTTAACAAA TGCACCTGGT TCTTTTACTA AGTGTTCAAA TACCAGTGAA CTTAAAGAAT    2280
147
148  TTGTCAATCC TAGCCTTGCA AGAGAAGAAA AAGAAGAGAA ACTAGAAACA GTTAAAGTGT    2340
149
150  CTAATAATGC TGAAGACGCC AAAGATCTCA TGTTAAGTGG AGAAAGGGTT TTGCAAACTG    2400
151
152  AAAGATCTGT AGAGAGTAGC AGTATTTCAC TGGTACCTGG TACTGATTAT GGCACTCAGG    2460
```

153
154 AAAGTATCTC GTTACTGGAA GTTAGCACTC TAGGGAAGGC AAAAACAGAA CCAAATAAAT 2520
155
156 GTGTGAGTCA GTGTGCAGCA TTTGAAAACC CCAAGGGACT AATTCATGGT TGTTCCAAAG 2580
157
158 ATAATAGAAA TGACACAGAA GGCTTTAAGT ATCCATTGGG ACATGAAGTT AACCACAGTC 2640
159
160 GGGAAACAAG CATAGAAATG GAAGAAAGTG AACTTGATGC TCAGTATTTG CAGAATACAT 2700
161
162 TCAAGGTTTC AAAGCGCCAG TCATTTGCTC TGTTTTCAAA TCCAGGAAAT GCAGAAGAGG 2760
163
164 AATGTGCAAC ATTCTCTGCC CACTCTGGGT CCTTAAAGAA ACAAAGTCCA AAAGTCACTT 2820
165
166 TTGAATGTGA ACAAAAGGAA GAAAATCAAG GAAAGAATGA GTCTAATATC AAGCCTGTAC 2880
167
168 AGACAGTTAA TATCACTGCA GGCTTTCCTG TGGTTGGTCA GAAAGATAAG CCAGTTGATA 2940
169
170 ATGCCAAATG TAGTATCAAA GGAGGCTCTA GGTTTTGTCT ATCATCTCAG TTCAGAGGCA 3000
171
172 ACGAAACTGG ACTCATTACT CCAAATAAAC ATGGACTTTT ACAAAACCCA TATCGTATAC 3060
173
174 CACCACTTTT TCCCATCAAG TCATTTGTTA AAACTAAATG TAAGAAAAAT CTGCTAGAGG 3120
175
176 AAAACTTTGA GGAACATTCA ATGTCACCTG AAAGAGAAAT GGGAAATGAG AACATTCCAA 3180
177
178 GTACAGTGAG CACAATTAGC CGTAATAACA TTAGAGAAAA TGTTTTTAAA GGAGCCAGCT 3240
179
180 CAAGCAATAT TAATGAAGTA GGTTCCAGTA CTAATGAAGT GGGCTCCAGT ATTAATGAAA 3300
181
182 TAGGTTCCAG TGATGAAAAC ATTCAAGCAG AACTAGGTAG AAACAGAGGG CCAAAATTGA 3360
183
184 ATGCTATGCT TAGATTAGGG GTTTTGCAAC CTGAGGTCTA TAAACAAAGT CTTCCTGGAA 3420
185
186 GTAATTGTAA GCATCCTGAA ATAAAAAAGC AAGAATATGA AGAAGTAGTT CAGACTGTTA 3480
187
188 ATACAGATTT CTCTCCATAT CTGATTTCAG ATAACTTAGA ACAGCCTATG GGAAGTAGTC 3540
189
190 ATGCATCTCA GGTTTGTTCT GAGACACCTG ATGACCTGTT AGATGATGGT GAAATAAAGG 3600
191
192 AAGATACTAG TTTTGCTGAA AATGACATTA AGGAAAGTTC TGCTGTTTTT AGCAAAAGCG 3660
193
194 TCCAGAGAGG AGAGCTTAGC AGGAGTCCTA GCCCTTTCAC CCATACACAT TTGGCTCAGG 3720
195
196 GTTACCGAAG AGGGGCCAAG AAATTAGAGT CCTCAGAAGA GAACTTATCT AGTGAGGATG 3780
197
198 AAGAGCTTCC CTGCTTCCAA CACTTGTTAT TTGGTAAAGT AAACAATATA CCTTCTCAGT 3840
199
200 CTACTAGGCA TAGCACCGTT GCTACCGAGT GTCTGTCTAA GAACACAGAG GAGAATTTAT 3900
201
202 TATCATTGAA GAATAGCTTA AATGACTGCA GTAACCAGGT AATATTGGCA AAGGCATCTC 3960
203
204 AGGAACATCA CCTTAGTGAG GAAACAAAAT GTTCTGCTAG CTTGTTTTCT TCACAGTGCA 4020
205

```
206  GTGAATTGGA AGACTTGACT GCAAATACAA ACACCCAGGA TCCTTTCTTG ATTGGTTCTT    4080
207
208  CCAAACAAAT GAGGCATCAG TCTGAAAGCC AGGGAGTTGG TCTGAGTGAC AAGGAATTGG    4140
209
210  TTTCAGATGA TGAAGAAAGA GGAACGGGCT TGGAAGAAAA TAATCAAGAA GAGCAAAGCA    4200
211
212  TGGATTCAAA CTTAGGTGAA GCAGCATCTG GGTGTGAGAG TGAAACAAGC GTCTCTGAAG    4260
213
214  ACTGCTCAGG GCTATCCTCT CAGAGTGACA TTTTAACCAC TCAGCAGAGG GATACCATGC    4320
215
216  AACATAACCT GATAAAGCTC CAGCAGGAAA TGGCTGAACT AGAAGCTGTG TTAGAACAGC    4380
217
218  ATGGGAGCCA GCCTTCTAAC AGCTACCCTT CCATCATAAG TGACTCCTCT GCCCTTGAGG    4440
219
220  ACCTGCGAAA TCCAGAACAA AGCACATCAG AAAAAGCAGT ATTAACTTCA CAGAAAAGTA    4500
221
222  GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT GAGGTGTCTG    4560
223
224  CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AAGGTCATCC CCTTCTAAAT    4620
225
226  GCCCATCATT AGATGATAGG TGGTACATGC ACAGTTGCTC TGGGAGTCTT CAGAATAGAA    4680
227
228  ACTACCCATC TCAAGAGGAG CTCATTAAGG TTGTTGATGT GGAGGAGCAA CAGCTGGAAG    4740
229
230  AGTCTGGGCC ACACGATTTG ACGGAAACAT CTTACTTGCC AAGGCAAGAT CTAGAGGGAA    4800
231
232  CCCCTTACCT GGAATCTGGA ATCAGCCTCT TCTCTGATGA CCCTGAATCT GATCCTTCTG    4860
233
234  AAGACAGAGC CCCAGAGTCA GCTCGTGTTG GCAACATACC ATCTTCAACC TCTGCATTGA    4920
235
236  AAGTTCCCCA ATTGAAAGTT GCAGAATCTG CCCAGGGTCC AGCTGCTGCT CATACTACTG    4980
237
238  ATACTGCTGG GTATAATGCA ATGGAAGAAA GTGTGAGCAG GGAGAAGCCA GAATTGACAG    5040
239
240  CTTCAACAGA AAGGGTCAAC AAAAGAATGT CCATGGTGGT GTCTGGCCTG ACCCCAGAAG    5100
241
242  AATTTATGCT CGTGTACAAG TTTGCCAGAA AACACCACAT CACTTTAACT AATCTAATTA    5160
243
244  CTGAAGAGAC TACTCATGTT GTTATGAAAA CAGATGCTGA GTTTGTGTGT GAACGGACAC    5220
245
246  TGAAATATTT TCTAGGAATT GCGGGAGGAA AATGGGTAGT TAGCTATTTC TGGGTGACCC    5280
247
248  AGTCTATTAA AGAAAGAAAA ATGCTGAATG AGCATGATTT TGAAGTCAGA GGAGATGTGG    5340
249
250  TCAATGGAAG AAACCACCAA GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGATCT    5400
251
252  TCAGGGGGCT AGAAATCTGT TGCTATGGGC CCTTCACCAA CATGCCCACA GATCAACTGG    5460
253
254  AATGGATGGT ACAGCTGTGT GGTGCTTCTG TGGTGAAGGA GCTTTCATCA TTCACCCTTG    5520
255
256  GCACAGGTGT CCACCCAATT GTGGTTGTGC AGCCAGATGC CTGGACAGAG GACAATGGCT    5580
257
258  TCCATGCAAT TGGGCAGATG TGTGAGGCAC CTGTGGTGAC CCGAGAGTGG GTGTTGGACA    5640
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/798,691*

INPUT SET: S20151.raw

Line	Error	Original Text

Patent
Attorney's Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of	)	
	)	
MURPHY, et al.	)	Group Art Unit: 1806
	)	
Application No.: 08/798,691	)	Examiner: Minh-Tam Davis
	)	
Filed: February 12, 1997	)	
	)	
For: **CODING SEQUENCES OF THE HUMAN BRCA1 GENE**	)	

#11/B
Amdt
10/16/97

10/23/97
Official
Davis

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Prior to the commencement of the Examination of the above-identified application, Applicants herewith submit the following Amendment to the Specification and Remarks.

AMENDMENT

IN THE SPECIFICATION:

Please delete pages 8-15 and insert the following text in lieu thereof:

--"Mutation" refers to a base change or a gain or loss of base pair(s) in a DNA sequence, which results in a DNA sequence which codes for a non-functioning protein or a protein with substantially reduced or altered function.

B1

Certificate of Transmission

I hereby certify that this correspondence is being facsimile transmitted to the Patent and Trademark Office on October 23, 1997 to Examiner Minh-Tam Davis at fax number: (703) 305-7939.

Albert P. Halluin
Typed or printed name of person signing certificate

Albert P. Halluin
Signature

94

"Polymorphism" refers to a base change which is not associated with known pathology.

"Tumor growth inhibitor protein" refers to the protein coded for by the BRCA1 gene. The functional protein is thought to suppress breast and ovarian tumor growth.

One embodiment of the invention is an isolated consensus DNA sequence of the BRCA1 coding sequence as set forth in SEQ. ID. NO.: 1.

A further embodiment of the invention is a consensus protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 2.

B1 cont

A further embodiment of the invention is an isolated coding sequence of the BRCA1 gene as set forth in SEQ. ID. NO.: 3.

A further embodiment of the invention is a protein sequence of the BRCA1 protein as set forth in SEQ. ID. NO.: 4.

A further embodiment of the invention is an isolated coding sequence of the BRCA1 gene as set forth in SEQ. ID. NO.: 5.

A further embodiment of the invention is a protein sequence of the BRCA 1 protein as set forth in SEQ. ID. NO.: 6.

A further embodiment of the invention is a BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer which comprises an alternative pair of codons, AGC and AGT, which occur at position 2201 at frequencies of about 35-45%, and from about 55-65%, respectively.

A further embodiment of the invention is a BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer which comprises an alternative pair of codons, AGC and AGT, which occur at position 2201 at frequencies of about 40%, and from about 55-65%,

95

respectively.

A further embodiment of the invention is a set of at least two alternative codon pairs which occur at polymorphic positions in a BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer, wherein codon pairs are selected from the group consisting of:

(1) AGC and AGT at position 2201;

(2) TTG and CTG at position 2430;

(3) CCG and CTG at position 2731;

(4) GAA and GGA at position 3232;

(5) AAA and AGA at position 3667;

(6) TCT and TCC at position 4427; and

(7) AGT and GGT at position 4956.

A further embodiment of the invention is a set of at least two alternative codon pairs from the group of alternative codon pairs which occur at polymorphic positions in a BRCA1 gene with a BRCA1 coding sequence not associated with breast or ovarian cancer as recited above, wherein the codon pairs occur in the following frequencies, respectively, in a population of individuals free of disease:

(1) at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(2) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(3) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

(4) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

96

(5) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(6) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

(7) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

A further embodiment of the invention is a set of at least three alternative codon pairs from the group of alternative codon pairs as recited above.

A further embodiment of the invention is a set of at least four alternative codon pairs from the group of alternative codon pairs as recited above.

A further embodiment of the invention is a set of at least five alternative codon pairs from the group of alternative codon pairs as recited above.

A further embodiment of the invention is a set of at least six alternative codon pairs from the group of alternative codon pairs as recited above.

A further embodiment of the invention is a set of at least seven alternative codon pairs from the group of alternative codon pairs as recited above.

A further embodiment of the invention is a method of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with disease comprising the steps of:

(1) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(2) sequencing said amplified DNA fragment by dideoxy sequencing;

(3) repeating steps (1) and (2) until said individual's BRCA1 coding sequence is completely sequenced;

(4) comparing the sequence of said amplified DNA fragment to BRCA1$^{(omi)}$ DNA

sequence, SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, or SEQ. ID. NO.: 5;

(5) determining the presence or absence of each of the following polymorphic variations in said individual's BRCA1 coding sequence:

(a) AGC and AGT at position 2201;

(b) TTG and CTG at position 2430;

(c) CCG and CTG at position 2731;

(d) GAA and GGA at position 3232;

(e) AAA and AGA at position 3667;

(f) TCT and TCC at position 4427; and

(g) AGT and GGT at position 4956;

(6) determining any sequence differences between said individual's BRCA1 coding sequences and SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, or SEQ. ID. NO.: 5 wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427, and 4956, is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

A further embodiment of the invention is a method as described above, wherein codon variations occur at the following frequencies, respectively, in a population of individuals free of disease:

(1) at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(2) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(3) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

98

(4) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(5) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(6) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

(7) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

A further embodiment of the invention is a method as described above, wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label, a bioluminescent label, a chemiluminescent label, or an enzyme label.

B1 cont

A further embodiment of the invention is a method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising the steps of:

(1) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(2) sequencing said amplified DNA fragment by dideoxy sequencing;

(3) repeating steps (1) and (2) until said individual's BRCA1 coding sequence is completely sequenced;

(4) comparing the sequence of said amplified DNA fragment to BRCA1$^{(omi)}$ DNA sequence, SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, or SEQ. ID. NO.: 5;

(5) determining any sequence differences between said individual's BRCA1 coding sequences and SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, or SEQ. ID. NO.: 5 to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a said base change is not any one of the following:

99

(a) AGC and AGT at position 2201;

(b) TTG and CTG at position 2430;

(c) CCG and CTG at position 2731;

(d) GAA and GGA at position 3232;

(e) AAA and AGA at position 3667;

(f) TCT and TCC at position 4427; and

(g) AGT and GGT at position 4956; and is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

A further embodiment of the invention is a method according to the method above, wherein codon variations occur at the following frequencies, respectively, in a population free of disease:

(1) at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(2) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(3) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

(4) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(5) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(6) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

(7) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and

100

from about 55-65%, respectively.

A further embodiment of the invention is a method according to the method above, wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label, a bioluminescent label, a chemiluminescent label, or an enzyme label.

A further embodiment of the invention is a set of codon pairs, which occur at polymorphic position in a BRCA1 gene with a BRCA1 coding sequence as set forth in SEQ. ID. NO.: 1, wherein said set of codon pairs is:

(a) AGC and AGT at position 2201;

(b) TTG and CTG at position 2430;

(c) CCG and CTG at position 2731;

(d) GAA and GGA at position 3232;

(e) AAA and AGA at position 3667;

(f) TCT and TCC at position 4427; and

(g) AGT and GGT at position 4956.

A further embodiment of the invention is a set of at least two alternative codon pairs from the set of codon pairs above, wherein said set of at least two alternative codon pairs occur at the following frequencies:

(1) at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;

(2) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(3) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

(4) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(5) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(6) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

(7) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

A further embodiment of the invention is a BRCA1 coding sequence as set forth in SEQ. ID. NO.: 1, wherein the codon pairs occur at the following frequencies:

(1) at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;

(2) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

3l cont

(3) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

(4) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(5) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

(6) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

(7) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

A further embodiment of the invention is a method of determining the consensus genomic sequence or consensus coding sequence for a target gene, comprising the steps of:

(1) screening a number of individuals in a population for a family history which

102

indicates inheritance of normal alleles for a target gene;

(2) isolating at least one allele of the target gene from individuals found to have a family history which indicates inheritance of normal alleles for a target gene;

(3) sequencing each allele;

(4) comparing the nucleic acid sequence of the genomic sequence or of the coding sequence of each allele of the target gene to determine similarities and differences in the nucleic acid sequence; and

(5) determining which allele of the target gene occurs with the greatest frequency.

A further embodiment of the invention is a method of performing gene therapy, comprising the steps of:

(1) transfecting cancer cells *in vivo* with an effective amount of a vector transformed with a BRCA1 coding sequences of SEQ. ID. NO.: 1, SEQ. ID. NO.: 3, OR SEQ. ID. NO.: 5;

(2) allowing the cells to take up the vector; and

(3) measuring a reduction in tumor growth.

A further embodiment of the invention is a method of performing protein therapy, comprising the steps of:

(1) injecting into a patient an effective amount of BRCA1 tumor growth inhibiting protein of SEQ. ID. NO.: 2, SEQ. ID. NO.: 4, or SEQ. ID. NO.: 6;

(2) allowing the cells to take up the protein; and

(3) measuring a reduction in tumor growth.

SEQUENCING

Any nucleic acid specimen, in purified or non-purified form, can be utilized as the starting nucleic acid or acids, providing it contains, or is suspected of containing, the specific

103

nucleic acid sequence containing a polymorphic locus. Thus, the process may amplify, for example, DNA or RNA, including messenger RNA, wherein DNA or RNA may be single stranded or double stranded. In the event that RNA is to be used as a template, enzymes, and/or conditions optimal for reverse transcribing the template to DNA would be utilized. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized.

A mixture–

IN THE CLAIMS:

Please cancel claims 1, 2, 4, 6-15, 22, 23 and 25-27.

Please amend claims 16-21 as follows:

Claim 16 (amended). A method of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with disease, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1 [coni] DNA sequence[, SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5;] selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;

(e) determining the presence of absence of each of the following polymorphic

104

variations in said individual's BRCA1 coding sequence:

[•](i) AGC and AGT at position 2201,

[•](ii) TTG and CTG at position 2430,

[•](iii) CCG and CTG at position 2731,

[•](iv) GAA and GGA at position 3232,

[•](v) AAA and AGA at position 3667,

[•](vi) TCT and TCC at position 4427, and

[•](vii) AGT and GGT at position 4956;

(f) determining any sequence differences between said individual's BRCA1 coding sequences and [SEQ. ID. NO1, SEQ. ID. NO3, OR SEQ. ID. NO5] a BRCA1$^{(ami)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427[,] and 4956[,] is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

Claim 17 (amended). A method [of claim 16] of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with disease, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;

(e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA1 coding sequence:

(i) AGC and AGT at position 2201

(ii) TTG and CTG at position 2430,

(iii) CCG and CTG at position 2731,

(iv) GAA and GGA at position 3232,

(v) AAA and AGA at position 3667,

(vi) TCT and TCC at position 4427, and

(vii) AGT and GGT at position 4956; and

Sub C1 B2 cont

(f) determining any sequence differences between said individual's BRCA1 coding sequences and a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427 and 4956 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence;

wherein[,] codon variations occur at the following frequencies, respectively, in a caucasian population of individuals free of disease:

[•](i) at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](ii) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](iii) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

[•](iv) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](v) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](vi) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

[•](vii) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

Claim 18 (amended). A method according to claims 16 or 17, wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label, a bioluminescent label, a chemiluminescent label, or an enzyme label.

Claim 19 (amended). A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using

105

an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence[, SEQ. ID. NO1, SEQ. ID. NO3, or SEQ. ID. NO5;] selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;

(e) determining any sequence differences between said individual's BRCA1 coding sequences and [SEQ. ID. NO1, SEQ. ID. NO3, OR SEQ. ID. NO5;] a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5 in order to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a base change which is not any one of the following:

[•](i) AGC and AGT at position 2201,

[•](ii) TTG and CTG at position 2430,

[•](iii) CCG and CTG at position 2731,

[•](iv) GAA and GGA at position 3232,

[•](v) AAA and AGA at position 3667,

[•](vi) TCT and TCC at position 4427, and

[•](vii) AGT and GGT at position 4956, is correlated with the potential

of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

Claim 20 (amended). A method of [claim 19] detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising:

(f) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(g) sequencing said amplified DNA fragment by dideoxy sequencing;

(h) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(i) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;

(j) determining any sequence differences between said individual's BRCA1 coding sequences and a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5 in order to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a base change which is not any one of the following:

(i) AGC and AGT at position 2201,

(ii) TTG and CTG at position 2430,

(iii) CCG and CTG at position 2731,

(iv) GAA and GGA at position 3232,

(v) AAA and AGA at position 3667,

(vi) TCT and TCC at position 4427, and

(vii) AGT and GGT at position 4956, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence, wherein[,]codon variations occur at the following frequencies, respectively, in a caucasian population of individuals free of disease:

Sub C2 B6 cont

[•](i) at position 2201, AGC and AGT occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](ii) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](iii) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

[•](iv) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](v) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](vi) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

[•](vii) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

Claim 21 (amended). A method according to claims ~~19~~ or 20, wherein said oligonucleotide primer is labeled with a radiolabel, a fluorescent label, a bioluminescent label, a chemiluminescent label, or an enzyme label.

Claim 24 (amended). A BRCA1 coding sequence according to claims [1] 3 or 5, wherein the codon pairs occur at the following frequencies:

[•](i) at position 2201, AGC and AGT occur at frequencies of about 40%, and from about 55-65%, respectively;

[•](ii) at position 2430, TTG and CTG occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](iii) at position 2731, CCG and CTG occur at frequencies from about 25-35%, and from about 65-75%, respectively;

[•](iv) at position 3232, GAA and GGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](v) at position 3667, AAA and AGA occur at frequencies from about 35-45%, and from about 55-65%, respectively;

[•](vi) at position 4427, TCT and TCC occur at frequencies from about 45-55%, and from about 45-55%, respectively; and

Sub C3 B

[•](vii) at position 4956, AGT and GGT occur at frequencies from about 35-45%, and from about 55-65%, respectively.

IN THE ABSTRACT:

Under 'ABSTRACT OF THE DISCLOSURE' please insert --**CODING SEQUENCES OF THE HUMAN BRCA1 GENE**--.

In line 1, please delete "isolated coding sequences and to the protein sequences they code for.".

In line 2, please delete "sequence" and substitute therefor --sequences--.

REMARKS

Applicants wish to thank the Examiner for the interview held on October 17, 1997. Applicants also thank the Examiner for expediting the Applicants' filing of the Terminal Disclaimer and the Information Disclosure Statement on that date.

Applicants have amended the Specification to delete pages 8-15 and substitute therefor new text in order to present the several embodiments of the invention in conventional syntax rather than merely as an exact copy of the claims. No new matter is added.

Pursuant to the agreement reached during the October 17, 1997 interview, Applicants have amended the claims to closely parallel the claims of parent Application Serial No. 08/598,591, which issued August 5, 1997 as U.S. Patent No. 5,654,155. Specifically, Applicants have cancelled claims 1, 2, 4, 6-15, 22, 23 and 25-27, and amended claims 16-21 and 24. Claim 16 now includes a Markush group consisting all the possible combinations of SEQ. ID. NO.: 1, SEQ. ID. NO.: 3 and SEQ. ID. NO.: 5. Claim 17, formerly dependent on claim 16, is now written in independent form. Accordingly, claim 18 is now dependent on claims 16 or 17. Claim 19 has been amended to include the Markush group as in claim 16.

Claim 20, formerly dependent on claim 19, is now written in independent form. Accordingly, claim 21 is now dependent on claims19 or 20. Finally, claim 24, formerly dependent on claim 1, is now dependent on claims 3 or 5. Additionally, typographical errors and errors of syntax have been corrected in all the claims. No new matter is added.

Finally, Applicants have amended the Abstract to add the title of the invention, to delete an incomplete sentence at the beginning of the Abstract, and to correct a typographical error. No new matter is added.

Respectfully submitted,

Dated: October 23, 1997

Albert P. Halluin Reg. No. 25,227

HOWREY & SIMON
Box 34
1299 Pennsylvania Avenue, N.W.
Washington, D. C. 20004-2402
Tel. (650) 463-8109
Fax (650) 563-8400

HOWREY & SIMON

301 Ravenswood Avenue
Menlo Park, CA 94025-3453
(650) 463-8100

550 South Hope Street, Suite 1400
Los Angeles, CA 90071-2627
(213) 892-1800
(213) 892-2300 FAX

1299 Pennsylvania Avenue, N.W.
Washington, DC 20004-2402
(202) 783-0800
(202) 383-6610 FAX

FACSIMILE COVER SHEET

Facsimile Number: (650) 463-8400

DATE: 10/23/97

TO:

NAME: Examiner Miinh-Tam Davis

COMPANY: USPTO

CITY/STATE: Washington, D.C.

FAX NO: 703/305-7939

FROM:

NAME: Albert P. Halluin

NUMBER OF PAGES, INCLUDING COVER: 21

TELEPHONE CONFIRMATION NUMBER: 703-305-2008

CHARGE NUMBER 5371

☐ **ORIGINAL WILL FOLLOW VIA:**

☐ Regular Mail ☐ Overnight Delivery ☐ Hand Delivery ☐ Other: _____

☒ **ORIGINAL WILL NOT FOLLOW**

SUPPLEMENTAL MESSAGE:

U.S. Serial No. 08/798,691
Filed: February 12, 1997
Our Reference: 05371.0014.999
For: Coding Sequences Of The Human BRCA1 Gene
Inventors: Murphy, et al.
Attached: PRELIMINARY AMENDMENT

If there are any questions or problems with the transmission of this facsimile, please call (650) 463-8100.

Patent
Attorney's Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of)
)
MURPHY, et al.)
)
Application No.: 08/798,691) Group Art Unit: 1809
)
Filed: February 12, 1997) Examiner: Minh-Tam Davis
)
For: CODING SEQUENCES OF)
THE HUMAN BRCA1 GENE)
)

#12

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In accordance with the duty of disclosure as set forth in 37 C.F.R. §1.56, Applicants hereby submit the following information in conformance with 37 C.F.R. §§ 1.97 and 1.98. Pursuant to 37 C.F.R. § 1.98, a copy of each of the documents cited is enclosed.

U.S. PATENT DOCUMENT

Patent No. 5,693,473 Issued: December 2, 1997 Inventors: Shattuck-Eidens et al.

OTHER DOCUMENT

Holt, et al., Gene Therapy Protocol ORDA #: 9603-149 ORDA approved Protocol for BRCA1 Gene Therapy, pp. 1003-1030 (1995).

U.S. Patent No. 5,693,473 corresponds to European Patent Application EP 705,903 which Applicants included in the Information Disclosure Statement previously filed.

These documents are being submitted within 3 months of the filing or entry of the national stage of this application or before the first Office Action on the merits, whichever is later, therefore no fee or certification is required under 37 C.F.R. § 1.97(b).

To assist the Examiner, the listed on the attached form PTO-1449. It is respectfully requested that an Examiner initialled copy of this form be returned to the undersigned.

Respectfully submitted,

HOWREY & SIMON

By: _____
Albert P. Halluin
Registration No. 25,227

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109

Date: December 8, 1997

Patent
Attorney's Docket No. 05371.0014.999

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of	)	
	)	
MURPHY, et al.	)	
	)	
Application No.: 08/798,691	)	Group Art Unit: 1809
	)	
Filed: February 12, 1997	)	Examiner: Minh-Tam Davis
	)	
For: CODING SEQUENCES OF THE HUMAN BRCA1 GENE	)	
	)	

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT TRANSMITTAL LETTER

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Enclosed is an Information Disclosure Statement and accompanying form PTO-1449 for the above-identified patent application.

[X] No additional fee for submission of an IDS is required.

[ ] The fee of $230.00 as set forth in 37 C.F.R. § 1.17(p) is also enclosed.

[ ] A certification under 37 C.F.R. § 1.97(e) is also enclosed.

[ ] A certification under 37 C.F.R. § 1.97(e), a petition requesting consideration of the information disclosure statement, and the petition fee of $130.00 as set forth in 37 C.F.R. § 1.17(i) are also enclosed.

[ ] Charge $_____ to Deposit Account No. 08-3038 for the fee due.

[ ] A check in the amount of $_____ is enclosed for the fee due.

The Commissioner is hereby authorized to charge any appropriate fees under 37 C.F.R. §§1.16, 1.17 and 1.21 that may be required by this paper, and to credit any overpayment, to Deposit Account No. 08-3038. This paper is submitted in duplicate.

Respectfully submitted,

HOWREY & SIMON

By: Albert P. Halluin
Albert P. Halluin
Registration No. 25,227

1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
(650) 463-8109

Date: December 8, 1997

Attorney's Docket No. 05371.0014.999

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

#13/C
(Sweet
12/16/97

In re Patent Application of	)	
	)	
MURPHY, et al.	)	Group Art Unit: 1809
	)	
Application No.: 08/798,691	)	Examiner: Minh-Tam Davis
	)	
Filed: February 12, 1997	)	
	)	
For: **CODING SEQUENCES OF THE HUMAN BRCA1 GENE**	)	

SUPPLEMENTAL PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Prior to the commencement of the Examination of the above-identified application, Applicants herewith submit the following Amendment to the claims and Remarks. This Preliminary Amendments supplements the Preliminary Amendment previously submitted on October 23, 1997.

AMENDMENT

IN THE CLAIMS:

Please amend claims 16, 17, 19, 20 and 24 as follows:

Claim ~~16~~ 2 (twice amended). A method of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with ovarian or breast cancer disease, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the

gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO: 3 and SEQ. ID. NO. 5;] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5;

(e) determining the presence [of] or absence of each of the following polymorphic variations in said individual's BRCA1 coding sequence:

C1 cont

(i) [AGC and AGT] C and T at position 2201,

(ii) [TTG and CTG] T and C at position 2430,

(iii) [CCG and CTG] C and T at position 2731,

(iv) [GAA and GGA] A and G at position 3232,

(v) [AAA and AGA] A and G at position 3667,

(vi) [TCT and TCC] T and C at position 4427, and

(vii) [AGT and GGT] A and G at position 4956;

(f) determining any sequence differences between said individual's BRCA1 coding sequences and a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5,] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ

108

ID NO: 3 and SEQ ID NO: 5, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427 and 4956 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

Claim ~~11~~ 3 (twice amended). A method of identifying individuals having a BRCA1 gene with a BRCA1 coding sequence not associated with ovarian or breast cancer disease, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5;

(e) determining the presence [of] or absence of each of the following polymorphic variations in said individual's BRCA1 coding sequence:

(i) [AGC and AGT] C and T at position 2201,

(ii) [TTG and CTG] T and C at position 2430,

(iii) [CCG and CTG] C and T at position 2731,

(iv) [GAA and GGA] A and G at position 3232,

(v) [AAA and AGA] A and G at position 3667,

(vi) [TCT and TCC] T and C at position 4427, and

(vii) [AGT and GGT] A and G at position 4956; and

(f) determining any sequence differences between said individual's BRCA1 coding sequences and a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5,] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 2201, 2430, 2731, 3232, 3667, 4427 and 4956 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence;

C' cont

wherein codon variations occur at the following frequencies, respectively, in a [caucasian] Caucasian population of individuals ~~free of disease:~~ with no family history of breast or ovarian cancer

D

(i) at position 2201, [AGC and AGT] C and T occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(ii) at position 2430, [TTG and CTG] T and C occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(iii) at position 2731, [CCG and CTG] C and T occur at frequencies from about 25[-] to about 35%, and from about 65[-] to about 75%, respectively;

(iv) at position 3232, [GAA and GGA] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%,

respectively;

(v) at position 3667, [AAA and AGA] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(vi) at position 4427, [TCT and TCC] T and C occur at frequencies from about 45[-] to about 55%, and from about 45[-] to about 55%, respectively; and

(vii) at position 4956, [AGT and GGT] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively.

Claim ~~19~~ 5 (twice amended). A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising:

(a) amplifying a DNA fragment of an individual's BRCA1 coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

(b) sequencing said amplified DNA fragment by dideoxy sequencing;

(c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

(d) comparing the sequence of said amplified DNA fragment to a $BRCA1^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5;

(e) determining any sequence differences between said individual's BRCA1 coding sequences and a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: SEQ. ID. NO.: 1 together with [SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5] SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5 in order to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a base change which is not any one of the following:

(i) [AGC and AGT] C and T at position 2201,

(ii) [TTG and CTG] T and C at position 2430,

(iii) [CCG and CTG] C and T at position 2731,

(iv) [GAA and GGA] A and G at position 3232,

(v) [AAA and AGA] A and G at position 3667,

(vi) [TCT and TCC] T and C at position 4427, and

(vii) [AGT and GGT] A and G at position 4956, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence.

Claim ~~20~~ 6 (twice amended). A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA1 coding sequence, comprising:

[(f)] (a) amplifying a DNA fragment of an individual's BRCA1 [coding sequence] coding sequence using an oligonucleotide primer which specifically hybridizes to sequences within the gene;

[(g)] (b) sequencing said amplified DNA fragment by dideoxy sequencing;

[(h)] (c) repeating steps (a) and (b) until said individual's BRCA1 coding sequence is completely sequenced;

[(i)] (d) comparing the sequence of said amplified DNA fragment to a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5;] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5;

Cont

[(j)] (e) determining any sequence differences between said individual's BRCA1 coding sequences and a BRCA1$^{(omi)}$ DNA sequence selected from the group consisting of: [SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 3, SEQ. ID. NO.: 1 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 together with SEQ. ID. NO.: 5, SEQ. ID. NO.: 3 and SEQ. ID. NO. 5] SEQ ID NO: 1 together with SEQ ID NO: 3, SEQ ID NO: 1 together with SEQ ID NO: 5, SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5, SEQ ID NO: 3 and SEQ ID NO: 5 in order to determine the presence or absence of base changes in said individual's BRCA1 coding sequence wherein a base change which is not any one of the following:

(i) [AGC and AGT] C and T at position 2201,

(ii) [TTG and CTG] T and C at position 2430,

(iii) [CCG and CTG] C and T at position 2731,

(iv) [GAA and GGA] A and G at position 3232,

(v) [AAA and AGA] A and G at position 3667,

(vi) [TCT and TCC] T and C at position 4427, and

113

(vii) [AGT and GGT] A and G at position 4956, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA1 mutation in the BRCA1 coding sequence, wherein codon variations occur at the following frequencies, respectively, in a [caucasian] Caucasian population of individuals ~~free of disease~~ with no family history of breast or ovarian cancer:

(i) at position 2201, [AGC and AGT] C and T occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(ii) at position 2430, [TTG and CTG] T and C occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(iii) at position 2731, [CCG and CTG] C and T occur at frequencies from about 25[-] to about 35%, and from about 65[-] to about 75%, respectively;

(iv) at position 3232, [GAA and GGA] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(v) at position 3667, [AAA and AGA] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(vi) at position 4427, [TCT and TCC] T and C occur at frequencies from about 45[-] to about 55%, and from about 45[-] to about 55%, respectively; and

(vii) at position 4956, [AGT and GGT] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively.

Claim 24 (twice amended). A BRCA1 coding sequence according to claims 3 or 5, wherein the codon pairs occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:

(i) at position 2201, [AGC and AGT] C and T occur at frequencies of about 40%, and from about 55[-] to about 65%, respectively;

(ii) at position 2430, [TTG and CTG] T and C occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(iii) at position 2731, [CCG and CTG] C and T occur at frequencies from about 25[-] to about 35%, and from about 65[-] to about 75%, respectively;

(iv) at position 3232, [GAA and GGA] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(v) at position 3667, [AAA and AGA] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively;

(vi) at position 4427, [TCT and TCC] T and C occur at frequencies from about 45[-] to about 55%, and from about 45[-] to about 55%, respectively; and

(vii) at position 4956, [AGT and GGT] A and G occur at frequencies from about 35[-] to about 45%, and from about 55[-] to about 65%, respectively.

IN THE ABSTRACT:

Please delete the current abstract and replace with the following substitute abstract:

ABSTRACT OF THE DISCLOSURE

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

This invention is directed to three coding sequences of the BRCA1 gene. The three coding sequences, BRCA1$^{(omi1)}$, BRCA1$^{(omi2)}$ and BRCA1$^{(omi3)}$ as well as their frequencies of occurrence are provided together with the protein sequences they code for. Another aspect of this invention is a method of determining the consensus sequence for any gene. Another aspect of the invention is a method of identifying an individual having an increased genetic susceptibility to breast or ovarian cancer because they have inherited a causative mutation in their BRCA1 gene. This invention is also related to a method of performing gene therapy with any of the isolated BRCA1 coding sequences.

REMARKS

Pursuant to Applicants' discussions with the Examiner, the foregoing Supplemental Preliminary Amendment is submitted in order to correct minor typographical errors, to recite the specific nucleotides comprising the polymorphisms rather than the 3 base-codons containing these polymorphisms, and to limit the polymorphic variation frequencies recited in claim 24 to those of "a caucasian population of individuals free of disease" (as in claims 17 and 20). In addition, claims 16, 17, 19 and 20 were amended to include the limitation of all three sequences together, i.e., "SEQ ID NO: 1 together with SEQ ID NO: 3 together with SEQ ID NO: 5";

Applicants' failure to include this final permutation in the previous Preliminary Amendment was an inadvertent error. No new matter is added by any of the amendments.

Respectfully submitted,

Dated: December 5, 1997

Albert P. Halluin Reg. No. 25,227

HOWREY & SIMON
Box 34
1299 Pennsylvania Avenue, N.W.
Washington, D. C. 20004-2402
Tel: (650) 463-8109
Fax: (650) 463-8400

Interview Summary

Application No.	Applicant(s)	
08/798,691	Murphy et al	
Examiner	Group Art Unit	
Sheela J. Huff	1642	

All participants (applicant, applicant's representative, PTO personnel):

(1) *Sheela J. Huff* (3) _____

(2) *Albert Halliun* (4) _____

Date of Interview *Feb 11, 1998*

Type: ☒ Telephonic ☐ Personal (copy is given to ☐ applicant ☐ applicant's representative).

Exhibit shown or demonstration conducted: ☐ Yes ☒ No. If yes, brief description:

Agreement ☒ was reached. ☐ was not reached.

Claim(s) discussed: *all pending claims.*

Identification of prior art discussed:
none

Description of the general nature of what was agreed to if an agreement was reached, or any other comments:
The attorney agreed to the changes in the Examiner's amendment.

(A fuller description, if necessary, and a copy of the amendments, if available, which the examiner agreed would render the claims allowable must be attached. Also, where no copy of the amendents which would render the claims allowable is available, a summary thereof must be attached.)

1. ☒ It is not necessary for applicant to provide a separate record of the substance of the interview.

Unless the paragraph above has been checked to indicate to the contrary, A FORMAL WRITTEN RESPONSE TO THE LAST OFFICE ACTION IS NOT WAIVED AND MUST INCLUDE THE SUBSTANCE OF THE INTERVIEW. (See MPEP Section 713.04). If a response to the last Office action has already been filed, APPLICANT IS GIVEN ONE MONTH FROM THIS INTERVIEW DATE TO FILE A STATEMENT OF THE SUBSTANCE OF THE INTERVIEW.

2. ☐ Since the Examiner's interview summary above (including any attachments) reflects a complete response to each of the objections, rejections and requirements that may be present in the last Office action, and since the claims are now allowable, this completed form is considered to fulfill the response requirements of the last Office action. Applicant is not relieved from providing a separate record of the interview unless box 1 above is also checked.

Examiner Note: You must sign and stamp this form unless it is an attachment to a signed Office action.

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

SERIAL NUMBER	FILING DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NO.
08/798,691	02/12/97	MURPHY	P PA-0054CIP

HM31/0217

ALBERT P. HALLUIN
HOWERY & SIMON
1299 PENNSYLVANIA AVE., N.W.
WASHINGTON DC 20004

EXAMINER	
ART UNIT 1642	PAPER NUMBER

DATE MAILED: 02/17/98

Please find below a communication from the EXAMINER in charge of this application.

Commissioner of Patents

1. PATENT APPLICATION FILE COPY

Notice of Allowability

Application No.	Applicant(s)
08/798,691	Murphy et al
Examiner	**Group Art Unit**
Sheela J. Huff	1642

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance and Issue Fee Due or other appropriate communication will be mailed in due course.

☒ This communication is responsive to *interviews on 12/4/97, 12/11/97, 12/12/97 and 2/10/98*.

☒ The allowed claim(s) is/are *5, 16-21 and 24 renumbered as 1-8 respectively.*.

☐ The drawings filed on _____ are acceptable.

☐ Acknowledgement is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d).

☐ All ☐ Some* ☐ None of the CERTIFIED copies of the priority documents have been

☐ received.

☐ received in Application No. (Series Code/Serial Number) _____.

☐ received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

*Certified copies not received: _____.

☐ Acknowledgement is made of a claim for domestic priority under 35 U.S.C. § 119(e).

A SHORTENED STATUTORY PERIOD FOR RESPONSE to comply with the requirements noted below is set to EXPIRE **THREE MONTHS** FROM THE "DATE MAILED" of this Office action. Failure to timely comply will result in ABANDONMENT of this application. Extensions of time may be obtained under the provisions of 37 CFR 1.136(a).

☐ Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL APPLICATION, PTO-152, which discloses that the oath or declaration is deficient. A SUBSTITUTE OATH OR DECLARATION IS REQUIRED.

☒ Applicant MUST submit NEW FORMAL DRAWINGS

☐ because the originally filed drawings were declared by applicant to be informal.

☒ including changes required by the Notice of Draftsperson's Patent Drawing Review, PTO-948, attached hereto or to Paper No. _____.

☐ including changes required by the proposed drawing correction filed on _____, which has been approved by the examiner.

☐ including changes required by the attached Examiner's Amendment/Comment.

Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the reverse side of the drawings. The drawings should be filed as a separate paper with a transmittal lettter addressed to the Official Draftsperson.

☐ Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Any response to this letter should include, in the upper right hand corner, the APPLICATION NUMBER (SERIES CODE/SERIAL NUMBER). If applicant has received a Notice of Allowance and Issue Fee Due, the ISSUE BATCH NUMBER and DATE of the NOTICE OF ALLOWANCE should also be included.

Attachment(s)

☒ Notice of References Cited, PTO-892

☒ Information Disclosure Statement(s), PTO-1449, Paper No(s). *7, 9 and 12*

☒ Notice of Draftsperson's Patent Drawing Review, PTO-948

☐ Notice of Informal Patent Application, PTO-152

☒ Interview Summary, PTO-413 (2 sheets)

☒ Examiner's Amendment/Comment

☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material

☒ Examiner's Statement of Reasons for Allowance

Sheela Huff

SHEELA HUFF
PRIMARY EXAMINER

EXAMINER'S AMENDMENT 1D

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Albert Halluin on December 12,1997 and February 10, 1998.

2. The application has been amended as follows:

a. Claim 3 has been cancelled.

b. Claim 4 has been amended as follows:

8. A BRCA1 coding sequence according to claim[s or] 5 wherein the codon pairs occur at the following frequencies, [respectively,] in a Caucasian population of individuals with no family history of breast or ovarian cancer [free of disease]:

(i) at position 2201, C and T occur at frequencies of about 40%, and from about 55% to about 65%, respectively;

(ii) at position 2430, T and C occur at frequencies of about 35 to about 45%, and from about 55% to about 65%, respectively;

(iii) at position 2731, C and T occur at frequencies of about 25 to about 35%, and from about 65% to about 75%, respectively;

(iv) at position 3232, A and G occur at frequencies of about 35 to about 45%, and from about 55% to about 65%, respectively;

(v) at position 3667, A and G occur at frequencies of about 35 to about 45%, and from about 55% to about 65%, respectively;

(vi) at position 4427, T and C occur at frequencies of about 45 to about 55%, and from about 45% to about 55%, respectively; and

(ii) at position 4956, A and G occur at frequencies of about 35 to about 45%, and from about 55% to about 65%, respectively.

c. In claim 16, line 2, --ovarian or breast cancer-- has been inserted before "disease".

d. In claim 17, line 2, --ovarian or breast cancer-- has been inserted before "disease".

e. In claim 17, line 38, --with no family history of breast or ovarian cancer-- has been inserted after "individuals" and "free of disease" has been deleted.

f. In claim 20, line 39, --with no family history of breast or ovarian cancer-- has been inserted after "individuals" and "free of disease" has been deleted.

3. The following is an examiner's statement of reasons for allowance:

Sequence ID No 5 is free from the art of record. The seqeunce of claim 24 is free from the art of record because the prior art netiher teaches nor suggests all 7 specific polymorphism sites in the same seqeunce.

The method claims are free from the art of record because it is not predictable to combine 7 specific polymorphic sites into a consensus sequence, and to test only these 7 specific polymorphic sites for rapid detection of individuals not having breast cancer.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

4. The Group and/or Art Unit location of your application in the PTO has changed. To aid in correlating any papers for this application, all further correspondence regarding this application should be directed to Technology Center 1600, Group 1640, Group Art Unit 1642.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Sheela J. Huff whose telephone number is (703) 305-7866. The examiner can normally be reached on Monday-Thursday from 6:30am to 3:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Lila Feisee, can be reached on (703)308-2731. The FAX phone number for this Group is (703)308-4242.

Communications via Internet e-mail regarding this application, other than those under 35 U.S.C. 132 or which otherwise require a signature, may be used by the applicant and should be addressed to [**lila.feisee@uspto.gov**].

All Internet e-mail communications will be made of record in the application file. **PTO employees do not engage in Internet communications where there exists a possibility that sensitive information could be identified or exchanged unless the record includes a properly signed express waiver of the confidentiality requirements of 35 U.S.C. 122.** This is more clearly set forth in the Interim Internet Usage Policy published in the Official Gazette of the Patent and Trademark on February 25, 1997 at 1195 OG 89.

Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist whose telephone number is (703) 308-0196.

Sheela J. Huff
February 11, 1998

Sheela J. Huff
Primary Examiner

Notice of References Cited

Application No.	Applicant(s)		
08/798,691	Murphy et al		
Examiner		Group Art Unit	Page 1 of 1
Sheela J. Huff		1642	

U.S. PATENT DOCUMENTS

		DOCUMENT NO.	DATE	NAME	CLASS	SUBCLASS
	A					
	B					
	C					
	D					
	E					
	F					
	G					
	H					
	I					
	J					
	K					
	L					
	M					

FOREIGN PATENT DOCUMENTS

		DOCUMENT NO.	DATE	COUNTRY	NAME	CLASS	SUBCLASS
	N						
	O						
	P						
	Q						
	R						
	S						
	T						

NON-PATENT DOCUMENTS

		DOCUMENT (Including Author, Title, Source, and Pertinent Pages)	DATE
	U	Database Genbank, Accession No. U14680	1994
	V		
	W		
	X		

FORM PTO 948 (REV. 01-97) U.S. DEPARTMENT OF COMMERCE-Patent and Trademark Office Application No. 798691.

NOTICE OF DRAFTPERSON'S PATENT DRAWING REVIEW

The drawing filed (insert date) 2/12/97. are:

A. _____ not objected to by the Draftperson under 37 CFR 1.84 or 1.152.

B. _____ objected to by the Draftperson under 37 CFR 1.84 or 1.152 as indicated below. The Examiner will require submission of new, corrected drawings whe necessary. Corrected drawings must be submitted according to the instructions on the back of this notice.

1. DRAWINGS. 37 CFR 1.84(a): Acceptable categories of drawings:
Black ink. Color.
_____ Color drawing are not acceptable until petition is granted.
Fig.(s) _____
_____ Pencil and non black ink is not permitted. Fig(s)_____

2. PHOTOGRAPHS. 37 CFR 1.84(b)
_____ Photographs are not acceptable until petition is granted,
_____ 3 full-tone sets are required. Fig(s)_____
_____ Photographs not properly mounted (must brystol board or photographic double-weight paper). Fig(s)_____
_____ Poor quailty (half-tone). Fig(s)_____

3. TYPE OF PAPER. 37 CFR 1.84(e)
_____ Paper not flexible, strong, white and durable.
Fig.(s)_____
_____ Erasures, alterations, overwritings, interlineations, folds, copy machine marks not acceptable. (too thin)
_____ Mylar, vellum paper is not acceptable (too thin).
Fig(s)_____

4. SIZE OF PAPER. 37 CFR 1.84(F): Acceptable sizes:
_____ 21.0 cm by 29.7 cm (DIN size A4)
_____ 21.6 cm by 27.9 cm (8 1/2 x 11 inches)
_____ All drawings sheets not the same size.
Sheet(s)_____

5. MARGINS. 37 CFR 18.4(g): Acceptable margins:
Top 2.5 cm Left 2.5 cm Right 1.5 cm Bottom 1.0 cm
SIZE: A4 Size
Top 2.5 cm Left 2.5 cm Right 1.5 cm Bottom 1.0 cm
SIZE: 8 1/2 x 11
_____ Margins not acceptable. Fig(s)_____
_____ Top (T) _____ Left (L)
_____ Right (R) _____ Bottom (B)

6. VIEWS. CFR 1.84(h)
REMINDER: Specification may require revision to correspond to drawing changes.
_____ Views connected by projection lines or lead lines.
Fig.(s)_____
Partial views. 37 CFR 1.84(h)(2)
_____ Brackets needed to show figure as one entity.
Fig.(s)_____
_____ Views not labeled separately or properly.
Fig.(s)_____
_____ Enlarged view not labeled separately or properly.
Fig.(s)_____

7. SECTIONAL VIEWS. 37 CFR 1.84(h)(3)
_____ Hatching not indicated for sectional portions of an object.
Fig.(s)_____
_____ Sectional designation should be noted with Arabic or Roman numbers. Fig.(s)_____

8. ARRANGEMENT OF VIEWS. 37 CFR 1.84(i)
_____ Words do not appear on a horizontal, left-to-right fashion when page is either upright or turned, so that the top becomes the right side, except for graphs. Fig.(s)_____
_____ Views not on the same plane on drawing sheet. Fig.(s)_____

9. SCALE. 37 CFR 1.84(k)
_____ Scale not large enough to show mechansim with crowding when drawing is reduced in size to two-thirds in reproduction.
Fig.(s)_____

10. CHARACTER OF LINES, NUMBERS, & LETTERS. 37 CFR 1.84(l)
✓ Lines, numbers & letters not uniformly thick and well defined, clean, durable and black (poor line quality).
Fig.(s) 1-2.

11. SHADING. 37 CFR 1.84(m)
_____ Solid black areas pale. Fig.(s)_____
_____ Solid black shading not permitted. Fig.(s)_____
_____ Shade lines, pale, rough and blurred. Fig.(s)

12. NUMBERS, LETTERS, & REFERENCE CHARACTERS. 37 CFR 1.48(p)
_____ Numbers and reference characters not plain and legible.
Fig.(s)_____
✓ Figure legends are poor. Fig.(s) 12 placed incorrectly
_____ Numbers and reference characters not oriented in the same direction as the view. 37 CFR 1.84(p)(3) Fig.(s)_____
_____ Engligh alphabet not used. 37 CFR 1.84(p)(3) Fig.(s)_____
_____ Numbers, letters and reference characters must be at least .32 cm (1/8 inch) in height. 37 CFR 1.84(p)(3) Fig.(s)_____

13. LEAD LINES. 37 CFR 1.84(q)
_____ Lead lines cross each other. Fig.(s)_____
_____ Lead lines missing. Fig.(s)_____

14. NUMBERING OF SHEETS OF DRAWINGS. 37 CFR 1.48(t)
_____ Sheets not numbered consecutively, and in Ababic numerals beginning with number 1. Fig.(s)_____

15. NUMBERING OF VIEWS. 37 CFR 1.84(u)
_____ Views not numbered consecutively, and in Abrabic numerals, beginning with number 1. Fig.(s)_____

16. CORRECTIONS. 37 CFR 1.84(w)
_____ Corrections not made from PTO-948 dated _____

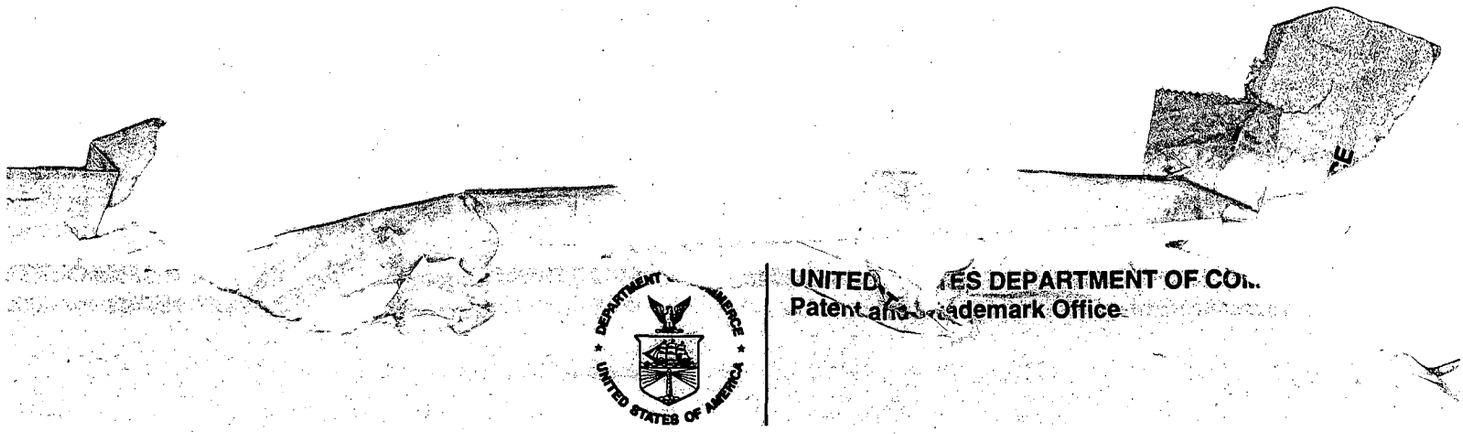
17. DESIGN DRAWINGS. 37 CFR 1.152
_____ Surface shading shown not appropriate. Fig.(s)_____
_____ Solid black shading not used for color contrast.
Fig.(s)_____

COMMENTS

REVIEWER TANK DATE 8/13/97 TELEPHONE NO. 703 305 8404

ATTACHMENT TO PAPER NO. 105

PTO COPY



UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office

NOTICE OF ALLOWANCE AND ISSUE FEE DUE

HM31/0213

ALBERT P. HALLUIN
HOWERY & SIMON
1299 PENNSYLVANIA AVE., N.W.
WASHINGTON DC 20004

APPLICATION NO.	FILING DATE	TOTAL CLAIMS	EXAMINER AND GROUP ART UNIT	DATE MAILED
First Named Applicant 08/798,691	02/12/97	008	HUFF, S 1642	02/17/98

TITLE OF INVENTION: MURPHY, PATRICIA D.

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

ATTY'S DOCKET NO.	CLASS-SUBCLASS	BATCH NO.	APPLN. TYPE	SMALL ENTITY	FEE DUE	DATE DUE
1 PA-0054CIP	435-006.000	A52	UTILITY	YES	$660.00	05/18/98

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED.

THE ISSUE FEE MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED.

HOW TO RESPOND TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.
If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

A. If the status is changed, pay twice the amount of the FEE DUE shown above and notify the Patent and Trademark Office of the change in status, or
B. If the status is the same, pay the FEE DUE shown above.

If the SMALL ENTITY is shown as NO:

A. Pay FEE DUE shown above, or

B. File verified statement of Small Entity Status before, or with, payment of 1/2 the FEE DUE shown above.

II. Part B-Issue Fee Transmittal should be completed and returned to the Patent and Trademark Office (PTO) with your ISSUE FEE. Even if the ISSUE FEE has already been paid by charge to deposit account, Part B Issue Fee Transmittal should be completed and returned. If you are charging the ISSUE FEE to your deposit account, section "4b" of Part B-Issue Fee Transmittal should be completed and an extra copy of the form should be submitted.

III. All communications regarding this application must give application number and batch number.
Please direct all communications prior to issuance to Box ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PATENT AND TRADEMARK OFFICE COPY

PART B—ISSUE FEE TRANSMITTAL

Complete and mail this form, together with applicable fees, to: **Box ISSUE FEE**
Assistant Commissioner for Patents
Washington, D.C. 20231

MAILING INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE. Blocks 1 through 4 should be completed where appropriate. All further correspondence including the Issue Fee Receipt, the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Legibly mark-up with any corrections or use Block 1)

HM31/0213

ALBERT P. HALLUIN
HOWERY & SIMON
1299 PENNSYLVANIA AVE., N.W.
WASHINGTON DC 20004

Note: The certificate of mailing below can only be used for domestic mailings of the Issue Fee Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing.

Certificate of Mailing

I hereby certify that this Issue Fee Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Box Issue Fee address above on the date indicated below.

(Depositor's name)

(Signature)

(Date)

APPLICATION NO.	FILING DATE	TOTAL CLAIMS	EXAMINER AND GROUP ART UNIT	DATE MAILED
First Named Applicant 08/798,691	02/12/97	008	HUFF, S 1642	02/17/98

TITLE OF INVENTION: MURPHY, PATRICIA D.

CODING SEQUENCES OF THE HUMAN BRCA1 GENE

ATTY'S DOCKET NO.	CLASS-SUBCLASS	BATCH NO.	APPLN. TYPE	SMALL ENTITY	FEE DUE	DATE DUE
1 PA-0054CIP	435-006.000	452	UTILITY	YES	$660.00	05/18/98

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363). Use of PTO form(s) and Customer Number are recommended, but not required.

☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47) attached.

2. For printing on the patent front page, list (1) the names of up to 3 registered patent attorneys or agents OR, alternatively, (2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 Albert P. Halluin
2 R. Thomas Gallegos
3 HOWREY & SIMON

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)
PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. Inclusion of assignee data is only appropriate when an assignment has been previously submitted to the PTO or is being submitted under separate cover. Completion of this form is NOT a subsititue for filing an assignment.

(A) NAME OF ASSIGNEE: OncorMed, Inc.

(B) RESIDENCE: (CITY & STATE OR COUNTRY) Gaithersburg, Maryland

Please check the appropriate assignee category indicated below (will not be printed on the patent)

☐ individual ☒ corporation or other private group entity ☐ government

4a. The following fees are enclosed (make check payable to Commissioner of Patents and Trademarks):

☒ Issue Fee
☒ Advance Order - # of Copies 10

4b. The following fees or deficiency in these fees should be charged to:

DEPOSIT ACCOUNT NUMBER _____
(ENCLOSE AN EXTRA COPY OF THIS FORM)

☐ Issue Fee
☐ Advance Order - # of Copies _____

The COMMISSIONER OF PATENTS AND TRADEMARKS IS requested to apply the Issue Fee to the application identified above.

(Authorized Signature) Albert P. Halluin (Date) 2/20/98

NOTE; The Issue Fee will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the Patent and Trademark Office.

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending on the needs of the individual case. Any comments on the amount of time required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, D.C. 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND FEES AND THIS FORM TO: Box Issue Fee, Assistant Commissioner for Patents, Washington D.C. 20231

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

03/04/1998 GPOINDEX 00000005 08798691
01 FC:251 660.00 OP
02 FC:252 30.00 OP

GeneDX 1023, pg. 290

TRANSMIT THIS FORM WITH FEE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:	Attention: **Drafting Branch**
Patricia D. MURPHY, et al.	Art Unit: 1642
Application No. 08/798,691	Examiner: Huff, S.
Filed: February 12, 1997	Attorney's Docket No: 05371.0014.999
For: CODING SEQUENCES OF THE HUMAN BRCA1 GENE	Batch: A52

SUBMISSION OF FORMAL DRAWINGS

BOX ISSUE FEE
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Enclosed please find one (1) sheets of formal drawings for review by the Patent and Trademark Office in connection with the Notice of Allowance mailed February 17, 1998. Also, attached is a copy of the Notice of Draftperson's Patent Drawing Review (Form PTO 948) noting the changes required in Paragraphs 10 and 12 . In this Notice, it is indicated that there are two (2) figures, when in fact there is only one (1).

Should the enclosed drawings require changes, it is respectfully requested that the Patent and Trademark Office notify the undersigned attorney of same.

Respectfully submitted,

Date: February 26, 1998

Albert P. Halluin (Reg. No. 25,227)

HOWREY & SIMON
Box No. 34
1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004-2402
(650) 463-8100

FORM PTO 948 (REV. 01-97) U.S. DEPARTMENT OF COMMERCE-Patent and Trademark Office Application No. 798691

NOTICE OF DRAFTPERSON'S PATENT DRAWING REVIEW

The drawing filed (insert date) 2/12/97 are:

A. _____ not objected to by the Draftperson under 37 CFR 1.84 or 1.152.

B. _____ objected to by the Draftperson under 37 CFR 1.84 or 1.152 as indicated below. The Examiner will require submission of new, corrected drawings when necessary. Corrected drawings must be submitted according to the instructions on the back of this notice.

1. DRAWINGS. 37 CFR 1.84(a): Acceptable categories of drawings:
Black ink. Color.
_____ Color drawing are not acceptable until petition is granted.
Fig.(s)_____
_____ Pencil and non black ink is not permitted. Fig(s)_____

2. PHOTOGRAPHS. 37 CFR 1.84(b)
_____ Photographs are not acceptable until petition is granted.
_____ 3 full-tone sets are required. Fig(s)_____
_____ Photographs not properly mounted (must brystol board or photographic double-weight paper). Fig(s)_____
_____ Poor quailty (half-tone). Fig(s)_____

3. TYPE OF PAPER. 37 CFR 1.84(e)
_____ Paper not flexible, strong, white and durable.
Fig.(s)_____
_____ Erasures, alterations, overwritings, interlineations, folds, copy machine marks not acceptable. (too thin)
_____ Mylar, vellum paper is not acceptable (too thin).
Fig(s)_____

4. SIZE OF PAPER. 37 CFR 1.84(F): Acceptable sizes:
_____ 21.0 cm by 29.7 cm (DIN size A4)
_____ 21.6 cm by 27.9 cm (8 1/2 x 11 inches)
_____ All drawings sheets not the same size.
Sheet(s)_____

5. MARGINS. 37 CFR 18.4(g): Acceptable margins:
Top 2.5 cm Left 2.5 cm Right 1.5 cm Bottom 1.0 cm
SIZE: A4 Size
Top 2.5 cm Left 2.5 cm Right 1.5 cm Bottom 1.0 cm
SIZE: 8 1/2 x 11
_____ Margins not acceptable. Fig(s)_____
_____ Top (T) _____ Left (L)
_____ Right (R) _____ Bottom (B)

6. VIEWS. CFR 1.84(h)
REMINDER: Specification may require revision to correspond to drawing changes.
_____ Views connected by projection lines or lead lines.
Fig.(s)_____
Partial views. 37 CFR 1.84(h)(2)
_____ Brackets needed to show figure as one entity.
Fig.(s)_____
_____ Views not labeled separately or properly.
Fig.(s)_____
_____ Enlarged view not labeled separately or properly.
Fig.(s)_____

7. SECTIONAL VIEWS. 37 CFR 1.84(h)(3)
_____ Hatching not indicated for sectional portions of an object.
Fig.(s)_____
_____ Sectional designation should be noted with Arabic or Roman numbers. Fig.(s)_____

8. ARRANGEMENT OF VIEWS. 37 CFR 1.84(i)
_____ Words do not appear on a horizontal, left-to-right fashion when page is either upright or turned. so that the top becomes the right side, except for graphs. Fig.(s)_____
_____ Views not on the same plane on drawing sheet. Fig.(s)_____

9. SCALE. 37 CFR 1.84(k)
_____ Scale not large enough to show mechansim with crowding when drawing is reduced in size to two-thirds in reproduction.
Fig.(s)_____

10. CHARACTER OF LINES, NUMBERS, & LETTERS. 37 CFR 1.84(l)
✓ Lines, numbers & letters not uniformly thick and well defined, clean, durable and black (poor line quality).
Fig.(s) 1-2

11. SHADING. 37 CFR 1.84(m)
_____ Solid black areas pale. Fig.(s)_____
_____ Solid black shading not permitted. Fig.(s)_____
_____ Shade lines, pale, rough and blurred. Fig.(s)

12. NUMBERS, LETTERS, & REFERENCE CHARACTERS. 37 CFR 1.48(p)
_____ Numbers and reference characters not plain and legible.
Fig.(s)_____
✓ Figure legends are poor. Fig.(s) 12 placed incorrectly
_____ Numbers and reference characters not oriented in the same direction as the view. 37 CFR 1.84(p)(3) Fig.(s)_____
_____ Engligh alphabet not used. 37 CFR 1.84(p)(3) Fig.(s)_____
_____ Numbers, letters and reference characters must be at least 32 cm (1/8 inch) in height. 37 CFR 1.84(p)(3) Fig.(s)_____

13. LEAD LINES. 37 CFR 1.84(q)
_____ Lead lines cross each other. Fig.(s)_____
_____ Lead lines missing. Fig.(s)_____

14. NUMBERING OF SHEETS OF DRAWINGS. 37 CFR 1.48(t)
_____ Sheets not numbered consecutively, and in Ababic numerals beginning with number 1. Fig.(s)_____

15. NUMBERING OF VIEWS. 37 CFR 1.84(u)
_____ Views not numbered consecutively, and in Abrabic numerals, beginning with number 1. Fig.(s)_____

16. CORRECTIONS. 37 CFR 1.84(w)
_____ Corrections not made from PTO-948 dated_____

_____ Surface shading shown not appropriate for Fig.(s)_____

_____ Solid black shading not used for color contrast.

Fig.(s)_____ ____

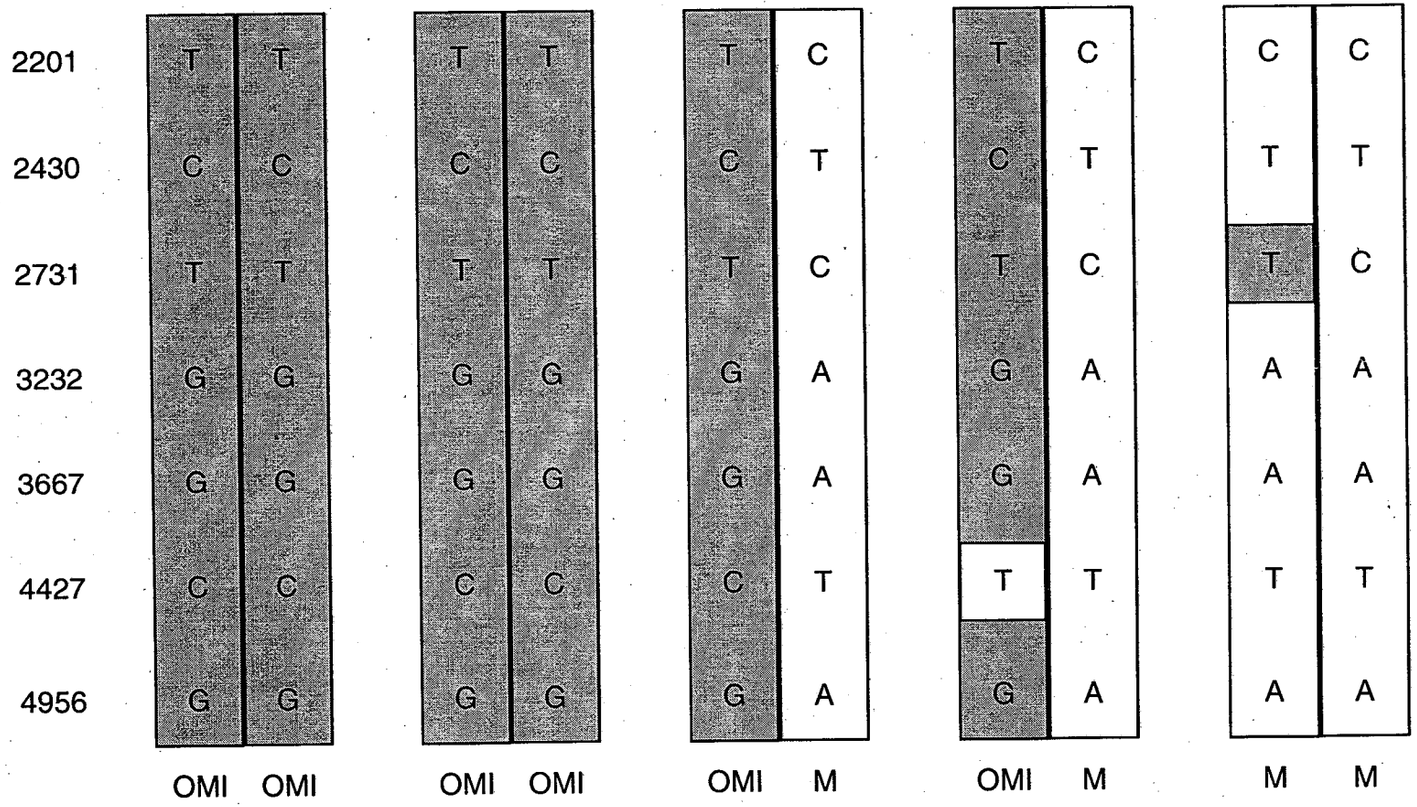
COMMENTS

REVIEWER TAWR DATE 8/13/97 TELEPHONE NO. 703 305 8404

ATTACHMENT TO PAPER NO. #15

APPLICANT'S COPY

FIGURE 1



PTO UTILITY GRANT
Paper Number 17

The United States of America

The Commissioner of Patents and Trademarks

Has received an application for a patent for a new and useful invention. The title and description of the invention are enclosed. The requirements of law have been complied with, and it has been determined that a patent on the invention shall be granted under the law.

Therefore, this

United States Patent

Grants to the person(s) having title to this patent the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States of America or importing the invention into the United States of America for the term set forth below, subject to the payment of maintenance fees as provided by law.

If this application was filed prior to June 8, 1995, the term of this patent is the longer of seventeen years from the date of grant of this patent or twenty years from the earliest effective U.S. filing date of the application, subject to any statutory extension.

If this application was filed on or after June 8, 1995, the term of this patent is twenty years from the U.S. filing date, subject to an statutory extension. If the application contains a specific reference to an earlier filed application or applications under 35 U.S.C. 120, 121 or 365(c), the term of the patent is twenty years from the date on which the earliest application was filed, subject to any statutory extension.

Bruce Lehman
Commissioner of Patents and Trademarks

Sandra J Morton
Attest

Form PTO-1584 (Rev. 2/97)

UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office

ASSISTANT SECRETARY AND COMMISSIONER
OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

CHANGE OF ADDRESS/POWER OF ATTORNEY

LOCATION 9200 SERIAL NUMBER 08798691 PATENT NUMBER 5750400

THE CORRESPONDENCE ADDRESS HAS BEEN CHANGED TO CUSTOMER # 22930

THE PRACTITIONERS OF RECORD HAVE BEEN CHANGED TO CUSTOMER # 22930

THE FEE ADDRESS HAS BEEN CHANGED TO CUSTOMER # 22930

ON 02/04/00 THE ADDRESS OF RECORD FOR CUSTOMER NUMBER 22930 IS:

HOWREY & SIMON
1299 PENNSYLVANIA AVENUE NW
BOX 34
WASHINGTON DC 20004

AND THE PRACTITIONERS OF RECORD FOR CUSTOMER NUMBER 22930 ARE:

21072	25101	25227	26599	29209	30104	30186	31036	32040	32212
32680	34206	34301	35944	36253	36576	37969	38069	38424	39604
39623	39772	39839	39841	39948	39965	40205	40509	40659	40750
40772	41114	41131	41408	41410	41612	41698	42012	42714	42759
42887	43610	43741	43881	44400	45045				

PTO INSTRUCTIONS: PLEASE TAKE THE FOLLOWING ACTION WHEN THE CORRESPONDENCE ADDRESS HAS BEEN CHANGED TO CUSTOMER NUMBER: RECORD, ON THE NEXT AVAILABLE CONTENTS LINE OF THE FILE JACKET, 'ADDRESS CHANGE TO CUSTOMER NUMBER'. LINE THROUGH THE OLD ADDRESS ON THE FILE JACKET LABEL AND ENTER ONLY THE 'CUSTOMER NUMBER' AS THE NEW ADDRESS. FILE THIS LETTER IN THE FILE JACKET. WHEN ABOVE CHANGES ARE ONLY TO FEE ADDRESS AND/OR PRACTITIONERS OF RECORD, FILE LETTER IN THE FILE JACKET.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Patent of: **Patricia D. Murphy *et al.***

Patent No. **5,750,400**

Issued: **May 21, 1998**

Application **08/798,691**

Filed: **February 12, 1997**

For: **Coding sequences of the human BRCA1 gene**

Mail Stop Certificate of Corrections Branch
U.S. Patent and Trademark Office
2011 South Clark Place
Customer Window
Crystal Plaza Two, Lobby, Room 1B03
Arlington, VA 22202

REQUEST FOR CERTIFICATE OF CORRECTION

This is a request for the issuance of a Certificate of Correction under 37 C.F.R. 1.323 in the above-identified patent. Two (2) copies of form PTO-1050 are enclosed. Applicants respectfully submit that the errors, misidentification of nucleic acid and protein sequences in the Sequence Listing of the patent, are errors of a clerical nature and that their correction does not involve changes which would constitute new matter or require reexamination. Accordingly, correction under 37 C.F.R. 1.323 is requested. An authorization to charge the deposit account of the undersigned, Deposit Account No. 50-0310 in the amount of $100.00, as set forth under 37 CFR 1.20(a), is given herewith.

This request is filed to correct sequence identifiers that were inadvertently applied to the wrong sequences. Specifically, SEQ ID NO: 3 and 4 of the Sequence Listing correspond to a gene known as BRCA1$^{(omi3)}$ but these sequences were misidentified in the specification as SEQ ID NO: 5 and 6, respectively. SEQ ID NO: 5 and 6 of the Sequence Listing correspond to a gene known as BRCA1$^{(omi2)}$ but these sequences were misidentified in the specification as SEQ ID NO: 3 and 4, respectively. The sequences corresponding to SEQ ID NO: 3 to 6 of the Sequence Listing are correctly identified in the attached Form PTO-1050, which serves to correct the specification.

If there are any additional fees due in connection with the filing of this Request, the Commissioner is hereby authorized to charge any fees due to Deposit Account No. 50-0310.

Respectfully submitted
Morgan, Lewis & Bockius LLP

Dated: **October 6, 2003**
Morgan, Lewis & Bockius LLP
Customer No. **09629**
1111 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
202-739-3000

Robert Smyth
Registration No. 50,801

Form PTO-1050

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. **5,750,400**

ISSUED: **May 12, 1998**

INVENTORS: Patricia D. **MURPHY** *et al.*

It is hereby certified that errors appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

In column 2, lines 42-43, delete "BRCA1$^{(omi2)}$ SEQ ID NO: 3 and BRCA1$^{(omi3)}$ SEQ ID NO: 5" and add -- BRCA1$^{(omi2)}$ SEQ ID NO: 5 and BRCA1$^{(omi3)}$ SEQ ID NO: 3--.

In column 4, lines 29-30, delete "SEQ ID NO: 3 and SEQ ID NO: 5 respectively" and add --SEQ ID NO: 5 and SEQ ID NO: 3 respectively--.

In column 21, line 41, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 22, line 42, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 22, line 45, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 22, line 58, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 23, line 63, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 22, line 67, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 24, line 2, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

In column 24, line 5, delete "SEQ ID NO: 3" and add --SEQ ID NO: 5--.

MAILING ADDRESS OF SENDER
MORGAN, LEWIS & BOCKIUS LLP
Attn: Patent Support Unit
1111 Pennsylvania Avenue, NW
Washington, DC 20004

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 5,750,400 Page 1 of 1
DATED : May 12, 1998
INVENTOR(S) : Patricia D. Murphy et al.

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 2,
Lines 42-43, delete "BRCA1$^{(omi2)}$ SEQ ID NO: 3 and BRCA1$^{(omi3)}$ SEQ ID NO: 5" and add -- BRCA1$^{(omi2)}$ SEQ ID NO: 5 and BRCA1$^{(omi3)}$ SEQ ID NO: 3 --.

Column 4,
Lines 29-30, delete "SEQ ID NO: 3 and SEQ ID NO: 5 respectively" and add -- SEQ ID NO: 5 and SEQ ID NO: 3 respectively --.

Column 21,
Line 41, delete "SEQ ID NO: 3" and add -- SEQ ID NO: 5 --.

Column 22,
Lines 42, 45, 58 and 67, delete "SEQ ID NO: 3" and add -- SEQ ID NO: 5 --.

Column 23,
Line 63, delete "SEQ ID NO: 3" and add -- SEQ ID NO: 5 --.

Column 24,
Lines 2 and 5, delete "SEQ ID NO: 3" and add -- SEQ ID NO: 5 --.

Signed and Sealed this

Seventeenth Day of February, 2004

Jon W. Dudas

JON W. DUDAS
Acting Director of the United States Patent and Trademark Office

NOTICE RE: CERTIFICATES OF CORRECTION

DATE : 12/5/03 Paper No.: ~~#19~~

TO : Supervisor, Art Unit 1642

SUBJECT : Certificate of Correction Request in Patent No.: 5750400

A response to the following question is requested with respect to the accompanying request for a certificate of correction.

With respect to the change(s) requested, correcting Office and/or Applicant's errors, should the patent read as shown in the certificate of correction? No new matter should be introduced, nor should the scope or meaning of the claims be changed.

PLEASE COMPLETE THIS FORM AND RETURN WITH FILE, WITHIN 7 DAYS, TO CERTIFICATES OF CORRECTION BRANCH - PK 3-915/922 PALM LOCATION 7580 - TEL. NO. 305-8309

THANK YOU FOR YOUR ASSISTANCE!

Note your decision, regarding the changes requested in the Request for Certificate of Correction, by placing a check mark (✓) in the box that reflects your decision, which corresponds to the question checked above.

☑ YES ☐ NO ☐ Comments below

☐ Comments:

Supervisor Art Unit 1642

All of the Cited Publications were missing from the Official PTO File History.

PATENT APPLICATION FEE DETERMINATION RECORD

Effective October 1, 1996

Application or Docket Number: 08/798691

CLAIMS AS FILED - PART I

	(Column 1) NUMBER FILED		(Column 2) NUMBER EXTRA	SMALL ENTITY RATE	SMALL ENTITY FEE	OR	OTHER THAN SMALL ENTITY RATE	OTHER THAN SMALL ENTITY FEE
FOR								
BASIC FEE					385.00	OR		770.00
TOTAL CLAIMS	27	minus 20 =	* 7	x$11=	77	OR	x$22=	154
INDEPENDENT CLAIMS	13	minus 3 =	* 10	x40=	400	OR	x80=	
MULTIPLE DEPENDENT CLAIM PRESENT				+130=		OR	+260=	
				TOTAL	862	OR	TOTAL	

* If the difference in column 1 is less than zero, enter "0" in column 2

CLAIMS AS AMENDED - PART II

AMENDMENT A

	(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY RATE	SMALL ENTITY ADDITIONAL FEE	OR	OTHER THAN SMALL ENTITY RATE	OTHER THAN SMALL ENTITY ADDITIONAL FEE
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA					
Total	*	Minus	**	=	x$11=		OR	x$22=	
Independent	*	Minus	***	=	x40=		OR	x80=	
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM					+130=		OR	+260=	
					TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT B

	(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY RATE	SMALL ENTITY ADDITIONAL FEE	OR	OTHER THAN SMALL ENTITY RATE	OTHER THAN SMALL ENTITY ADDITIONAL FEE
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA					
Total	*	Minus	**	=	x$11=		OR	x$22=	
Independent	*	Minus	***	=	x40=		OR	x80=	
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM					+130=		OR	+260=	
					TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT C

	(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY RATE	SMALL ENTITY ADDITIONAL FEE	OR	OTHER THAN SMALL ENTITY RATE	OTHER THAN SMALL ENTITY ADDITIONAL FEE
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA					
Total	*	Minus	**	=	x$11=		OR	x$22=	
Independent	*	Minus	***	=	x40=		OR	x80=	
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM					+130=		OR	+260=	
					TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20."
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3."
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Form PTO 1130
(REV 2/94)

PACE DATA ENTRY CODING SHEET

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

1ST EXAMINER	A. RIVERS	DATE	3/23/97
2ND EXAMINER		DATE	

APPLICATION NUMBER: 08/798691 (U.S. PTO 02/12/97)

TYPE APPL	FILING DATE (MONTH DAY YEAR)	SPECIAL HANDLING	GROUP ART UNIT	CLASS	SHEETS OF DRAWING
1	02 12 97	00	1806	514	--2

TOTAL CLAIMS	INDEPENDENT CLAIMS	SMALL ENTITY?	FILING FEE	FOREIGN LICENSE	ATTORNEY DOCKET NUMBER
-27	-13	1	-862	Y	PA-0054C1P

CONTINUITY DATA

CONT CODE	STATUS CODE	PARENT APPLICATION SERIAL NUMBER	PCT APPLICATION SERIAL NUMBER	PARENT PATENT NUMBER	PARENT FILING DATE (MONTH DAY YEAR)
05	2	08598591	PCT/		02 12 96
			PCT/		
			PCT/		
			PCT/		
			PCT/		

PCT/FOREIGN APPLICATION DATA

FOREIGN PRIORITY CLAIMED	COUNTRY CODE	PCT/FOREIGN APPLICATION SERIAL NUMBER	FOREIGN FILING DATE (MONTH DAY YEAR)