

08/308104

Class

Subclass

ISSUE CLASSIFICATION

SCANNED 4

ABANDONED

UTILITY
SERIAL
NUMBER

08/308104

PATENT DATE

PATENT
NUMBERSERIAL NUMBER
08/308,104FILING DATE
09/16/94

CLASS

SUBCLASS

GROUP ART UNIT

EXAMINER

APPLICANTS MARK H. SKOLNICK, SALT LAKE CITY, UT; DAVID E. GOLDBERG, SALT LAKE CITY, UT; YOSHIO MIKI, SALT LAKE CITY, UT; JEFF SWENSON, SALT LAKE CITY, UT; ALEXANDER KAMB, SALT LAKE CITY, UT; KEITH D. HARSHMAN, SALT LAKE CITY, UT; DONNA M. SHATTUCK-EIDENS, SALT LAKE CITY, UT; SEAN V. TAVTIGIAN, SALT LAKE CITY, UT.

CONTINUING DATA***

VERIFIED THIS APPLN IS A CIP OF 08/300,266 09/02/94 ABN

FOREIGN APPLICATIONS***

VERIFIED

FOREIGN FILING LICENSE GRANTED 05/08/98

Foreign priority claimed
35 USC 119 conditions met☐

yes

☐

no

☐

yes

☐

no

AS
FILEDSTATE OR
COUNTRYSHEETS
DRWGS.TOTAL
CLAIMSINDEP.
CLAIMSFILING FEE
RECEIVEDATTORNEY'S
DOCKET NO.

Verified and Acknowledged

Examiner's Initials

UT

10

74

21

\$3,416.00

248841093473

ADDRESS

VENABLE, BAETJER, HOWARD & CIVILETTI
SUITE 1000
1201 NEW YORK AVENUE, N.W.
WASHINGTON DC 20005

17Q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

TITLE

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

Total Claims

Print Claim

Assistant Examiner

ISSUE FEE

DRAWING

Amount Due

Date Paid

Sheets Drwg.

Figs. Drwg.

Print Fig.

Label
Area

Primary Examiner

PREPARED FOR ISSUE

ISSUE
BATCH
NUMBER

WARNING: The information disclosed herein may be restricted. Unauthorized disclosure may be prohibited by the United States Code Title 35, Sections 122, 181 and 368. Possession outside the U.S. Patent & Trademark Office is restricted to authorized employees and contractors only.

08/308104

PATENT APPLICATION



08308104

APPROVED FOR LICENSE ☐

INITIAL SEP 309428

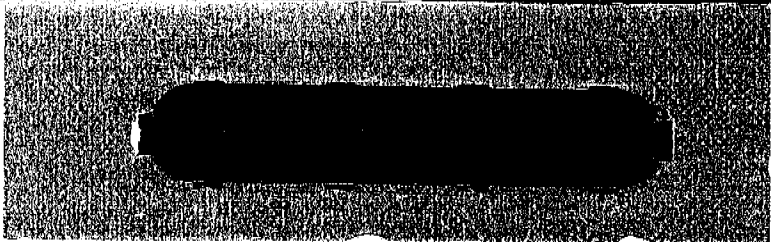
Date
Entered
or
Counted

CONTENTS

Date
Received
or
Mailed

IPEP

1. Application _____ papers.
2. *Spec. Declaration / 700 / DNA Seq* 12/6/94
3. *Petition under 1.48* 2-13-95
4. *3 1/2 RAN history (copy)* 3-17-95
4. Requirement For Information 8-28-95
5. Response 8-15-95
6. Suppl. Requirement For Infor. 9-29-95
7. Response To Requirement For Inf. 10-26-95
8. *Def. (Copied)* 3/11/06
9. *1.53 ABN.* 5-12-98
10. *Request for Canceled* 10-2-01
11. *Request for Appeal* 4/30/03
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____
21. _____
22. _____
23. _____
24. _____
25. _____
26. _____
27. _____
28. _____
29. _____
30. _____
31. _____
32. _____



SEARCHED			
Class	Sub.	Date	Exmr.

SEARCH NOTES		
	Date	Exmr.

INTERFERENCE SEARCHED			
Class	Sub.	Date	Exmr.

Staple Issue Slip Here

POSITION		ID NO.	DATE
CLASSIFIER		6	9-30-94
EXAMINER		4/10	11/30/94
TYPIST			
VERIFIER			
CORPS CORR.			
SPEC. HAND			
FILE MAINT.			
DRAFTING			

INDEX OF CLAIMS

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

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

Claim		Date					
Final	Original						
	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						

PATENT APPLICATION SERIAL NO. 08/308104

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

Commission of Patents and Trademarks
Washington, D.C. 20231

08/308104

Sir:

Transmitted herewith for filing is the patent application of:

Inventors: Mark H. SKOLNICK, David E. GOLDGAR, Yoshio MIKI, Jeff SWENSON, Alexander KAMB,
Keith D. HARSHMAN, Donna M. SHATTUCK-EIDENS and Sean V. TAVTIGIAN

For: 17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTABILITY GENE

Enclosed are:

- [X] Ten (10) sheet(s) of drawings, in triplicate.
[] Assignment of the invention to _____
[] Verified statement establishing small entity status under 37 CFR 1.9 and 37 CFR 1.27
[] Statement Pursuant to 37 CFR §1.821(f) and Computer Disk.
[X] Application filed with NO DECLARATION and NO FEE pursuant to 37 CFR §1.37(d).

(COL. 1)		(COL. 2)	SMALL ENTITY			OTHER THAN SMALL ENTITY	
FOR:	NO. FILED	EXTRA	RATE	FEE		RATE	FEE
BASIC FEE:				\$355.00	OR		\$710.00
TOTAL CLAIMS:	-20 =		X \$ 11	\$	OR	X \$ 22	\$
INDEP CLAIMS:	- 3 =		X \$ 37	\$	OR	X \$ 74	\$
MULTIPLE DEPENDENT CLAIM PRESENTED:			+ \$115	\$	OR	+ \$230	\$
			TOTAL	\$	OR	TOTAL	\$

- [] Charge Deposit Account No. 22-0261 in the amount of \$_____. Duplicate copies of this sheet are attached.
[] A check in the amount of \$_____ to cover the filing fee and assignment recordation fee is attached.

Please direct all correspondence to: Jeffrey L. Ihnen, Reg. No. 28,957

Respectfully submitted,

Jeffrey L. Ihnen
Jeffrey L. Ihnen
Registration No. 28,957
Attorney for Applicant

VENABLE, BAETJER, HOWARD & CIVILETTI
1201 New York Avenue, N.W., Suite 1000
Washington, D.C. 20005
202-962-4810

Dated: 16 September 1994
Atty. Docket No. 24884-109347-3

TITLE OF THE INVENTION

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

ABSTRACT OF THE DISCLOSURE

5 The present invention relates generally to the field of human genetics. Specifically, the present invention relates to methods and materials used to isolate and detect a human breast and ovarian cancer predisposing gene (BRCA1), some mutant alleles of which cause susceptibility to cancer, in particular breast and ovarian cancer. More specifically, the invention relates to germline mutations in the BRCA1 gene and their use in the diagnosis of predisposition to breast and ovarian cancer. The present invention further relates to somatic mutations in the BRCA1 gene in human breast and ovarian cancer and their use in the diagnosis and prognosis of human breast and ovarian cancer. Additionally, the invention relates to somatic mutations in the BRCA1 gene in other human cancers and their use in the diagnosis and prognosis of human cancers. The invention also relates to the therapy of human cancers which have a mutation in the BRCA1 gene, including gene therapy, protein replacement therapy and protein mimetics. The invention further relates to the screening of drugs for cancer therapy. Finally, the invention relates to the screening of the BRCA1 gene for mutations, which are useful for diagnosing the predisposition to breast and ovarian cancer.



Handwritten signature/initials

Handwritten signature/initials

08/308104

TITLE OF THE INVENTION

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

This work has been supported in part by National Cancer Institute grant Numbers CA-48711 and CA-55914. The United States government has certain rights in the invention.

5 CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of application Serial No. 08/300,266, filed on 2 September 1994, which is a continuation-in-part of application Serial No. 08/289,221, filed on 12 August 1994, which is a continuation-in-part of application Serial No. 08/246,602, filed on 20 May 1994, both incorporated herein by reference.

10
15
20
25

FIELD OF THE INVENTION

The present invention relates generally to the field of human genetics. Specifically, the present invention relates to methods and materials used to isolate and detect a human breast and ovarian cancer predisposing gene (BRCA1), some mutant alleles of which cause susceptibility to cancer, in particular, breast and ovarian cancer. More specifically, the invention relates to germline mutations in the BRCA1 gene and their use in the diagnosis of predisposition to breast and ovarian cancer. The present invention further relates to somatic mutations in the BRCA1 gene in human breast and ovarian cancer and their use in the diagnosis and prognosis of human breast and ovarian cancer. Additionally, the invention relates to somatic mutations in the BRCA1 gene in other human cancers and their use in the diagnosis and prognosis of human cancers. The invention also relates to the therapy of human cancers which have a mutation in the BRCA1 gene, including gene therapy, protein replacement therapy and protein mimetics. The invention further relates to the screening of drugs for cancer therapy. Finally, the invention relates to the screening of the BRCA1 gene for mutations, which are useful for diagnosing the predisposition to breast and ovarian cancer.

The publications and other materials used herein to illuminate the background of the invention, and in particular, cases to provide additional details respecting the practice, are incorporated herein by reference, and for convenience, are referenced by author and date in the following text and respectively grouped in the appended bibliography.

BACKGROUND OF THE INVENTION

The genetics of cancer is complicated, involving multiple dominant, positive regulators of the transformed state (oncogenes) as well as multiple recessive, negative regulators (tumor suppressor genes). Over one hundred oncogenes have been characterized. Less than a dozen tumor suppressor genes have been identified, but the number is expected to increase beyond fifty (Knudson, 1993).

The involvement of so many genes underscores the complexity of the growth control mechanisms that operate in cells to maintain the integrity of normal tissue. This complexity is manifest in another way. So far, no single gene has been shown to participate in the development of all, or even the majority of human cancers. The most common oncogenic mutations are in the H-ras gene, found in 10-15% of all solid tumors (Anderson *et al.*, 1992). The most frequently mutated tumor suppressor genes are the TP53 gene, homozygously deleted in roughly 50% of all tumors, and CDKN2, which was homozygously deleted in 46% of tumor cell lines examined (Kamb *et al.*, 1994). Without a target that is common to all transformed cells, the dream of a "magic bullet" that can destroy or revert cancer cells while leaving normal tissue unharmed is improbable. The hope for a new generation of specifically targeted antitumor drugs may rest on the ability to identify tumor suppressor genes or oncogenes that play general roles in control of cell division.

The tumor suppressor genes which have been cloned and characterized influence susceptibility to: 1) retinoblastoma (RB1); 2) Wilms' tumor (WT1); 3) Li-Fraumeni (TP53); 4) Familial adenomatous polyposis (APC); 5) Neurofibromatosis type 1 (NF1); 6) Neurofibromatosis type 2 (NF2); 7) von Hippel-Lindau syndrome (VHL); 8) Multiple endocrine neoplasia type 2A (MEN2A); and 9) Melanoma (CDKN2).

Tumor suppressor loci that have been mapped genetically but not yet isolated include genes for: Multiple endocrine neoplasia type 1 (MEN1); Lynch cancer family syndrome 2 (LCFS2); Neuroblastoma (NB); Basal cell nevus syndrome (BCNS); Beckwith-Wiedemann syndrome (BWS); Renal cell carcinoma (RCC); Tuberous sclerosis 1 (TSC1); and Tuberous sclerosis 2 (TSC2). The tumor suppressor genes that have been characterized to date encode products with similarities to a variety of protein types, including DNA binding proteins (WT1), ancillary transcription regulators (RB1), GTPase activating proteins or GAPs (NF1), cytoskeletal components (NF2), membrane bound receptor kinases (MEN2A), cell cycle regulators (CDKN2) and others with no obvious similarity to known proteins (APC and VHL).

In many cases, the tumor suppressor gene originally identified through genetic studies has been shown to be lost or mutated in some sporadic tumors. This result suggests that regions of chromosomal aberration may signify the position of important tumor suppressor genes involved both in genetic predisposition to cancer and in sporadic cancer.

5 One of the hallmarks of several tumor suppressor genes characterized to date is that they are deleted at high frequency in certain tumor types. The deletions often involve loss of a single allele, a so-called loss of heterozygosity (LOH), but may also involve homozygous deletion of both alleles. For LOH, the remaining allele is presumed to be nonfunctional, either because of a preexisting inherited mutation, or because of a secondary sporadic mutation.

10 Breast cancer is one of the most significant diseases that affects women. At the current rate, American women have a 1 in 8 risk of developing breast cancer by age 95 (American Cancer Society, 1992). Treatment of breast cancer at later stages is often futile and disfiguring, making early detection a high priority in medical management of the disease. Ovarian cancer, although less frequent than breast cancer is often rapidly fatal and is the fourth most common cause of cancer mortality in American women. Genetic factors contribute to an ill-defined proportion of breast cancer incidence, estimated to be about 5% of all cases but approximately 25% of cases diagnosed before age 40 (Claus *et al.*, 1991). Breast cancer has been subdivided into two types, early-age onset and late-age onset, based on an inflection in the age-specific incidence curve around age 50. Mutation of one gene, BRCA1, is thought to account for approximately 45% of familial breast cancer, but at least 80% of families with both breast and ovarian cancer (Easton *et al.*, 1993).

20 Intense efforts to isolate the BRCA1 gene have proceeded since it was first mapped in 1990 (Hall *et al.*, 1990; Narod *et al.*, 1991). A second locus, BRCA2, has recently been mapped to chromosome 13q (Wooster *et al.*, 1994) and appears to account for a proportion of early-onset breast cancer roughly equal to BRCA1, but confers a lower risk of ovarian cancer. The remaining susceptibility to early-onset breast cancer is divided between as yet unmapped genes for familial cancer, and rarer germline mutations in genes such as TP53 (Malkin *et al.*, 1990). It has also been suggested that heterozygote carriers for defective forms of the Ataxia-Telangectasia gene are at higher risk for breast cancer (Swift *et al.*, 1976; Swift *et al.*, 1991). Late-age onset breast cancer is also often familial although the risks in relatives are not as high as those for early-onset breast cancer, (Cannon-Albright *et al.*, 1994; Mettlin *et al.*, 1990). However, the percentage of such cases due to genetic susceptibility is unknown.

Breast cancer has long been recognized to be, in part, a familial disease (Anderson, 1972). Numerous investigators have examined the evidence for genetic inheritance and concluded that the data are most consistent with dominant inheritance for a major susceptibility locus or loci. (Hill *et al.*, 1978; Bishop and Gardner, 1980; Go *et al.*, 1983; Willams and Anderson, 1984; Bishop *et al.*, 1988; Newman *et al.*, 1988; Claus *et al.*, 1991). Recent results demonstrate that at least three loci exist which convey susceptibility to breast cancer as well as other cancers. These loci are the TP53 locus on chromosome 17p (Malkin *et al.*, 1990), a 17q-linked susceptibility locus known as BRCA1 (Hall *et al.*, 1990), and one or more loci responsible for the unmapped residual. Hall *et al.* (1990) indicated that the inherited breast cancer susceptibility in kindreds with early age onset is linked to chromosome 17q21; although subsequent studies by this group using a more appropriate genetic model partially refuted the limitation to early onset breast cancer (Margaritte *et al.*, 1992).

Most strategies for cloning the 17q-linked breast cancer predisposing gene (BRCA1) require precise genetic localization studies. The simplest model for the functional role of BRCA1 holds that alleles of BRCA1 that predispose to cancer are recessive to wild type alleles; that is, cells that contain at least one wild type BRCA1 allele are not cancerous. However, cells that contain one wild type BRCA1 allele and one predisposing allele may occasionally suffer loss of the wild type allele either by random mutation or by chromosome loss during cell division (nondisjunction). All the progeny of such a mutant cell lack the wild type function of BRCA1 and may develop into tumors. According to this model, predisposing alleles of BRCA1 are recessive, yet susceptibility to cancer is inherited in a dominant fashion: women who possess one predisposing allele (and one wild type allele) risk developing cancer, because their mammary epithelial cells may spontaneously lose the wild type BRCA1 allele. This model applies to a group of cancer susceptibility loci known as tumor suppressors or antioncogenes, a class of genes that includes the retinoblastoma gene and neurofibromatosis gene. By inference this model may also explain the BRCA1 function, as has recently been suggested (Smith *et al.*, 1992).

A second possibility is that BRCA1 predisposing alleles are truly dominant; that is, a wild type allele of BRCA1 cannot overcome the tumor forming role of the predisposing allele. Thus, a cell that carries both wild type and mutant alleles would not necessarily lose the wild type copy of BRCA1 before giving rise to malignant cells. Instead, mammary cells in predisposed individuals would undergo some other stochastic change(s) leading to cancer.

If BRCA1 predisposing alleles are recessive, the BRCA1 gene is expected to be expressed in normal mammary tissue but not functionally expressed in mammary tumors. In contrast, if BRCA1

predisposing alleles are dominant, the wild type BRCA1 gene may or may not be expressed in normal mammary tissue. However, the predisposing allele will likely be expressed in breast tumor cells.

5 The 17q linkage of BRCA1 was independently confirmed in three of five kindreds with both breast cancer and ovarian cancer (Narod *et al.*, 1991). These studies claimed to localize the gene within a very large region, 15 centimorgans (Cm), or approximately 15 million base pairs, to either side of the linked marker pCMM86 (D17S74). However, attempts to define the region further by genetic studies, using markers surrounding pCMMS6, proved unsuccessful. Subsequent studies indicated that the gene was considerably more proximal (Easton *et al.*, 1993) and that the original
10 analysis was flawed (Margaritte *et al.*, 1992). Hall *et al.*, (1992) recently localized the BRCA1 gene to an approximately 8 Cm interval (approximately 8 million basepairs) bounded by Mfd15 (D17S250) on the proximal side and the human GIP gene on the distal side. A slightly narrower interval for the BRCA1 locus, based on publicly available data, was agreed upon at the Chromosome 17 workshop in March of 1992 (Fain, 1992). The size of these regions and the uncertainty associated with them has made it exceedingly difficult to design and implement physical mapping and/or cloning strategies for isolating the BRCA1 gene.

15 Identification of a breast cancer susceptibility locus would permit the early detection of susceptible individuals and greatly increase our ability to understand the initial steps which lead to cancer. As susceptibility loci are often altered during tumor progression, cloning these genes could also be important in the development of better diagnostic and prognostic products, as well as better cancer therapies.

SUMMARY OF THE INVENTION

20 The present invention relates generally to the field of human genetics. Specifically, the present invention relates to methods and materials used to isolate and detect a human breast cancer predisposing gene (BRCA1), some alleles of which cause susceptibility to cancer, in particular breast
25 and ovarian cancer. More specifically, the present invention relates to germline mutations in the BRCA1 gene and their use in the diagnosis of predisposition to breast and ovarian cancer. The invention further relates to somatic mutations in the BRCA1 gene in human breast cancer and their use in the diagnosis and prognosis of human breast and ovarian cancer. Additionally, the invention
30 relates to somatic mutations in the BRCA1 gene in other human cancers and their use in the

diagnosis and prognosis of human cancers. The invention also relates to the therapy of human cancers which have a mutation in the BRCA1 gene, including gene therapy, protein replacement therapy and protein mimetics. The invention further relates to the screening of drugs for cancer therapy. Finally, the invention relates to the screening of the BRCA1 gene for mutations, which are useful for
5 diagnosing the predisposition to breast and ovarian cancer.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagram showing the order of loci neighboring BRCA1 as determined by the chromosome 17 workshop. Figure 1 is reproduced from Fain, 1992.

Figure 2 is a schematic map of YACs which define part of Mfd15-Mfd188 region.

Figure 3 is a schematic map of STS's, P1s and BAC's in the BRCA1 region.

Figure 4 is a schematic map of human chromosome 17. The pertinent region containing BRCA1 is expanded to indicate the relative positions of two previously identified genes, CA125 and RNU2, BRCA1 spans the marker D17S855.

Figure 5 shows alignment of the BRCA1 zinc-finger domain with 3 other zinc-finger domains that scored highest in a Smith-Waterman alignment. RPT1 encodes a protein that appears to be a negative regulator of the IL-2 receptor in mouse. RIN1 encodes a DNA-binding protein that includes a RING-finger motif related to the zinc-finger. RFP1 encodes a putative transcription factor that is the N-terminal domain of the RET oncogene product. The bottom line contains the C3HC4 consensus zinc-finger sequence showing the positions of cystines and one histidine that form the zinc
15 ion binding pocket.
20

Figure 6 is a diagram of BRCA1 mRNA showing the locations of introns and the variants of BRCA1 mRNA produced by alternative splicing. Intron locations are shown by dark triangles and the exons are numbered below the line representing the cDNA. The top cDNA is the composite used to generate the peptide sequence of BRCA1. Alternative forms identified as cDNA clones or hybrid
25 selection clones are shown below.

Figure 7 shows the tissue expression pattern of BRCA1. The blot was obtained from Clontech and contains RNA from the indicated tissues. Hybridization conditions were as recommended by the manufacturer using a probe consisting of nucleotide positions 3575 to 3874 of BRCA1. Note that both breast and ovary are heterogeneous tissues and the percentage of relevant epithelial cells can be
30 variable. Molecular weight standards are in kilobases.

Figure 8A shows a mutation and cosegregation analysis in BRCA1 kindreds. Carrier individuals are represented as filled circles and squares in the pedigree diagrams. Nonsense mutation in Kindred 2082. P indicates the person originally screened, b and c are haplotype carriers, a, d, f, and g do not carry the BRCA1 haplotype. The C to T mutation results in a stop codon and creates a site for the restriction enzyme AvrII. PCR amplification products are cut with this enzyme. The carriers are heterozygous for the site and therefore show three bands. Non-carriers remain uncut.

Figure 8B shows a mutation and cosegregation analysis in BRCA1 kindreds. Carrier individuals are represented as filled circles and squares in the pedigree diagrams. Frameshift mutation in Kindred 1910. The first three lanes are control, noncarrier samples. Lanes labeled 1-3 contain sequences from carrier individuals. Lane 4 contains DNA from a kindred member who does not carry the BRCA1 mutation. The diamond is used to prevent identification of the kindred. The frameshift resulting from the additional C is apparent in lanes labeled 1, 2, and 3.

Figure 8C shows a mutation and cosegregation analysis in BRCA1 kindreds. Carrier individuals are represented as filled circles and squares in the pedigree diagrams. Inferred regulatory mutation in Kindred 2035. ASO analysis of carriers and noncarriers of 2 different polymorphisms (PM1 and PM7) which were examined for heterozygosity in the germline and compared to the heterozygosity of lymphocyte mRNA. The top 2 rows of each panel contain PCR products amplified from genomic DNA and the bottom 2 rows contain PCR products amplified from cDNA. "A" and "G" are the two alleles detected by the ASO. The dark spots indicate that a particular allele is present in the sample. The first three lanes of PM7 represent the three genotypes in the general population.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates generally to the field of human genetics. Specifically, the present invention relates to methods and materials used to isolate and detect a human breast cancer predisposing gene (BRCA1), some alleles of which cause susceptibility to cancer, in particular breast and ovarian cancer. More specifically, the present invention relates to germline mutations in the BRCA1 gene and their use in the diagnosis of predisposition to breast and ovarian cancer. The invention further relates to somatic mutations in the BRCA1 gene in human breast cancer and their use in the diagnosis and prognosis of human breast and ovarian cancer. Additionally, the invention relates to somatic mutations in the BRCA1 gene in other human cancers and their use in the

diagnosis and prognosis of human cancers. The invention also relates to the therapy of human cancers which have a mutation in the BRCA1 gene, including gene therapy, protein replacement therapy and protein mimetics. The invention further relates to the screening of drugs for cancer therapy. Finally, the invention relates to the screening of the BRCA1 gene for mutations, which are useful for
5 diagnosing the predisposition to breast and ovarian cancer.

The present invention provides an isolated polynucleotide comprising all, or a portion of the BRCA1 locus or of a mutated BRCA1 locus, preferably at least eight bases and not more than about 100 Kb in length. Such polynucleotides may be antisense polynucleotides. The present invention also provides a recombinant construct comprising such an isolated polynucleotide, for example, a
10 recombinant construct suitable for expression in a transformed host cell.

Also provided by the present invention are methods of detecting a polynucleotide comprising a portion of the BRCA1 locus or its expression product in an analyte. Such methods may further comprise the step of amplifying the portion of the BRCA1 locus, and may further include a step of providing a set of polynucleotides which are primers for amplification of said portion of the BRCA1
15 locus. The method is useful for either diagnosis of the predisposition to cancer or the diagnosis or prognosis of cancer.

The present invention also provides isolated antibodies, preferably monoclonal antibodies, which specifically bind to an isolated polypeptide comprised of at least five amino acid residues encoded by the BRCA1 locus.

20 The present invention also provides kits for detecting in an analyte a polynucleotide comprising a portion of the BRCA1 locus, the kits comprising a polynucleotide complementary to the portion of the BRCA1 locus packaged in a suitable container, and instructions for its use.

The present invention further provides methods of preparing a polynucleotide comprising polymerizing nucleotides to yield a sequence comprised of at least eight consecutive nucleotides of
25 the BRCA1 locus; and methods of preparing a polypeptide comprising polymerizing amino acids to yield a sequence comprising at least five amino acids encoded within the BRCA1 locus.

The present invention further provides methods of screening the BRCA1 gene to identify mutations. Such methods may further comprise the step of amplifying a portion of the BRCA1 locus, and may further include a step of providing a set of polynucleotides which are primers for
30 amplification of said portion of the BRCA1 locus. The method is useful for identifying mutations for use in either diagnosis of the predisposition to cancer or the diagnosis or prognosis of cancer.

The present invention further provides methods of screening suspected BRCA1 mutant alleles to identify mutations in the BRCA1 gene.

In addition, the present invention provides methods of screening drugs for cancer therapy to identify suitable drugs for restoring BRCA1 gene product function.

5 Finally, the present invention provides the means necessary for production of gene-based therapies directed at cancer cells. These therapeutic agents may take the form of polynucleotides comprising all or a portion of the BRCA1 locus placed in appropriate vectors or delivered to target cells in more direct ways such that the function of the BRCA1 protein is reconstituted. Therapeutic agents may also take the form of polypeptides based on either a portion of, or the entire protein sequence of BRCA1. These may functionally replace the activity of BRCA1 *in vivo*.

10 It is a discovery of the present invention that the BRCA1 locus which predisposes individuals to breast cancer and ovarian cancer, is a gene encoding a BRCA1 protein, which has been found to have no significant homology with known protein or DNA sequences. This gene is termed BRCA1 herein. It is a discovery of the present invention that mutations in the BRCA1 locus in the germline are indicative of a predisposition to breast cancer and ovarian cancer. Finally, it is a discovery of the present invention that somatic mutations in the BRCA1 locus are also associated with breast cancer, ovarian cancer and other cancers, which represents an indicator of these cancers or of the prognosis of these cancers. The mutational events of the BRCA1 locus can involve deletions, insertions and point mutations within the coding sequence and the non-coding sequence.

15 20 Starting from a region on the long arm of human chromosome 17 of the human genome, 17q, which has a size estimated at about 8 million basepairs, a region which contains a genetic locus, BRCA1, which causes susceptibility to cancer, including breast and ovarian cancer, has been identified.

25 The region containing the BRCA1 locus was identified using a variety of genetic techniques. Genetic mapping techniques initially defined the BRCA1 region in terms of recombination with genetic markers. Based upon studies of large extended families ("kindreds") with multiple cases of breast cancer (and ovarian cancer cases in some kindreds), a chromosomal region has been pinpointed that contains the BRCA1 gene as well as other putative susceptibility alleles in the BRCA1 locus. Two meiotic breakpoints have been discovered on the distal side of the BRCA1 locus which are expressed as recombinants between genetic markers and the disease, and one recombinant on the proximal side of the BRCA1 locus. Thus, a region which contains the BRCA1 locus is physically bounded by these markers.

30

The use of the genetic markers provided by this invention allowed the identification of clones which cover the region from a human yeast artificial chromosome (YAC) or a human bacterial artificial chromosome (BAC) library. It also allowed for the identification and preparation of more easily manipulated cosmid, P1 and BAC clones from this region and the construction of a contig from a subset of the clones. These cosmids, P1s, YACs and BACs provide the basis for cloning the BRCA1 locus and provide the basis for developing reagents effective, for example, in the diagnosis and treatment of breast and/or ovarian cancer. The BRCA1 gene and other potential susceptibility genes have been isolated from this region. The isolation was done using software trapping (a computational method for identifying sequences likely to contain coding exons, from contiguous or discontinuous genomic DNA sequences), hybrid selection techniques and direct screening, with whole or partial cDNA inserts from cosmids, P1s and BACs, in the region to screen cDNA libraries. These methods were used to obtain sequences of loci expressed in breast and other tissue. These candidate loci were analyzed to identify sequences which confer cancer susceptibility. We have discovered that there are mutations in the coding sequence of the BRCA1 locus in kindreds which are responsible for the 17q-linked cancer susceptibility known as BRCA1. This gene was not known to be in this region. The present invention not only facilitates the early detection of certain cancers, so vital to patient survival, but also permits the detection of susceptible individuals before they develop cancer.

Population Resources

Large, well-documented Utah kindreds are especially important in providing good resources for human genetic studies. Each large kindred independently provides the power to detect whether a BRCA1 susceptibility allele is segregating in that family. Recombinants informative for localization and isolation of the BRCA1 locus could be obtained only from kindreds large enough to confirm the presence of a susceptibility allele. Large sibships are especially important for studying breast cancer, since penetrance of the BRCA1 susceptibility allele is reduced both by age and sex, making informative sibships difficult to find. Furthermore, large sibships are essential for constructing haplotypes of deceased individuals by inference from the haplotypes of their close relatives.

While other populations may also provide beneficial information, such studies generally require much greater effort, and the families are usually much smaller and thus less informative. Utah's age-adjusted breast cancer incidence is 20% lower than the average U.S. rate. The lower incidence in Utah is probably due largely to an early age at first pregnancy, increasing the probability that cases found in Utah kindreds carry a genetic predisposition.

Genetic Mapping

Given a set of informative families, genetic markers are essential for linking a disease to a region of a chromosome. Such markers include restriction fragment length polymorphisms (RFLPs) (Botstein *et al.*, 1980), markers with a variable number of tandem repeats (VNTRs) (Jeffreys *et al.*, 1985; Nakamura *et al.*, 1987), and an abundant class of DNA polymorphisms based on short tandem repeats (STRs), especially repeats of CpA (Weber and May, 1989; Litt *et al.*, 1989). To generate a genetic map, one selects potential genetic markers and tests them using DNA extracted from members of the kindreds being studied.

Genetic markers useful in searching for a genetic locus associated with a disease can be selected on an *ad hoc* basis, by densely covering a specific chromosome, or by detailed analysis of a specific region of a chromosome. A preferred method for selecting genetic markers linked with a disease involves evaluating the degree of informativeness of kindreds to determine the ideal distance between genetic markers of a given degree of polymorphism, then selecting markers from known genetic maps which are ideally spaced for maximal efficiency. Informativeness of kindreds is measured by the probability that the markers will be heterozygous in unrelated individuals. It is also most efficient to use STR markers which are detected by amplification of the target nucleic acid sequence using PCR; such markers are highly informative, easy to assay (Weber and May, 1989), and can be assayed simultaneously using multiplexing strategies (Skolnick M.H., and Wallace B.R., 1988), greatly reducing the number of experiments required.

Once linkage has been established, one needs to find markers that flank the disease locus, i.e., one or more markers proximal to the disease locus, and one or more markers distal to the disease locus. Where possible, candidate markers can be selected from a known genetic map. Where none is known, new markers can be identified by the STR technique, as shown in the Examples.

Genetic mapping is usually an iterative process. In the present invention, it began by defining flanking genetic markers around the BRCA1 locus, then replacing these flanking markers with other markers that were successively closer to the BRCA1 locus. As an initial step, recombination events, defined by large extended kindreds, helped specifically to localize the BRCA1 locus as either distal or proximal to a specific genetic marker (Goldgar *et al.*, 1993).

The region surrounding BRCA1, until the disclosure of the present invention, was not well mapped and there were few markers. Therefore, short repetitive sequences on cosmids subcloned from YACs, which had been physically mapped, were analyzed in order to develop new genetic markers. Using this approach, one marker of the present invention, 42D6, was discovered which

replaced pCMM86 as the distal flanking marker for the BRCA1 region. Since 42D6 is approximately 14 Cm from pCMM86, the BRCA1 region was thus reduced by approximately 14 centimorgans (Easton *et al.*, 1993). The present invention thus began by finding a much more closely linked distal flanking marker of the BRCA1 region. BRCA1 was then discovered to be distal to the genetic marker Mfd15. Therefore, BRCA1 was shown to be in a region of 6 to 10 million bases bounded by Mfd15 and 42D6. Marker Mfd191 was subsequently discovered to be distal to Mfd15 and proximal to BRCA1. Thus, Mfd15 was replaced with Mfd191 as the closest proximal genetic marker. Similarly, it was discovered that genetic marker Mfd188 could replace genetic marker 42D6, narrowing the region containing the BRCA1 locus to approximately 1.5 million bases. Then the marker Mfd191 was replaced with tdj1474 as the proximal marker and Mfd188 was replaced with U5R as the distal marker, further narrowing the BRCA1 region to a small enough region to allow isolation and characterization of the BRCA1 locus (see Figure 3), using techniques known in the art and described herein.

Physical Mapping

Three distinct methods were employed to physically map the region. The first was the use of yeast artificial chromosomes (YACs) to clone the region which is flanked by tdj1474 and U5R. The second was the creation of a set of P1, BAC and cosmid clones which cover the region containing the BRCA1 locus.

Yeast Artificial Chromosomes (YACs). Once a sufficiently small region containing the BRCA1 locus was identified, physical isolation of the DNA in the region proceeded by identifying a set of overlapping YACs which covers the region. Useful YACs can be isolated from known libraries, such as the St. Louis and CEPH YAC libraries, which are widely distributed and contain approximately 50,000 YACs each. The YACs isolated were from these publicly accessible libraries and can be obtained from a number of sources including the Michigan Genome Center. Clearly, others who had access to these YACS, without the disclosure of the present invention, would not have known the value of the specific YACs we selected since they would not have known which YACs were within, and which YACs outside of, the smallest region containing the BRCA1 locus.

Cosmid, P1 and BAC Clones. In the present invention, it is advantageous to proceed by obtaining cosmid, P1, and BAC clones to cover this region. The smaller size of these inserts, compared to YAC inserts, makes them more useful as specific hybridization probes. Furthermore, having the cloned DNA in bacterial cells, rather than in yeast cells, greatly increases the ease with

which the DNA of interest can be manipulated, and improves the signal-to-noise ratio of hybridization assays. For cosmid subclones of YACs, the DNA is partially digested with the restriction enzyme Sau3A and cloned into the BamHI site of the pWE15 cosmid vector (Stratagene, cat. #1251201). The cosmids containing human sequences are screened by hybridization with human repetitive DNA (e.g., Gibco/BRL, Human Cot-1 DNA, cat. 5279SA), and then fingerprinted by a variety of techniques, as detailed in the Examples.

P1 and BAC clones are obtained by screening libraries constructed from the total human genome with specific sequence tagged sites (STSs) derived from the YACs, cosmids or P1s and BACs, isolated as described herein.

These P1, BAC and cosmid clones can be compared by interspersed repetitive sequence (IRS) PCR and/or restriction enzyme digests followed by gel electrophoresis and comparison of the resulting DNA fragments ("fingerprints"). (Maniatis *et al.*, 1982). The clones can also be characterized by the presence of STSs. The fingerprints are used to define an overlapping contiguous set of clones which covers the region but is not excessively redundant, referred to herein as a "minimum tiling path". Such a minimum tiling path forms the basis for subsequent experiments to identify cDNAs which may originate from the BRCA1 locus.

Coverage of the Gap with P1 and BAC Clones. To cover any gaps in the BRCA1 contig between the identified cosmids with genomic clones, clones in P1 and BAC vectors which contain inserts of genomic DNA roughly twice as large as cosmids for P1s and still greater for BACs (Sternberg, 1990; Sternberg *et al.*, 1990; Pierce *et al.*, 1992; Shizuya *et al.*, (1992)) were used. P1 clones were isolated by Genome Sciences using PCR primers provided by us for screening. BACs were provided by hybridization techniques in Dr. Mel Simon's laboratory. The strategy of using P1 clones also permitted the covering of the genomic region with an independent set of clones not derived from YACs. This guards against the possibility of other deletions in YACs that have not been detected. These new sequences derived from the P1 clones provide the material for further screening for candidate genes, as described below.

Gene Isolation.

There are many techniques for testing genomic clones for the presence of sequences likely to be candidates for the coding sequence of a locus one is attempting to isolate, including but not limited to:

- a. zoo blots
- b. identifying HTF islands
- c. exon trapping
- d. hybridizing cDNA to cosmids or YACs.
- e. screening cDNA libraries.

(a) Zoo blots. The first technique is to hybridize cosmids to Southern blots to identify DNA sequences which are evolutionarily conserved, and which therefore give positive hybridization signals with DNA from species of varying degrees of relationship to humans (such as monkey, cow, chicken, pig, mouse and rat). Southern blots containing such DNA from a variety of species are commercially available (Clontech, Cat. 7753-1).

(b) Identifying HTF islands. The second technique involves finding regions rich in the nucleotides C and G, which often occur near or within coding sequences. Such sequences are called HTF (HpaI tiny fragment) or CpG islands, as restriction enzymes which contain CpG dimers cut frequently in these regions (Lindsay *et al.*, 1987).

(c) Exon trapping. The third technique is exon trapping, a method that identifies sequences in genomic DNA which contain splice junctions and therefore are likely to comprise coding sequences of genes. Exon amplification (Buckler *et al.*, 1988) is used to select and amplify exons from DNA clones described above. Exon amplification is based on the selection of RNA sequences which are flanked by functional 5' and/or 3' splice sites. The products of the exon amplification are used to screen the breast cDNA libraries to identify a manageable number of candidate genes for further study. Exon trapping can also be performed on small segments of sequenced DNA using computer programs or by software trapping.

(d) Hybridizing cDNA to Cosmids, P1s, BACs or YACS. The fourth technique is a modification of the selective enrichment technique which utilizes hybridization of cDNA to cosmids, P1s, BACs or YACs and permits transcribed sequences to be identified in, and recovered from cloned genomic DNA (Kandpal *et al.*, 1990). The selective enrichment technique, as modified for the present purpose, involves binding DNA from the region of BRCA1 present in a YAC to a column matrix and selecting cDNAs from the relevant libraries which hybridize with the bound DNA, followed by amplification and purification of the bound DNA, resulting in a great enrichment for cDNAs in the region represented by the cloned genomic DNA.

(e) Identification of cDNAs. The fifth technique is to identify cDNAs that correspond to the BRCA1 locus. Hybridization probes containing putative coding sequences, selected using any of the above techniques, are used to screen various libraries, including breast tissue cDNA libraries, ovarian cDNA libraries, and any other necessary libraries.

Another variation on the theme of direct selection of cDNA was also used to find candidate genes for BRCA1 (Lovett *et al.*, 1991; Futreal, 1993). This method uses cosmid, P1 or BAC DNA as the probe. The probe DNA is digested with a blunt cutting restriction enzyme such as HaeIII. Double stranded adaptors are then ligated onto the DNA and serve as binding sites for primers in subsequent PCR amplification reactions using biotinylated primers. Target cDNA is generated from mRNA derived from tissue samples, e.g., breast tissue, by synthesis of either random primed or oligo(dT) primed first strand followed by second strand synthesis. The cDNA ends are rendered blunt and ligated onto double-stranded adaptors. These adaptors serve as amplification sites for PCR. The target and probe sequences are denatured and mixed with human Cot-1 DNA to block repetitive sequences. Solution hybridization is carried out to high Cot-1/2 values to ensure hybridization of rare target cDNA molecules. The annealed material is then captured on avidin beads, washed at high stringency and the retained cDNAs are eluted and amplified by PCR. The selected cDNA is subjected to further rounds of enrichment before cloning into a plasmid vector for analysis.

Testing the cDNA for Candidacy

Proof that the cDNA is the BRCA1 locus is obtained by finding sequences in DNA extracted from affected kindred members which create abnormal BRCA1 gene products or abnormal levels of BRCA1 gene product. Such BRCA1 susceptibility alleles will co-segregate with the disease in large kindreds. They will also be present at a much higher frequency in non-kindred individuals with breast and ovarian cancer than in individuals in the general population. Finally, since tumors often mutate somatically at loci which are in other instances mutated in the germline, we expect to see normal germline BRCA1 alleles mutated into sequences which are identical or similar to BRCA1 susceptibility alleles in DNA extracted from tumor tissue. Whether one is comparing BRCA1 sequences from tumor tissue to BRCA1 alleles from the germline of the same individuals, or one is comparing germline BRCA1 alleles from cancer cases to those from unaffected individuals, the key is to find mutations which are serious enough to cause obvious disruption to the normal function of the gene product. These mutations can take a number of forms. The most severe forms would be frame shift mutations or large deletions which would cause the gene to code for an abnormal protein

or one which would significantly alter protein expression. Less severe disruptive mutations would include small in-frame deletions and nonconservative base pair substitutions which would have a significant effect on the protein produced, such as changes to or from a cysteine residue, from a basic to an acidic amino acid or vice versa, from a hydrophobic to hydrophilic amino acid or vice versa, or other mutations which would affect secondary, tertiary or quaternary protein structure. Silent mutations or those resulting in conservative amino acid substitutions would not generally be expected to disrupt protein function.

According to the diagnostic and prognostic method of the present invention, alteration of the wild-type BRCA1 locus is detected. In addition, the method can be performed by detecting the wild-type BRCA1 locus and confirming the lack of a predisposition to cancer at the BRCA1 locus. "Alteration of a wild-type gene" encompasses all forms of mutations including deletions, insertions and point mutations in the coding and noncoding regions. Deletions may be of the entire gene or of only a portion of the gene. Point mutations may result in stop codons, frameshift mutations or amino acid substitutions. Somatic mutations are those which occur only in certain tissues, e.g., in the tumor tissue, and are not inherited in the germline. Germline mutations can be found in any of a body's tissues and are inherited. If only a single allele is somatically mutated, an early neoplastic state is indicated. However, if both alleles are somatically mutated, then a late neoplastic state is indicated. The finding of BRCA1 mutations thus provides both diagnostic and prognostic information. A BRCA1 allele which is not deleted (e.g., found on the sister chromosome to a chromosome carrying a BRCA1 deletion) can be screened for other mutations, such as insertions, small deletions, and point mutations. It is believed that many mutations found in tumor tissues will be those leading to decreased expression of the BRCA1 gene product. However, mutations leading to non-functional gene products would also lead to a cancerous state. Point mutational events may occur in regulatory regions, such as in the promoter of the gene, leading to loss or diminution of expression of the mRNA. Point mutations may also abolish proper RNA processing, leading to loss of expression of the BRCA1 gene product, or to a decrease in mRNA stability or translation efficiency.

Useful diagnostic techniques include, but are not limited to fluorescent *in situ* hybridization (FISH), direct DNA sequencing, PFGE analysis, Southern blot analysis, single stranded conformation analysis (SSCA), RNase protection assay, allele-specific oligonucleotide (ASO), dot blot analysis and PCR-SSCP, as discussed in detail further below.

Predisposition to cancers, such as breast and ovarian cancer, and the other cancers identified herein, can be ascertained by testing any tissue of a human for mutations of the BRCA1 gene. For

example, a person who has inherited a germline BRCA1 mutation would be prone to develop cancers. This can be determined by testing DNA from any tissue of the person's body. Most simply, blood can be drawn and DNA extracted from the cells of the blood. In addition, prenatal diagnosis can be accomplished by testing fetal cells, placental cells or amniotic cells for mutations of the BRCA1 gene. Alteration of a wild-type BRCA1 allele, whether, for example, by point mutation or deletion, can be detected by any of the means discussed herein.

In order to detect the alteration of the wild-type BRCA1 gene in a tissue, it is helpful to isolate the tissue free from surrounding normal tissues. Means for enriching tissue preparation for tumor cells are known in the art. For example, the tissue may be isolated from paraffin or cryostat sections. Cancer cells may also be separated from normal cells by flow cytometry. These techniques, as well as other techniques for separating tumor cells from normal cells, are well known in the art. If the tumor tissue is highly contaminated with normal cells, detection of mutations is more difficult.

A rapid preliminary analysis to detect polymorphisms in DNA sequences can be performed by looking at a series of Southern blots of DNA cut with one or more restriction enzymes, preferably with a large number of restriction enzymes. Each blot contains a series of normal individuals and a series of cancer cases, tumors, or both. Southern blots displaying hybridizing fragments (differing in length from control DNA when probed with sequences near or including the BRCA1 locus) indicate a possible mutation. If restriction enzymes which produce very large restriction fragments are used, then pulsed field gel electrophoresis (PFGE) is employed.

Detection of point mutations may be accomplished by molecular cloning of the BRCA1 allele(s) and sequencing the allele(s) using techniques well known in the art. Alternatively, the gene sequences can be amplified directly from a genomic DNA preparation from the tumor tissue, using known techniques. The DNA sequence of the amplified sequences can then be determined.

There are six well known methods for a more complete, yet still indirect, test for confirming the presence of a susceptibility allele: 1) single stranded conformation analysis (SSCA) (Orita *et al.*, 1989); 2) denaturing gradient gel electrophoresis (DGGE) (Wartell *et al.*, 1990; Sheffield *et al.*, 1989); 3) RNase protection assays (Finkelstein *et al.*, 1990; Kinszler *et al.*, 1991); 4) allele-specific oligonucleotides (ASOs) (Conner *et al.*, 1983); 5) the use of proteins which recognize nucleotide mismatches, such as the *E. coli* mutS protein (Modrich, 1991); and, 6) allele-specific PCR (Rano & Kidd, 1989). For allele-specific PCR, primers are used which hybridize at their 3' ends to a particular BRCA1 mutation. If the particular BRCA1 mutation is not present, an amplification product is not observed. Amplification Refractory Mutation System (ARMS) can also be used, as disclosed in

European Patent Application Publication No. 0332435 and in Newton *et al.*, 1989. Insertions and deletions of genes can also be detected by cloning, sequencing and amplification. In addition, restriction fragment length polymorphism (RFLP) probes for the gene or surrounding marker genes can be used to score alteration of an allele or an insertion in a polymorphic fragment. Such a method is particularly useful for screening relatives of an affected individual for the presence of the BRCA1 mutation found in that individual. Other techniques for detecting insertions and deletions as known in the art can be used.

In the first three methods (SSCA, DGGE and RNase protection assay), a new electrophoretic band appears. SSCA detects a band which migrates differentially because the sequence change causes a difference in single-strand, intramolecular base pairing. RNase protection involves cleavage of the mutant polynucleotide into two or more smaller fragments. DGGE detects differences in migration rates of mutant sequences compared to wild-type sequences, using a denaturing gradient gel. In an allele-specific oligonucleotide assay, an oligonucleotide is designed which detects a specific sequence, and the assay is performed by detecting the presence or absence of a hybridization signal. In the mutS assay, the protein binds only to sequences that contain a nucleotide mismatch in a heteroduplex between mutant and wild-type sequences.

Mismatches, according to the present invention, are hybridized nucleic acid duplexes in which the two strands are not 100% complementary. Lack of total homology may be due to deletions, insertions, inversions or substitutions. Mismatch detection can be used to detect point mutations in the gene or in its mRNA product. While these techniques are less sensitive than sequencing, they are simpler to perform on a large number of tumor samples. An example of a mismatch cleavage technique is the RNase protection method. In the practice of the present invention, the method involves the use of a labeled riboprobe which is complementary to the human wild-type BRCA1 gene coding sequence. The riboprobe and either mRNA or DNA isolated from the tumor tissue are annealed (hybridized) together and subsequently digested with the enzyme RNase A which is able to detect some mismatches in a duplex RNA structure. If a mismatch is detected by RNase A, it cleaves at the site of the mismatch. Thus, when the annealed RNA preparation is separated on an electrophoretic gel matrix, if a mismatch has been detected and cleaved by RNase A, an RNA product will be seen which is smaller than the full length duplex RNA for the riboprobe and the mRNA or DNA. The riboprobe need not be the full length of the BRCA1 mRNA or gene but can be a segment of either. If the riboprobe comprises only a segment of the BRCA1 mRNA or gene,

it will be desirable to use a number of these probes to screen the whole mRNA sequence for mismatches.

In similar fashion, DNA probes can be used to detect mismatches, through enzymatic or chemical cleavage. See, e.g., Cotton *et al.*, 1988; Shenk *et al.*, 1975; Novack *et al.*, 1986. Alternatively, mismatches can be detected by shifts in the electrophoretic mobility of mismatched duplexes relative to matched duplexes. See, e.g., Cariello, 1988. With either riboprobes or DNA probes, the cellular mRNA or DNA which might contain a mutation can be amplified using PCR (see below) before hybridization. Changes in DNA of the BRCA1 gene can also be detected using Southern hybridization, especially if the changes are gross rearrangements, such as deletions and insertions.

DNA sequences of the BRCA1 gene which have been amplified by use of PCR may also be screened using allele-specific probes. These probes are nucleic acid oligomers, each of which contains a region of the BRCA1 gene sequence harboring a known mutation. For example, one oligomer may be about 30 nucleotides in length, corresponding to a portion of the BRCA1 gene sequence. By use of a battery of such allele-specific probes, PCR amplification products can be screened to identify the presence of a previously identified mutation in the BRCA1 gene. Hybridization of allele-specific probes with amplified BRCA1 sequences can be performed, for example, on a nylon filter. Hybridization to a particular probe under stringent hybridization conditions indicates the presence of the same mutation in the tumor tissue as in the allele-specific probe.

The most definitive test for mutations in a candidate locus is to directly compare genomic BRCA1 sequences from cancer patients with those from a control population. Alternatively, one could sequence messenger RNA after amplification, e.g., by PCR, thereby eliminating the necessity of determining the exon structure of the candidate gene.

Mutations from cancer patients falling outside the coding region of BRCA1 can be detected by examining the non-coding regions, such as introns and regulatory sequences near or within the BRCA1 gene. An early indication that mutations in noncoding regions are important may come from Northern blot experiments that reveal messenger RNA molecules of abnormal size or abundance in cancer patients as compared to control individuals.

Alteration of BRCA1 mRNA expression can be detected by any techniques known in the art. These include Northern blot analysis, PCR amplification and RNase protection. Diminished mRNA expression indicates an alteration of the wild-type BRCA1 gene. Alteration of wild-type BRCA1

genes can also be detected by screening for alteration of wild-type BRCA1 protein. For example, monoclonal antibodies immunoreactive with BRCA1 can be used to screen a tissue. Lack of cognate antigen would indicate a BRCA1 mutation. Antibodies specific for products of mutant alleles could also be used to detect mutant BRCA1 gene product. Such immunological assays can be done in any convenient formats known in the art. These include Western blots, immunohistochemical assays and ELISA assays. Any means for detecting an altered BRCA1 protein can be used to detect alteration of wild-type BRCA1 genes. Functional assays, such as protein binding determinations, can be used. In addition, assays can be used which detect BRCA1 biochemical function. Finding a mutant BRCA1 gene product indicates alteration of a wild-type BRCA1 gene.

Mutant BRCA1 genes or gene products can also be detected in other human body samples, such as serum, stool, urine and sputum. The same techniques discussed above for detection of mutant BRCA1 genes or gene products in tissues can be applied to other body samples. Cancer cells are sloughed off from tumors and appear in such body samples. In addition, the BRCA1 gene product itself may be secreted into the extracellular space and found in these body samples even in the absence of cancer cells. By screening such body samples, a simple early diagnosis can be achieved for many types of cancers. In addition, the progress of chemotherapy or radiotherapy can be monitored more easily by testing such body samples for mutant BRCA1 genes or gene products.

The methods of diagnosis of the present invention are applicable to any tumor in which BRCA1 has a role in tumorigenesis. The diagnostic method of the present invention is useful for clinicians, so they can decide upon an appropriate course of treatment.

The primer pairs of the present invention are useful for determination of the nucleotide sequence of a particular BRCA1 allele using PCR. The pairs of single-stranded DNA primers can be annealed to sequences within or surrounding the BRCA1 gene on chromosome 17q21 in order to prime amplifying DNA synthesis of the BRCA1 gene itself. A complete set of these primers allows synthesis of all of the nucleotides of the BRCA1 gene coding sequences, i.e., the exons. The set of primers preferably allows synthesis of both intron and exon sequences. Allele-specific primers can also be used. Such primers anneal only to particular BRCA1 mutant alleles, and thus will only amplify a product in the presence of the mutant allele as a template.

In order to facilitate subsequent cloning of amplified sequences, primers may have restriction enzyme site sequences appended to their 5' ends. Thus, all nucleotides of the primers are derived from BRCA1 sequences or sequences adjacent to BRCA1, except for the few nucleotides necessary to form a restriction enzyme site. Such enzymes and sites are well known in the art. The primers

themselves can be synthesized using techniques which are well known in the art. Generally, the primers can be made using oligonucleotide synthesizing machines which are commercially available. Given the sequence of the BRCA1 open reading frame shown in SEQ ID NO:1, design of particular primers is well within the skill of the art.

5 The nucleic acid probes provided by the present invention are useful for a number of purposes. They can be used in Southern hybridization to genomic DNA and in the RNase protection method for detecting point mutations already discussed above. The probes can be used to detect PCR amplification products. They may also be used to detect mismatches with the BRCA1 gene or mRNA using other techniques.

10 It has been discovered that individuals with the wild-type BRCA1 gene do not have cancer which results from the BRCA1 allele. However, mutations which interfere with the function of the BRCA1 protein are involved in the pathogenesis of cancer. Thus, the presence of an altered (or a mutant) BRCA1 gene which produces a protein having a loss of function, or altered function, directly correlates to a increased risk of cancer. In order to detect a BRCA1 gene mutation, a biological sample is prepared and analyzed for a difference between the sequence of the BRCA1 allele being
15 analyzed and the sequence of the wild-type BRCA1 allele. Mutant BRCA1 alleles can be initially identified by any of the techniques described above. The mutant alleles are then sequenced to identify the specific mutation of the particular mutant allele. Alternatively, mutant BRCA1 alleles can be initially identified by identifying mutant (altered) BRCA1 proteins, using conventional techniques. The mutant alleles are then sequenced to identify the specific mutation for each allele. The mutations, especially those which lead to an altered function of the BRCA1 protein, are then used for the diagnostic and prognostic methods of the present invention.

Definitions

The present invention employs the following definitions:

25 **"Amplification of Polynucleotides"** utilizes methods such as the polymerase chain reaction (PCR), ligation amplification (or ligase chain reaction, LCR) and amplification methods based on the use of Q-beta replicase. These methods are well known and widely practiced in the art. See, e.g., U.S. Patents 4,683,195 and 4,683,202 and Innis *et al.*, 1990 (for PCR); and Wu *et al.*, 1989 (for LCR). Reagents and hardware for conducting PCR are commercially available. Primers useful to
30 amplify sequences from the BRCA1 region are preferably complementary to, and hybridize specifically to sequences in the BRCA1 region or in regions that flank a target region therein.

BRCA1 sequences generated by amplification may be sequence directly. Alternatively, but less desirably, the amplified sequence(s) may be cloned prior to sequence analysis. A method for the direct cloning and sequence analysis of enzymatically amplified genomic segments has been described by Scharf, 1986.

5 "**Analyte polynucleotide**" and "**analyte strand**" refer to a single- or double-stranded polynucleotide which is suspected of containing a target sequence, and which may be present in a variety of types of samples, including biological samples.

10 "**Antibodies.**" The present invention also provides polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, which are capable of specifically binding to the BRCA1 polypeptides and fragments thereof or to polynucleotide sequences from the BRCA1 region, particularly from the BRCA1 locus or a portion thereof. The term "**antibody**" is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Polypeptides may be prepared synthetically in a peptide synthesizer and coupled to a carrier molecule (e.g., keyhole limpet hemocyanin) and injected over
15 several months into rabbits. Rabbit sera is tested for immunoreactivity to the BRCA1 polypeptide or fragment. Monoclonal antibodies may be made by injecting mice with the protein polypeptides, fusion proteins or fragments thereof. Monoclonal antibodies will be screened by ELISA and tested for specific immunoreactivity with BRCA1 polypeptide or fragments thereof. See, Harlow & Lane, 1988. These antibodies will be useful in assays as well as pharmaceuticals.

20 Once a sufficient quantity of desired polypeptide has been obtained, it may be used for various purposes. A typical use is the production of antibodies specific for binding. These antibodies may be either polyclonal or monoclonal, and may be produced by *in vitro* or *in vivo* techniques well known in the art. For production of polyclonal antibodies, an appropriate target immune system, typically mouse or rabbit, is selected. Substantially purified antigen is presented to the immune
25 system in a fashion determined by methods appropriate for the animal and by other parameters well known to immunologists. Typical sites for injection are in footpads, intramuscularly, intraperitoneally, or intradermally. Of course, other species may be substituted for mouse or rabbit. Polyclonal antibodies are then purified using techniques known in the art, adjusted for the desired specificity.

30 An immunological response is usually assayed with an immunoassay. Normally, such immunoassays involve some purification of a source of antigen, for example, that produced by the

same cells and in the same fashion as the antigen. A variety of immunoassay methods are well known in the art. See, e.g., Harlow & Lane, 1988, or Goding, 1986.

Monoclonal antibodies with affinities of 10^{-8} M⁻¹ or preferably 10^{-9} to 10^{-10} M⁻¹ or stronger will typically be made by standard procedures as described, e.g., in Harlow & Lane, 1988 or Goding, 1986. Briefly, appropriate animals will be selected and the desired immunization protocol followed. After the appropriate period of time, the spleens of such animals are excised and individual spleen cells fused, typically, to immortalized myeloma cells under appropriate selection conditions. Thereafter, the cells are clonally separated and the supernatants of each clone tested for their production of an appropriate antibody specific for the desired region of the antigen.

Other suitable techniques involve *in vitro* exposure of lymphocytes to the antigenic polypeptides, or alternatively, to selection of libraries of antibodies in phage or similar vectors. See Huse *et al.*, 1989. The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, polypeptides and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, magnetic particles and the like. Patents teaching the use of such labels include U.S. Patents 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced (see U.S. Patent 4,816,567).

"Binding partner" refers to a molecule capable of binding a ligand molecule with high specificity, as for example, an antigen and an antigen-specific antibody or an enzyme and its inhibitor. In general, the specific binding partners must bind with sufficient affinity to immobilize the analyte copy/ complementary strand duplex (in the case of polynucleotide hybridization) under the isolation conditions. Specific binding partners are known in the art and include, for example, biotin and avidin or streptavidin, IgG and protein A, the numerous, known receptor-ligand couples, and complementary polynucleotide strands. In the case of complementary polynucleotide binding partners, the partners are normally at least about 15 bases in length, and may be at least 40 bases in length. The polynucleotides may be composed of DNA, RNA, or synthetic nucleotide analogs.

A **"biological sample"** refers to a sample of tissue or fluid suspected of containing an analyte polynucleotide or polypeptide from an individual including, but not limited to, e.g., plasma, serum, spinal fluid, lymph fluid, the external sections of the skin, respiratory, intestinal, and genitourinary

tracts, tears, saliva, blood cells, tumors, organs, tissue and samples of *in vitro* cell culture constituents.

As used herein, the terms "**diagnosing**" or "**prognosing**," as used in the context of neoplasia, are used to indicate 1) the classification of lesions as neoplasia, 2) the determination of the severity of the neoplasia, or 3) the monitoring of the disease progression, prior to, during and after treatment.

"Encode". A polynucleotide is said to "encode" a polypeptide if, in its native state or when manipulated by methods well known to those skilled in the art, it can be transcribed and/or translated to produce the mRNA for and/or the polypeptide or a fragment thereof. The anti-sense strand is the complement of such a nucleic acid, and the encoding sequence can be deduced therefrom.

"Isolated" or "**substantially pure**". An "isolated" or "substantially pure" nucleic acid (e.g., an RNA, DNA or a mixed polymer) is one which is substantially separated from other cellular components which naturally accompany a native human sequence or protein, e.g., ribosomes, polymerases, many other human genome sequences and proteins. The term embraces a nucleic acid sequence or protein which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates and chemically synthesized analogs or analogs biologically synthesized by heterologous systems.

"BRCA1 Allele" refers to normal alleles of the BRCA1 locus as well as alleles carrying variations that predispose individuals to develop cancer of many sites including, for example, breast, ovarian, colorectal and prostate cancer. Such predisposing alleles are also called "**BRCA1 susceptibility alleles**".

"BRCA1 Locus," "BRCA1 gene," "BRCA1 Nucleic Acids" or "BRCA1 Polynucleotide" refer to polynucleotides, all of which are in the BRCA1 region, that are likely to be expressed in normal tissue, certain alleles of which predispose an individual to develop breast, ovarian, colorectal and prostate cancers. Mutations at the BRCA1 locus may be involved in the initiation and/or progression of other types of tumors. The locus is indicated in part by mutations that predispose individuals to develop cancer. These mutations fall within the BRCA1 region described *infra*. The BRCA1 locus is intended to include coding sequences, intervening sequences and regulatory elements controlling transcription and/or translation. The BRCA1 locus is intended to include all allelic variations of the DNA sequence.

These terms, when applied to a nucleic acid, refer to a nucleic acid which encodes a BRCA1 polypeptide, fragment, homolog or variant, including, e.g., protein fusions or deletions. The nucleic acids of the present invention will possess a sequence which is either derived from, or substantially

similar to a natural BRCA1-encoding gene or one having substantial homology with a natural BRCA1-encoding gene or a portion thereof. The coding sequence for a BRCA1 polypeptide is shown in SEQ ID NO:1, with the amino acid sequence shown in SEQ ID NO:2.

5 The polynucleotide compositions of this invention include RNA, cDNA, genomic DNA, synthetic forms, and mixed polymers, both sense and antisense strands, and may be chemically or biochemically modified or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those skilled in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, carbamates, etc.), charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), pendant moieties (e.g., polypeptides), intercalators (e.g., acridine, psoralen, etc.), chelators, alkylators, and modified linkages (e.g., alpha anomeric nucleic acids, etc.). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

10 The present invention provides recombinant nucleic acids comprising all or part of the BRCA1 region. The recombinant construct may be capable of replicating autonomously in a host cell. Alternatively, the recombinant construct may become integrated into the chromosomal DNA of the host cell. Such a recombinant polynucleotide comprises a polynucleotide of genomic, cDNA, semi-synthetic, or synthetic origin which, by virtue of its origin or manipulation, 1) is not associated with all or a portion of a polynucleotide with which it is associated in nature; 2) is linked to a polynucleotide other than that to which it is linked in nature; or 3) does not occur in nature.

20 Therefore, recombinant nucleic acids comprising sequences otherwise not naturally occurring are provided by this invention. Although the wild-type sequence may be employed, it will often be altered, e.g., by deletion, substitution or insertion.

25 cDNA or genomic libraries of various types may be screened as natural sources of the nucleic acids of the present invention, or such nucleic acids may be provided by amplification of sequences resident in genomic DNA or other natural sources, e.g., by PCR. The choice of cDNA libraries normally corresponds to a tissue source which is abundant in mRNA for the desired proteins. Phage libraries are normally preferred, but other types of libraries may be used. Clones of a library are

30

spread onto plates, transferred to a substrate for screening, denatured and probed for the presence of desired sequences.

The DNA sequences used in this invention will usually comprise at least about five codons (15 nucleotides), more usually at least about 7-15 codons, and most preferably, at least about 35 codons. One or more introns may also be present. This number of nucleotides is usually about the minimal length required for a successful probe that would hybridize specifically with a BRCA1-encoding sequence.

Techniques for nucleic acid manipulation are described generally, for example, in Sambrook *et al.*, 1989 or Ausubel *et al.*, 1992. Reagents useful in applying such techniques, such as restriction enzymes and the like, are widely known in the art and commercially available from such vendors as New England BioLabs, Boehringer Mannheim, Amersham, Promega Biotec, U. S. Biochemicals, New England Nuclear, and a number of other sources. The recombinant nucleic acid sequences used to produce fusion proteins of the present invention may be derived from natural or synthetic sequences. Many natural gene sequences are obtainable from various cDNA or from genomic libraries using appropriate probes. See, GenBank, National Institutes of Health.

"BRCA1 Region" refers to a portion of human chromosome 17q21 bounded by the markers tdj1474 and U5R. This region contains the BRCA1 locus, including the BRCA1 gene.

As used herein, the terms **"BRCA1 locus," "BRCA1 allele"** and **"BRCA1 region"** all refer to the double-stranded DNA comprising the locus, allele, or region, as well as either of the single-stranded DNAs comprising the locus, allele or region.

As used herein, a **"portion"** of the BRCA1 locus or region or allele is defined as having a minimal size of at least about eight nucleotides, or preferably about 15 nucleotides, or more preferably at least about 25 nucleotides, and may have a minimal size of at least about 40 nucleotides.

"BRCA1 protein" or **"BRCA1 polypeptide"** refer to a protein or polypeptide encoded by the BRCA1 locus, variants or fragments thereof. The term "polypeptide" refers to a polymer of amino acids and its equivalent and does not refer to a specific length of the product; thus, peptides, oligopeptides and proteins are included within the definition of a polypeptide. This term also does not refer to, or exclude modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations, and the like. Included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), polypeptides with substituted linkages as well as other modifications known in the art, both

naturally and non-naturally occurring. Ordinarily, such polypeptides will be at least about 50% homologous to the native BRCA1 sequence, preferably in excess of about 90%, and more preferably at least about 95% homologous. Also included are proteins encoded by DNA which hybridize under high or low stringency conditions, to BRCA1-encoding nucleic acids and closely related polypeptides or proteins retrieved by antisera to the BRCA1 protein(s).

The length of polypeptide sequences compared for homology will generally be at least about 16 amino acids, usually at least about 20 residues, more usually at least about 24 residues, typically at least about 28 residues, and preferably more than about 35 residues.

"Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression.

"Probes". Polynucleotide polymorphisms associated with BRCA1 alleles which predispose to certain cancers or are associated with most cancers are detected by hybridization with a polynucleotide probe which forms a stable hybrid with that of the target sequence, under stringent to moderately stringent hybridization and wash conditions. If it is expected that the probes will be perfectly complementary to the target sequence, stringent conditions will be used. Hybridization stringency may be lessened if some mismatching is expected, for example, if variants are expected with the result that the probe will not be completely complementary. Conditions are chosen which rule out nonspecific/adventitious bindings, that is, which minimize noise. Since such indications identify neutral DNA polymorphisms as well as mutations, these indications need further analysis to demonstrate detection of a BRCA1 susceptibility allele.

Probes for BRCA1 alleles may be derived from the sequences of the BRCA1 region or its cDNAs. The probes may be of any suitable length, which span all or a portion of the BRCA1 region, and which allow specific hybridization to the BRCA1 region. If the target sequence contains a sequence identical to that of the probe, the probes may be short, e.g., in the range of about 8-30 base pairs, since the hybrid will be relatively stable under even stringent conditions. If some degree of mismatch is expected with the probe, i.e., if it is suspected that the probe will hybridize to a variant region, a longer probe may be employed which hybridizes to the target sequence with the requisite specificity.

The probes will include an isolated polynucleotide attached to a label or reporter molecule and may be used to isolate other polynucleotide sequences, having sequence similarity by standard methods. For techniques for preparing and labelling probes see, e.g., Sambrook *et al.*, 1989 or

Ausebel *et al.*, 1992. Other similar polynucleotides may be selected by using homologous polynucleotides. Alternatively, polynucleotides encoding these or similar polypeptides may be synthesized or selected by use of the redundancy in the genetic code. Various codon substitutions may be introduced, e.g., by silent changes (thereby producing various restriction sites) or to optimize expression for a particular system. Mutations may be introduced to modify the properties of the polypeptide, perhaps to change ligand-binding affinities, interchain affinities, or the polypeptide degradation or turnover rate.

Probes comprising synthetic oligonucleotides or other polynucleotides of the present invention may be derived from naturally occurring or recombinant single- or double-stranded polynucleotides, or be chemically synthesized. Probes may also be labelled by nick translation, Klenow fill-in reaction, or other methods known in the art.

Portions of the polynucleotide sequence having at least about eight nucleotides, usually at least about 15 nucleotides, and fewer than about 6 Kb, usually fewer than about 1.0 Kb, from a polynucleotide sequence encoding BRCA1 are preferred as probes. The probes may also be used to determine whether mRNA encoding BRCA1 is present in a cell or tissue.

"Protein modifications or fragments" are provided by the present invention for BRCA1 polypeptides or fragments thereof which are substantially homologous to primary structural sequence but which include, e.g., *in vivo* or *in vitro* chemical and biochemical modifications or which incorporate unusual amino acids. Such modifications include, for example, acetylation, carboxylation, phosphorylation, glycosylation, ubiquitination, labelling, e.g., with radionuclides, and various enzymatic modifications, as will be readily appreciated by those well skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes are well known in the art, and include radioactive isotopes such as ^{32}P , ligands, which bind to labeled antiligands (e.g., antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands which can serve as specific binding pair members for a labeled ligand. The choice of label depends on the sensitivity required, ease of conjugation with the primer, stability requirements, and available instrumentation. Methods of labelling polypeptides are well known in the art. See, e.g., Sambrook *et al.*, 1989 or Ausebel *et al.*, 1992.

Besides substantially full-length polypeptides, the present invention provides for biologically active fragments of the polypeptides. Significant biological activities include ligand-binding, immunological activity and other biological activities characteristic of BRCA1 polypeptides. Immunological activities include both immunogenic function in a target immune system, as well as

sharing of immunological epitopes for binding, serving as either a competitor or substitute antigen for an epitope of the BRCA1 protein. As used herein, "epitope" refers to an antigenic determinant of a polypeptide. An epitope could comprise three amino acids in a spatial conformation which is unique to the epitope. Generally, an epitope consists of at least five such amino acids, and more usually consists of at least 8-10 such amino acids. Methods of determining the spatial conformation of such amino acids are known in the art.

For immunological purposes, tandem-repeat polypeptide segments may be used as immunogens, thereby producing highly antigenic proteins. Alternatively, such polypeptides will serve as highly efficient competitors for specific binding. Production of antibodies specific for BRCA1 polypeptides or fragments thereof is described below.

The present invention also provides for fusion polypeptides, comprising BRCA1 polypeptides and fragments. Homologous polypeptides may be fusions between two or more BRCA1 polypeptide sequences or between the sequences of BRCA1 and a related protein. Likewise, heterologous fusions may be constructed which would exhibit a combination of properties or activities of the derivative proteins. For example, ligand-binding or other domains may be "swapped" between different new fusion polypeptides or fragments. Such homologous or heterologous fusion polypeptides may display, for example, altered strength or specificity of binding. Fusion partners include immunoglobulins, bacterial β -galactosidase, trpE, protein A, β -lactamase, alpha amylase, alcohol dehydrogenase and yeast alpha mating factor. See, e.g., Godowski *et al.*, 1988.

Fusion proteins will typically be made by either recombinant nucleic acid methods, as described below, or may be chemically synthesized. Techniques for the synthesis of polypeptides are described, for example, in Merrifield, 1963.

"Protein purification" refers to various methods for the isolation of the BRCA1 polypeptides from other biological material, such as from cells transformed with recombinant nucleic acids encoding BRCA1, and are well known in the art. For example, such polypeptides may be purified by immunoaffinity chromatography employing, e.g., the antibodies provided by the present invention. Various methods of protein purification are well known in the art, and include those described in Deutscher, 1990 and Scopes, 1982.

The terms **"isolated"**, **"substantially pure"**, and **"substantially homogenous"** are used interchangeably to describe a protein or polypeptide which has been separated from components which accompany it in its natural state. A monomeric protein is substantially pure when at least

about 60 to 75% of a sample exhibits a single polypeptide sequence. A substantially pure protein will typically comprise about 60 to 90% W/W of a protein sample, more usually about 95%, and preferably will be over about 99% pure. Protein purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis or a protein sample, followed by visualizing a single polypeptide band upon staining the gel. For certain purposes, higher resolution may be provided by using HPLC or other means well known in the art for purification utilized.

A BRCA1 protein is substantially free of naturally associated components when it is separated from the native contaminants which accompany it in its natural state. Thus, a polypeptide which is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be substantially free from its naturally associated components. A protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art.

A polypeptide produced as an expression product of an isolated and manipulated genetic sequence is an "isolated polypeptide," as used herein, even if expressed in a homologous cell type. Synthetically made forms or molecules expressed by heterologous cells are inherently isolated molecules.

"Recombinant nucleic acid" is a nucleic acid which is not naturally occurring, or which is made by the artificial combination of two otherwise separated segments of sequence. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques. Such is usually done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a desired combination of functions.

"Regulatory sequences" refers to those sequences normally within 100 Kb of the coding region of a locus, but they may also be more distant from the coding region, which affect the expression of the gene (including transcription of the gene, and translation, splicing, stability or the like of the messenger RNA).

"Substantial homology or similarity". A nucleic acid or fragment thereof is "substantially homologous" ("or substantially similar") to another if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is

nucleotide sequence identity in at least about 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80%, preferably at least about 90%, and more preferably at least about 95-98% of the nucleotide bases.

Alternatively, substantial homology or (similarity) exists when a nucleic acid or fragment thereof will hybridize to another nucleic acid (or a complementary strand thereof) under selective hybridization conditions, to a strand, or to its complement. Selectivity of hybridization exists when hybridization which is substantially more selective than total lack of specificity occurs. Typically, selective hybridization will occur when there is at least about 55% homology over a stretch of at least about 14 nucleotides, preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90%. See, Kanehisa, 1984. The length of homology comparison, as described, may be over longer stretches, and in certain embodiments will often be over a stretch of at least about nine nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 32 nucleotides, and preferably at least about 36 or more nucleotides.

Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. Stringent temperature conditions will generally include temperatures in excess of 30°C, typically in excess of 37°C, and preferably in excess of 45°C. Stringent salt conditions will ordinarily be less than 1000 mM, typically less than 500 mM, and preferably less than 200 mM. However, the combination of parameters is much more important than the measure of any single parameter. See, e.g., Wetmur & Davidson, 1968.

Probe sequences may also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization conditions are well known in the art.

The terms "**substantial homology**" or "**substantial identity**", when referring to polypeptides, indicate that the polypeptide or protein in question exhibits at least about 30% identity with an entire naturally-occurring protein or a portion thereof, usually at least about 70% identity, and preferably at least about 95% identity.

"**Substantially similar function**" refers to the function of a modified nucleic acid or a modified protein, with reference to the wild-type BRCA1 nucleic acid or wild-type BRCA1 polypeptide. The modified polypeptide will be substantially homologous to the wild-type BRCA1 polypeptide and will

have substantially the same function. The modified polypeptide may have an altered amino acid sequence and/or may contain modified amino acids. In addition to the similarity of function, the modified polypeptide may have other useful properties, such as a longer half-life. The similarity of function (activity) of the modified polypeptide may be substantially the same as the activity of the wild-type BRCA1 polypeptide. Alternatively, the similarity of function (activity) of the modified polypeptide may be higher than the activity of the wild-type BRCA1 polypeptide. The modified polypeptide is synthesized using conventional techniques, or is encoded by a modified nucleic acid and produced using conventional techniques. The modified nucleic acid is prepared by conventional techniques. A nucleic acid with a function substantially similar to the wild-type BRCA1 gene function produces the modified protein described above.

Homology, for polypeptides, is typically measured using sequence analysis software. See, e.g., the Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 910 University Avenue, Madison, Wisconsin 53705. Protein analysis software matches similar sequences using measure of homology assigned to various substitutions, deletions, substitutions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

A polypeptide "**fragment**," "**portion**" or "**segment**" is a stretch of amino acid residues of at least about five to seven contiguous amino acids, often at least about seven to nine contiguous amino acids, typically at least about nine to 13 contiguous amino acids and, most preferably, at least about 20 to 30 or more contiguous amino acids.

The polypeptides of the present invention, if soluble, may be coupled to a solid-phase support, e.g., nitrocellulose, nylon, column packing materials (e.g., Sepharose beads), magnetic beads, glass wool, plastic, metal, polymer gels, cells, or other substrates. Such supports may take the form, for example, of beads, wells, dipsticks, membranes.

"**Target region**" refers to a region of the nucleic acid which is amplified and/or detected. The term "**target sequence**" refers to a sequence with which a probe or primer will form a stable hybrid under desired conditions.

The practice of the present invention employs, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, genetics, and immunology. See, e.g., Maniatis *et al.*, 1982; Sambrook *et al.*, 1989; Ausubel *et al.*, 1992; Glover, 1985; Anand, 1992; Guthrie & Fink, 1991. A general discussion of techniques and materials for

human gene mapping, including mapping of human chromosome 17q, is provided, e.g., in White & Lalouel, 1988.

Preparation of recombinant or chemically synthesized
nucleic acids; vectors, transformation, host cells

5 Large amounts of the polynucleotides of the present invention may be produced by replication in a suitable host cell. Natural or synthetic polynucleotide fragments coding for a desired fragment will be incorporated into recombinant polynucleotide constructs, usually DNA constructs, capable of introduction into and replication in a prokaryotic or eukaryotic cell. Usually the polynucleotide constructs will be suitable for replication in a unicellular host, such as yeast or bacteria, but may also be intended for introduction to (with and without integration within the genome) cultured mammalian or plant or other eukaryotic cell lines. The purification of nucleic acids produced by the methods of the present invention are described, e.g., in Sambrook *et al.*, 1989 or Ausebel *et al.*, 1992.

10 The polynucleotides of the present invention may also be produced by chemical synthesis, e.g., by the phosphoramidite method described by Beaucage & Carruthers, 1981 or the triester method according to Matteucci *et al.*, 1981, and may be performed on commercial, automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single-stranded product of chemical synthesis either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

15 20 Polynucleotide constructs prepared for introduction into a prokaryotic or eukaryotic host may comprise a replication system recognized by the host, including the intended polynucleotide fragment encoding the desired polypeptide, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the polypeptide encoding segment. Expression vectors may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Secretion signals may also be included where appropriate, whether from a native BRCA1 protein or from other receptors or from secreted polypeptides of the same or related species, which allow the protein to cross and/or lodge in cell membranes, and thus attain its functional topology, or be secreted from the cell. Such

25 30

vectors may be prepared by means of standard recombinant techniques well known in the art and discussed, for example, in Sambrook *et al.*, 1989 or Ausebel *et al.* 1992.

An appropriate promoter and other necessary vector sequences will be selected so as to be functional in the host, and may include, when appropriate, those naturally associated with BRCA1 genes. Examples of workable combinations of cell lines and expression vectors are described in Sambrook *et al.*, 1989 or Ausebel *et al.*, 1992; see also, e.g., Metzger *et al.*, 1988. Many useful vectors are known in the art and may be obtained from such vendors as Stratagene, New England Biolabs, Promega Biotech, and others. Promoters such as the trp, lac and phage promoters, tRNA promoters and glycolytic enzyme promoters may be used in prokaryotic hosts. Useful yeast promoters include promoter regions for metallothionein, 3-phosphoglycerate kinase or other glycolytic enzymes such as enolase or glyceraldehyde-3-phosphate dehydrogenase, enzymes responsible for maltose and galactose utilization, and others. Vectors and promoters suitable for use in yeast expression are further described in Hitzeman *et al.*, EP 73,675A. Appropriate non-native mammalian promoters might include the early and late promoters from SV40 (Fiers *et al.*, 1978) or promoters derived from murine molony leukemia virus, mouse tumor virus, avian sarcoma viruses, adenovirus II, bovine papilloma virus or polyoma. In addition, the construct may be joined to an amplifiable gene (e.g., DHFR) so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also *Enhancers and Eukaryotic Gene Expression*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1983).

While such expression vectors may replicate autonomously, they may also replicate by being inserted into the genome of the host cell, by methods well known in the art.

Expression and cloning vectors will likely contain a selectable marker, a gene encoding a protein necessary for survival or growth of a host cell transformed with the vector. The presence of this gene ensures growth of only those host cells which express the inserts. Typical selection genes encode proteins that a) confer resistance to antibiotics or other toxic substances, e.g. ampicillin, neomycin, methotrexate, etc.; b) complement auxotrophic deficiencies, or c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*. The choice of the proper selectable marker will depend on the host cell, and appropriate markers for different hosts are well known in the art.

The vectors containing the nucleic acids of interest can be transcribed *in vitro*, and the resulting RNA introduced into the host cell by well-known methods, e.g., by injection (see, Kubo *et al.*, 1988), or the vectors can be introduced directly into host cells by methods well known in the art, which vary

depending on the type of cellular host, including electroporation; transfection employing calcium chloride, rubidium chloride calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment; lipofection; infection (where the vector is an infectious agent, such as a retroviral genome); and other methods. See generally, Sambrook *et al.*, 1989 and Ausubel *et al.*, 1992. The introduction of the polynucleotides into the host cell by any method known in the art, including, *inter alia*, those described above, will be referred to herein as "transformation." The cells into which have been introduced nucleic acids described above are meant to also include the progeny of such cells.

Large quantities of the nucleic acids and polypeptides of the present invention may be prepared by expressing the BRCA1 nucleic acids or portions thereof in vectors or other expression vehicles in compatible prokaryotic or eukaryotic host cells. The most commonly used prokaryotic hosts are strains of *Escherichia coli*, although other prokaryotes, such as *Bacillus subtilis* or *Pseudomonas* may also be used.

Mammalian or other eukaryotic host cells, such as those of yeast, filamentous fungi, plant, insect, or amphibian or avian species, may also be useful for production of the proteins of the present invention. Propagation of mammalian cells in culture is *per se* well known. See, Kruse & Patterson, 1973. Examples of commonly used mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cells, and WI38, BHK, and COS cell lines, although it will be appreciated by the skilled practitioner that other cell lines may be appropriate, e.g., to provide higher expression, desirable glycosylation patterns, or other features.

Clones are selected by using markers depending on the mode of the vector construction. The marker may be on the same or a different DNA molecule, preferably the same DNA molecule. In prokaryotic hosts, the transformant may be selected, e.g., by resistance to ampicillin, tetracycline or other antibiotics. Production of a particular product based on temperature sensitivity may also serve as an appropriate marker.

Prokaryotic or eukaryotic cells transformed with the polynucleotides of the present invention will be useful not only for the production of the nucleic acids and polypeptides of the present invention, but also, for example, in studying the characteristics of BRCA1 polypeptides.

Antisense polynucleotide sequences are useful in preventing or diminishing the expression of the BRCA1 locus, as will be appreciated by those skilled in the art. For example, polynucleotide vectors containing all or a portion of the BRCA1 locus or other sequences from the BRCA1 region (particularly those flanking the BRCA1 locus) may be placed under the control of a promoter in an

antisense orientation and introduced into a cell. Expression of such an antisense construct within a cell will interfere with BRCA1 transcription and/or translation and/or replication.

The probes and primers based on the BRCA1 gene sequences disclosed herein are used to identify homologous BRCA1 gene sequences and proteins in other species. These BRCA1 gene sequences and proteins are used in the diagnostic/prognostic, therapeutic and drug screening methods described herein for the species from which they have been isolated.

Methods of Use: Nucleic Acid Diagnosis and Diagnostic Kits

In order to detect the presence of a BRCA1 allele predisposing an individual to cancer, a biological sample such as blood is prepared and analyzed for the presence or absence of susceptibility alleles of BRCA1. In order to detect the presence of neoplasia, the progression toward malignancy of a precursor lesion, or as a prognostic indicator, a biological sample of the lesion is prepared and analyzed for the presence or absence of mutant alleles of BRCA1. Results of these tests and interpretive information are returned to the health care provider for communication to the tested individual. Such diagnoses may be performed by diagnostic laboratories, or, alternatively, diagnostic kits are manufactured and sold to health care providers or to private individuals for self-diagnosis.

Initially, the screening method involves amplification of the relevant BRCA1 sequences, e.g., by PCR, followed by DNA sequence analysis.

When the probes are used to detect the presence of the target sequences (for example, in screening for cancer susceptibility), the biological sample to be analyzed, such as blood or serum, may be treated, if desired, to extract the nucleic acids. The sample nucleic acid may be prepared in various ways to facilitate detection of the target sequence; e.g. denaturation, restriction digestion, electrophoresis or dot blotting. The targeted region of the analyte nucleic acid usually must be at least partially single-stranded to form hybrids with the targeting sequence of the probe. If the sequence is naturally single-stranded, denaturation will not be required. However, if the sequence is double-stranded, the sequence will probably need to be denatured. Denaturation can be carried out by various techniques known in the art.

Analyte nucleic acid and probe are incubated under conditions which promote stable hybrid formation of the target sequence in the probe with the putative targeted sequence in the analyte. The region of the probes which is used to bind to the analyte can be made completely complementary to the targeted region of human chromosome 17q. Therefore, high stringency conditions are desirable in order to prevent false positives. However, conditions of high stringency are used only if the probes

are complementary to regions of the chromosome which are unique in the genome. The stringency of hybridization is determined by a number of factors during hybridization and during the washing procedure, including temperature, ionic strength, base composition, probe length, and concentration of formamide. These factors are outlined in, for example, Maniatis *et al.*, 1982 and Sambrook *et al.*, 1989. Under certain circumstances, the formation of higher order hybrids, such as triplexes, quadraplexes, etc., may be desired to provide the means of detecting target sequences.

Detection, if any, of the resulting hybrid is usually accomplished by the use of labeled probes. Alternatively, the probe may be unlabeled, but may be detectable by specific binding with a ligand which is labeled, either directly or indirectly. Suitable labels, and methods for labeling probes and ligands are known in the art, and include, for example, radioactive labels which may be incorporated by known methods (e.g., nick translation, random priming or kinasing), biotin, fluorescent groups, chemiluminescent groups (e.g., dioxetanes, particularly triggered dioxetanes), enzymes, antibodies and the like. Variations of this basic scheme are known in the art, and include those variations that facilitate separation of the hybrids to be detected from extraneous materials and/or that amplify the signal from the labeled moiety. A number of these variations are reviewed in, e.g., Matthews & Kricka, 1988; Landegren *et al.*, 1988; Mittlin, 1989, U.S. Patent 4,868,105, and in EPO Publication No. 225,807.

Methods of Use: Peptide Diagnosis and Diagnostic Kits

The neoplastic condition of lesions can also be detected on the basis of the alteration of wild-type BRCA1 polypeptide. Such alterations can be determined by sequence analysis in accordance with conventional techniques. More preferably, antibodies (polyclonal or monoclonal) are used to detect differences in, or the absence of BRCA1 peptides. In a preferred embodiment of the invention, antibodies will immunoprecipitate BRCA1 proteins from solution as well as react with BRCA1 protein on Western or immunoblots of polyacrylamide gels. In another preferred embodiment, antibodies will detect BRCA1 proteins in paraffin or frozen tissue sections, using immunocytochemical techniques. Techniques for raising and purifying antibodies are well known in the art, and any such techniques may be chosen to achieve the preparation of the invention.

Methods of Use: Drug Screening

This invention is particularly useful for screening compounds by using the BRCA1 polypeptide or binding fragment thereof in any of a variety of drug screening techniques.

The BRCA1 polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, or borne on a cell surface. One method of drug screening utilizes eucaryotic or procaryotic host cells which are stably transformed with recombinant polynucleotides expressing the polypeptide or fragment, preferably in competitive binding assays. Such cells, either
5 in viable or fixed form, can be used for standard binding assays. one may measure, for example, for the formation of complexes between a BRCA1 polypeptide or fragment and the agent being tested, or examine the degree to which the formation of a complex between a BRCA1 polypeptide or fragment and a known ligand is interfered with by the agent being tested.

Thus, the present invention provides methods of screening for drugs comprising contacting such
10 an agent with a BRCA1 polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the BRCA1 polypeptide or fragment, or (ii) for the presence of a complex between the BRCA1 polypeptide or fragment and a ligand, by methods well known in the art. In such competitive binding assays the BRCA1 polypeptide or fragment is typically labeled. Free BRCA1 polypeptide or fragment is separated from that present in a protein:protein complex, and
15 the amount of free (i.e., uncomplexed) label is a measure of the binding of the agent being tested to BRCA1 or its interference with BRCA1:ligand binding, respectively.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the BRCA1 polypeptides and is described in detail in Geysen, European Patent Application 84/03564, published on September 13, 1984. Briefly stated, large numbers of
20 different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with BRCA1 polypeptide and washed. Bound BRCA1 polypeptide is then detected by methods well known in the art.

Purified BRCA1 can be coated directly onto plates for use in the aforementioned drug screening techniques. However, non-neutralizing antibodies to the polypeptide can be used to capture antibodies
25 to immobilize the BRCA1 polypeptide on the solid phase.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of specifically binding the BRCA1 polypeptide compete with a test compound for binding to the BRCA1 polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic
30 determinants of the BRCA1 polypeptide.

A further technique for drug screening involves the use of host eukaryotic cell lines or cells (such as described above) which have a nonfunctional BRCA1 gene. These host cell lines or cells

are defective at the BRCA1 polypeptide level. The host cell lines or cells are grown in the presence of drug compound. The rate of growth of the host cells is measured to determine if the compound is capable of regulating the growth of BRCA1 defective cells.

Methods of Use: Rational Drug Design

5 The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g., enhance or interfere with the function of a polypeptide *in vivo*. See, e.g., Hodgson, 1991. In one approach, one first determines the three-dimensional structure of a protein of interest (e.g., BRCA1 polypeptide) or, for example, of the BRCA1-receptor or ligand complex, by x-ray crystallography, by computer modelling or most typically, by a combination of approaches. Less often, useful information regarding the structure of a polypeptide may be gained by modeling based on the structure of homologous proteins. An example of rational drug design is the development of HIV protease inhibitors (Erickson *et al.*, 1990). In addition, peptides (e.g., BRCA1 polypeptide) are analyzed by an alanine scan (Wells, 1991). In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

15
20 It is also possible to isolate a target-specific antibody, selected by a functional assay, and then to solve its crystal structure. In principle, this approach yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacore.

25
30 Thus, one may design drugs which have, e.g., improved BRCA1 polypeptide activity or stability or which act as inhibitors, agonists, antagonists, etc. of BRCA1 polypeptide activity. By virtue of the availability of cloned BRCA1 sequences, sufficient amounts of the BRCA1 polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, the knowledge of the BRCA1 protein sequence provided herein will guide those employing computer modeling techniques in place of, or in addition to x-ray crystallography.

Methods of Use: Gene Therapy

According to the present invention, a method is also provided of supplying wild-type BRCA1 function to a cell which carries mutant BRCA1 alleles. Supplying such a function should suppress neoplastic growth of the recipient cells. The wild-type BRCA1 gene or a part of the gene may be introduced into the cell in a vector such that the gene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. If a gene fragment is introduced and expressed in a cell carrying a mutant BRCA1 allele, the gene fragment should encode a part of the BRCA1 protein which is required for non-neoplastic growth of the cell. More preferred is the situation where the wild-type BRCA1 gene or a part thereof is introduced into the mutant cell in such a way that it recombines with the endogenous mutant BRCA1 gene present in the cell. Such recombination requires a double recombination event which results in the correction of the BRCA1 gene mutation. Vectors for introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector may be used. Methods for introducing DNA into cells such as electroporation, calcium phosphate co-precipitation and viral transduction are known in the art, and the choice of method is within the competence of the routineer. Cells transformed with the wild-type BRCA1 gene can be used as model systems to study cancer remission and drug treatments which promote such remission.

As generally discussed above, the BRCA1 gene or fragment, where applicable, may be employed in gene therapy methods in order to increase the amount of the expression products of such genes in cancer cells. Such gene therapy is particularly appropriate for use in both cancerous and pre-cancerous cells, in which the level of BRCA1 polypeptide is absent or diminished compared to normal cells. It may also be useful to increase the level of expression of a given BRCA1 gene even in those tumor cells in which the mutant gene is expressed at a "normal" level, but the gene product is not fully functional.

Gene therapy would be carried out according to generally accepted methods, for example, as described by Friedman, 1991. Cells from a patient's tumor would be first analyzed by the diagnostic methods described above, to ascertain the production of BRCA1 polypeptide in the tumor cells. A virus or plasmid vector (see further details below), containing a copy of the BRCA1 gene linked to expression control elements and capable of replicating inside the tumor cells, is prepared. Suitable vectors are known, such as disclosed in U.S. Patent 5,252,479 and PCT published application WO 93/07282. The vector is then injected into the patient, either locally at the site of the tumor or systemically (in order to reach any tumor cells that may have metastasized to other sites). If the

transfected gene is not permanently incorporated into the genome of each of the targeted tumor cells, the treatment may have to be repeated periodically.

Gene transfer systems known in the art may be useful in the practice of the gene therapy methods of the present invention. These include viral and nonviral transfer methods. A number of viruses have been used as gene transfer vectors, including papovaviruses (e.g., SV40, Madzak *et al.*, 1992), adenovirus (Berkner, 1992; Berkner *et al.*, 1988; Gorziglia and Kapikian, 1992; Quantin *et al.*, 1992; Rosenfeld *et al.*, 1992; Wilkinson *et al.*, 1992; Stratford-Perricaudet *et al.*, 1990), vaccinia virus (Moss, 1992), adeno-associated virus (Muzyczka, 1992; Ohi *et al.*, 1990), herpesviruses including HSV and EBV (Margolskee, 1992; Johnson *et al.*, 1992; Fink *et al.*, 1992; Breakfield and Geller, 1987; Freese *et al.*, 1990), and retroviruses of avian (Brandyopadhyay and Temin, 1984; Petropoulos *et al.*, 1992), murine (Miller, 1992; Miller *et al.*, 1985; Sorge *et al.*, 1984; Mann and Baltimore, 1985; Miller *et al.*, 1988), and human origin (Shimada *et al.*, 1991; Helseth *et al.*, 1990; Page *et al.*, 1990; Buchschacher and Panganiban, 1992). Most human gene therapy protocols have been based on disabled murine retroviruses.

Nonviral gene transfer methods known in the art include chemical techniques such as calcium phosphate coprecipitation (Graham and van der Eb, 1973; Pellicer *et al.*, 1980); mechanical techniques, for example microinjection (Anderson *et al.*, 1980; Gordon *et al.*, 1980; Brinster *et al.*, 1981; Constantini and Lacy, 1981); membrane fusion-mediated transfer via liposomes (Felgner *et al.*, 1987; Wang and Huang, 1989; Kaneda *et al.*, 1989; Stewart *et al.*, 1992; Nabel *et al.*, 1990; Lim *et al.*, 1992); and direct DNA uptake and receptor-mediated DNA transfer (Wolff *et al.*, 1990; Wu *et al.*, 1991; Zenke *et al.*, 1990; Wu *et al.*, 1989; Wolff *et al.*, 1991; Wagner *et al.*, 1990; Wagner *et al.*, 1991; Cotten *et al.*, 1990; Curiel *et al.*, 1991; Curiel *et al.*, 1991). Viral-mediated gene transfer can be combined with direct in vivo gene transfer using liposome delivery, allowing one to direct the viral vectors to the tumor cells and not into the surrounding nondividing cells. Alternatively, the retroviral vector producer cell line can be injected into tumors (Culver *et al.*, 1992). Injection of producer cells would then provide a continuous source of vector particles. This technique has been approved for use in humans with inoperable brain tumors.

In an approach which combines biological and physical gene transfer methods, plasmid DNA of any size is combined with a polylysine-conjugated antibody specific to the adenovirus hexon protein, and the resulting complex is bound to an adenovirus vector. The trimolecular complex is then used to infect cells. The adenovirus vector permits efficient binding, internalization, and degradation of the endosome before the coupled DNA is damaged.

Liposome/DNA complexes have been shown to be capable of mediating direct *in vivo* gene transfer. While in standard liposome preparations the gene transfer process is nonspecific, localized *in vivo* uptake and expression have been reported in tumor deposits, for example, following direct *in situ* administration (Nabel, 1992).

Gene transfer techniques which target DNA directly to breast and ovarian tissues, e.g., epithelial cells of the breast or ovaries, is preferred. Receptor-mediated gene transfer, for example, is accomplished by the conjugation of DNA (usually in the form of covalently closed supercoiled plasmid) to a protein ligand via polylysine. Ligands are chosen on the basis of the presence of the corresponding ligand receptors on the cell surface of the target cell/tissue type. One appropriate receptor/ligand pair may include the estrogen receptor and its ligand, estrogen (and estrogen analogues). These ligand-DNA conjugates can be injected directly into the blood if desired and are directed to the target tissue where receptor binding and internalization of the DNA-protein complex occurs. To overcome the problem of intracellular destruction of DNA, coinfection with adenovirus can be included to disrupt endosome function.

The therapy involves two steps which can be performed singly or jointly. In the first step, prepubescent females who carry a BRCA1 susceptibility allele are treated with a gene delivery vehicle such that some or all of their mammary ductal epithelial precursor cells receive at least one additional copy of a functional normal BRCA1 allele. In this step, the treated individuals have reduced risk of breast cancer to the extent that the effect of the susceptible allele has been countered by the presence of the normal allele. In the second step of a preventive therapy, predisposed young females, in particular women who have received the proposed gene therapeutic treatment, undergo hormonal therapy to mimic the effects on the breast of a full term pregnancy.

Methods of Use: Peptide Therapy

Peptides which have BRCA1 activity can be supplied to cells which carry mutant or missing BRCA1 alleles. The sequence of the BRCA1 protein is disclosed (SEQ ID NO:2). Protein can be produced by expression of the cDNA sequence in bacteria, for example, using known expression vectors. Alternatively, BRCA1 polypeptide can be extracted from BRCA1-producing mammalian cells. In addition, the techniques of synthetic chemistry can be employed to synthesize BRCA1 protein. Any of such techniques can provide the preparation of the present invention which comprises the BRCA1 protein. The preparation is substantially free of other human proteins. This is most readily accomplished by synthesis in a microorganism or *in vitro*.

Active BRCA1 molecules can be introduced into cells by microinjection or by use of liposomes, for example. Alternatively, some active molecules may be taken up by cells, actively or by diffusion. Extracellular application of the BRCA1 gene product may be sufficient to affect tumor growth. Supply of molecules with BRCA1 activity should lead to partial reversal of the neoplastic state. Other molecules with BRCA1 activity (for example, peptides, drugs or organic compounds) may also be used to effect such a reversal. Modified polypeptides having substantially similar function are also used for peptide therapy.

Methods of Use: Transformed Hosts

Similarly, cells and animals which carry a mutant BRCA1 allele can be used as model systems to study and test for substances which have potential as therapeutic agents. The cells are typically cultured epithelial cells. These may be isolated from individuals with BRCA1 mutations, either somatic or germline. Alternatively, the cell line can be engineered to carry the mutation in the BRCA1 allele, as described above. After a test substance is applied to the cells, the neoplastically transformed phenotype of the cell is determined. Any trait of neoplastically transformed cells can be assessed, including anchorage-independent growth, tumorigenicity in nude mice, invasiveness of cells, and growth factor dependence. Assays for each of these traits are known in the art.

Animals for testing therapeutic agents can be selected after mutagenesis of whole animals or after treatment of germline cells or zygotes. Such treatments include insertion of mutant BRCA1 alleles, usually from a second animal species, as well as insertion of disrupted homologous genes. Alternatively, the endogenous BRCA1 gene(s) of the animals may be disrupted by insertion or deletion mutation or other genetic alterations using conventional techniques. (Capecchi, 1989; Valancius and Smithies, 1991; Hasty *et al.*, 1991; Shinkai *et al.*, 1992; Mombaerts *et al.*, 1992; Philpott *et al.*, 1992; Snouwaert *et al.*, 1992; Donehower *et al.*, 1992) After test substances have been administered to the animals, the growth of tumors must be assessed. If the test substance prevents or suppresses the growth of tumors, then the test substance is a candidate therapeutic agent for the treatment of the cancers identified herein. These animal models provide an extremely important testing vehicle for potential therapeutic products.

The present invention is described by reference to the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner. Standard techniques well known in the art or the techniques specifically described below were utilized.

EXAMPLE 1

Ascertain and Study Kindreds Likely to Have
a 17q-Linked Breast Cancer Susceptibility Locus

Extensive cancer prone kindreds were ascertained from a defined population providing a large set of extended kindreds with multiple cases of breast cancer and many relatives available to study. The large number of meioses present in these large kindreds provided the power to detect whether the BRCA1 locus was segregating, and increased the opportunity for informative recombinants to occur within the small region being investigated. This vastly improved the chances of establishing linkage to the BRCA1 region, and greatly facilitated the reduction of the BRCA1 region to a manageable size, which permits identification of the BRCA1 locus itself.

Each kindred was extended through all available connecting relatives, and to all informative first degree relatives of each proband or cancer case. For these kindreds, additional breast cancer cases and individuals with cancer at other sites of interest (e.g. ovarian) who also appeared in the kindreds were identified through the tumor registry linked files. All breast cancers reported in the kindred which were not confirmed in the Utah Cancer Registry were researched. Medical records or death certificates were obtained for confirmation of all cancers. Each key connecting individual and all informative individuals were invited to participate by providing a blood sample from which DNA was extracted. We also sampled spouses and relatives of deceased cases so that the genotype of the deceased cases could be inferred from the genotypes of their relatives.

Ten kindreds which had three or more cancer cases with inferable genotypes were selected for linkage studies to 17q markers from a set of 29 kindreds originally ascertained from the linked databases for a study of proliferative breast disease and breast cancer (Skolnick *et al.*, 1990). The criterion for selection of these kindreds was the presence of two sisters or a mother and her daughter with breast cancer. Additionally, two kindreds which have been studied since 1980 as part of our breast cancer linkage studies (K1001, K9018), six kindreds ascertained from the linked databases for the presence of clusters of breast and/or ovarian cancer (K2019, K2073, K2079, K2080, K2039; K2082; and a self-referred kindred with early onset breast cancer (K2035) were included. These kindreds were investigated and expanded in our clinic in the manner described above. Table 1 displays the characteristics of these 19 kindreds which are the subject of subsequent examples. In Table 1, for each kindred the total number of individuals in our database, the number of typed individuals, and the minimum, median, and maximum age at diagnosis of breast/ovarian cancer are

reported. Kindreds are sorted in ascending order of median age at diagnosis of breast cancer. Four women diagnosed with both ovarian and breast cancer are counted in both categories.

169760 10190700

TABLE 1
Description of the 19 Kindreds

<u>KINDRED</u>	<u>No. of</u> <u>Individuals</u>		<u>Breast</u>				<u>Ovarian</u>			
	<u>Total</u>	<u>Sample</u>	<u># Aff.</u>	<u>Age at Dx</u>			<u># Aff.</u>	<u>Age at Dx</u>		
				<u>Min.</u>	<u>Med.</u>	<u>Max.</u>		<u>Min.</u>	<u>Med.</u>	<u>Max.</u>
1910	15	10	4	27	34	49	-	-	-	-
1001	133	98	13	28	37	64	-	-	-	-
2035	42	25	8	28	37	45	1	-	60	-
2027	21	11	4	34	38	41	-	-	-	-
9018	54	17	9	30	40	72	2	46	48	50
1925	50	27	4	39	42	53	-	-	-	-
1927	49	29	5	32	42	51	-	-	-	-
1911	28	21	7	28	42	76	-	-	-	-
1929	16	11	4	34	43	73	-	-	-	-
1901	35	19	10	31	44	76	-	-	-	-
2082	180	105	20	27	47	67	10	45	52	66
2019	42	19	10	42	53	79	-	-	-	-
1900	70	23	8	45	55	70	1	-	78	-
2080	264	74	22+	27	55	92	4	45	53	71
2073	57	29	9	35	57	80	-	-	-	-
1917	16	6	4	43	58	61	-	-	-	-
1920	22	14	3	62	63	68	-	-	-	-
2079	136	18	14	38	66	84	4	52	59	65
2039	87	40	14	44	68	88	4	41	51	75

+ Includes one case of male breast cancer.

EXAMPLE 2

Selection of Kindreds Which are Linked to Chromosome 17q
and Localization of BRCA1 to the Interval Mfd15 - Mfd188

For each sample collected in these 19 kindreds, DNA was extracted from blood (or in two cases from paraffin-embedded tissue blocks) using standard laboratory protocols. Genotyping in this study was restricted to short tandem repeat (STR) markers since, in general, they have high heterozygosity and PCR methods offer rapid turnaround while using very small amounts of DNA. To aid in this effort, four such STR markers on chromosome 17 were developed by screening a chromosome specific cosmid library for CA positive clones. Three of these markers localized to the long arm: (46E6, Easton *et al.*, 1993); (42D6, Easton *et al.*, 1993); 26C2 (D17S514, Oliphant *et al.*, 1993), while the other, 12G6, (D17S513, Oliphant *et al.*, 1991) localized to the short arm near the p53 tumor suppressor locus. Two of these, 42D6 and 46E6, were submitted to the Breast Cancer Linkage Consortium for typing of breast cancer families by investigators worldwide. Oligonucleotide sequences for markers not developed in our laboratory were obtained from published reports, or as part of the Breast Cancer Linkage Consortium, or from other investigators. All genotyping films were scored blindly with a standard lane marker used to maintain consistent coding of alleles. Key samples in the four kindreds presented here underwent duplicate typing for all relevant markers. All 19 kindreds have been typed for two polymorphic CA repeat markers: 42D6 (D17S588), a CA repeat isolated in our laboratory, and Mfd15 (D17S250), a CA repeat provided by J. Weber (Weber *et al.*, 1990). Several sources of probes were used to create genetic markers on chromosome 17, specifically chromosome 17 cosmid and lambda phage libraries created from sorted chromosomes by the Los Alamos National Laboratories (van Dilla *et al.*, 1986).

LOD scores for each kindred with these two markers (42D6, Mfd15) and a third marker, Mfd188 (D17S579, Hall *et al.* 1992), located roughly midway between these two markers, were calculated for two values of the recombination fraction, 0.001 and 0.1. (For calculation of LOD scores, see OH, 1985). Likelihoods were computed under the model derived by Claus *et al.*, 1991, which assumes an estimated gene frequency of 0.003, a lifetime risk in gene carriers of about 0.80, and population based age-specific risks for breast cancer in non-gene carriers. Allele frequencies for the three markers used for the LOD score calculations were calculated from our own laboratory typings of unrelated individuals in the CEPH panel (White *et al.*, 1988). Table 2 shows the results of the pairwise linkage analysis of each kindred with the three markers 42D6, Mfd188, and Mfd15.

TABLE 2

Pairwise Linkage Analysis of Kindreds

<u>KINDRED</u>	<u>Mfd15 (D17S250)</u> <u>Recombination</u>		<u>42D6 (D17S588)</u> <u>Recombination</u>		<u>Mfd188 (D17S579)</u> <u>Recombination</u>	
	0.001	0.1	0.001	0.1	0.001	0.1
1910	0.06	0.30	0.06	0.30	0.06	0.30
1001	-0.30	-0.09	NT	NT	-0.52	-0.19
2035	2.34	1.85	0.94	0.90	2.34	1.82
2027	-1.22	-0.33	-1.20	-0.42	-1.16	-0.33
9018	-0.54	-0.22	-0.17	-0.10	0.11	0.07
1925	1.08	0.79	0.55	0.38	-0.11	-0.07
1927	-0.41	0.01	-0.35	0.07	-0.44	-0.02
1911	-0.27	-0.13	-0.43	-0.23	0.49	0.38
1929	-0.49	-0.25	NT	NT	-0.49	-0.25
1901	1.50	1.17	0.78	0.57	0.65	0.37
2082	4.25	3.36	6.07	5.11	2.00	3.56
2019	-0.10	-0.01	-0.11	-0.05	-0.18	-0.10
1900	-0.14	-0.11	NT	NT	-0.12	-0.05
2080	-0.16	-0.04	0.76	0.74	-1.25	-0.58
2073	-0.41	-0.29	0.63	0.49	-0.23	-0.13
1917	-0.02	-0.02	NT	NT	-0.01	0.00
1920	-0.03	-0.02	NT	NT	0.00	0.00
2079	0.02	0.01	-0.01	-0.01	0.01	0.01
2039	-1.67	-0.83	0.12	0.59	-1.15	0.02

NT - Kindred not typed for Mfd188.

Using a criterion for linkage to 17q of a LOD score > 1.0 for at least one locus under the CASH model (Claus *et al.*, 1991), four of the 19 kindreds appeared to be linked to 17q (K1901, K1925, K2035, K2082). A number of additional kindreds showed some evidence of linkage but at this time could not be definitively assigned to the linked category. These included kindreds K1911, K2073, K2039, and K2080. Three of the 17q-linked kindreds had informative recombinants in this region and these are detailed below.

Kindred 2082 is the largest 17q-linked breast cancer family reported to date by any group. The kindred contains 20 cases of breast cancer, and ten cases of ovarian cancer. Two cases have both ovarian and breast cancer. The evidence of linkage to 17q for this family is overwhelming; the LOD score with the linked haplotype is over 6.0, despite the existence of three cases of breast cancer which appear to be sporadic, i.e., these cases share no part of the linked haplotype between Mfd15 and 42D6. These three sporadic cases were diagnosed with breast cancer at ages 46, 47, and 54. In smaller kindreds, sporadic cancers of this type greatly confound the analysis of linkage and the correct identification of key recombinants. The key recombinant in the 2082 kindred is a woman who developed ovarian cancer at age 45 whose mother and aunt had ovarian cancer at ages 58 and 66, respectively. She inherited the linked portion of the haplotype for both Mfd188 and 42D6 while inheriting unlinked alleles at Mfd15; this recombinant event placed BRCA1 distal to Mfd15.

K1901 is typical of early-onset breast cancer kindreds. The kindred contains 10 cases of breast cancer with a median age at diagnosis of 43.5 years of age; four cases were diagnosed under age 40. The LOD score for this kindred with the marker 42D6 is 1.5, resulting in a posterior probability of 17q-linkage of 0.96. Examination of haplotypes in this kindred identified a recombinant haplotype in an obligate male carrier and his affected daughter who was diagnosed with breast cancer at age 45. Their linked allele for marker Mfd15 differs from that found in all other cases in the kindred (except one case which could not be completely inferred from her children). The two haplotypes are identical for Mfd188 and 42D6. Accordingly, data from Kindred 1901 would also place the BRCA1 locus distal to Mfd15.

Kindred 2035 is similar to K1901 in disease phenotype. The median age of diagnosis for the eight cases of breast cancer in this kindred is 37. One case also had ovarian cancer at age 60. The breast cancer cases in this family descend from two sisters who were both unaffected with breast cancer until their death in the eighth decade. Each branch contains four cases of breast cancer with at least one case in each branch having markedly early onset. This kindred has a LOD score of 2.34 with Mfd15. The haplotypes segregating with breast cancer in the two branches share an identical

allele at Mfd15 but differ for the distal loci Mfd188 and nm23 (a marker typed as part of the consortium which is located just distal to 42D6, (Hall *et al.*, 1992). Although the two haplotypes are concordant for marker 42D6, it is likely that the alleles are shared identical by state (the same allele but derived from different ancestors), rather than identical by descent (derived from a common ancestor) since the shared allele is the second most common allele observed at this locus. By contrast the linked allele shared at Mfd15 has a frequency of 0.04. This is a key recombinant in our dataset as it is the sole recombinant in which BRCA1 segregated with the proximal portion of the haplotype, thus setting the distal boundary to the BRCA1 region. For this event not to be a key recombinant requires that a second mutant BRCA1 gene be present in a spouse marrying into the kindred who also shares the rare Mfd15 allele segregating with breast cancer in both branches of the kindred. This event has a probability of less than one in a thousand. The evidence from this kindred therefore placed the BRCA1 locus proximal to Mfd188.

EXAMPLE 3

Creation of a Fine Structure Map and Refinement of the BRCA1 Region to Mfd191-Mfd188 using Additional STR Polymorphisms

In order to improve the characterization of our recombinants and define closer flanking markers, a dense map of this relatively small region on chromosome 17q was required. The chromosome 17 workshop has produced a consensus map of this region (Figure 1) based on a combination of genetic and physical mapping studies (Fain, 1992). This map contains both highly polymorphic STR polymorphisms, and a number of nonpolymorphic expressed genes. Because this map did not give details on the evidence for this order nor give any measure of local support for inversions in the order of adjacent loci, we viewed it as a rough guide for obtaining resources to be used for the development of new markers and construction of our own detailed genetic and physical map of a small region containing BRCA1. Our approach was to analyze existing STR markers provided by other investigators and any newly developed markers from our laboratory with respect to both a panel of meiotic (genetic) breakpoints identified using DNA from the CEPH reference families and a panel of somatic cell hybrids (physical breakpoints) constructed for this region. These markers included 26C2 developed in our laboratory which maps proximal to Mfd15, Mfd191 (provided by James Weber), THRA1 (Futreal *et al.* 1992), and three polymorphisms kindly provided to us by Dr. Donald Black, NM23 (Hall *et al.* 1992), SCG40 (D17S181), and 6C1 (D17S293).

Genetic localization of markers. In order to localize new markers genetically within the region of interest, we have identified a number of key meiotic breakpoints within the region, both in the CEPH reference panel and in our large breast cancer kindred (K2082). Given the small genetic distance in this region, they are likely to be only a relatively small set of recombinants which can be used for this purpose, and they are likely to group markers into sets. The orders of the markers within each set can only be determined by physical mapping. However the number of genotypings necessary to position a new marker is minimized. These breakpoints are illustrated in Tables 3 and 4. Using this approach we were able to genetically order the markers THRA1, 6C1, SCG40, and Mfd191. As can be seen from Tables 3 and 4, THRA1 and MFD191 both map inside the Mfd15--Mfd188 region we had previously identified as containing the BRCA1 locus. In Tables 3 and 4, M/P indicates a maternal or paternal recombinant. A "1" indicates inherited allele is of grandpaternal origin, while a "0" indicates grandmaternal origin, and "-" indicates that the locus was untyped or uninformative.

TABLE 3

TABLE 3
CEPH Recombinants

<u>Family</u>	<u>ID</u>	<u>M/P</u>	<u>Mfd15</u>	<u>THRA1</u>	<u>Mfd191</u>	<u>Mfd188</u>	<u>SCG40</u>	<u>6C1</u>	<u>42D6</u>
13292	4	M	1	1	1	0	0	0	0
13294	4	M	1	1	1	0	0	0	0
13294	6	M	0	0	1	1	-	-	-
1334	3	M	1	1	1	1	1	0	0
1333	4	M	1	1	1	0	-	-	0
1333	6	M	0	0	1	1	-	-	1
1333	8	P	1	0	0	0	-	-	0
1377	8	M	0	-	0	0	0	0	1

-52-

TABLE 4
Kindred 2082 Recombinants

<u>Family</u>	<u>ID</u>	<u>M/P</u>	<u>Mfd15</u>	<u>Mfd191</u>	<u>Mfd188</u>	<u>SCG40</u>	<u>6C1</u>	<u>42D6</u>
75		M	0	1	1	1	-	-
63		M	0	0	1	1	-	1
125		M	1	1	1	0	-	0
40		M	1	1	0	0	-	0

-53-

Analysis of markers Mfd15, Mfd188, Mfd191, and THRA1 in our recombinant families.

Mfd15, Mfd188, Mfd191 and THRA1 were typed in our recombinant families and examined for additional information to localize the BRCA1 locus. In kindred 1901, the Mfd15 recombinant was recombinant for THRA1 but uninformative for Mfd191, thus placing BRCA1 distal to THRA1. In K2082, the recombinant with Mfd15 also was recombinant with Mfd191, thus placing the BRCA1 locus distal to Mfd191 (Goldgar *et al.*, 1994). Examination of THRA1 and Mfd191 in kindred K2035 yielded no further localization information as the two branches were concordant for both markers. However, SCG40 and 6C1 both displayed the same pattern as Mfd188, thus increasing our confidence in the localization information provided by the Mfd188 recombinant in this family. The BRCA1 locus, or at least a portion of it, therefore lies within an interval bounded by Mfd191 on the proximal side and Mfd188 on the distal side.

EXAMPLE 4

Development of Genetic and Physical
Resources in the Region of Interest

To increase the number of highly polymorphic loci in the Mfd191-Mfd188 region, we developed a number of STR markers in our laboratory from cosmids and YACs which physically map to the region. These markers allowed us to further refine the region.

STSs were identified from genes known to be in the desired region to identify YACs which contained these loci, which were then used to identify subclones in cosmids P1s or BAC. These subclones were then screened for the presence of a CA tandem repeat using a (CA)_n oligonucleotide (Pharmacia). Clones with a strong signal were selected preferentially, since they were more likely to represent CA-repeats which have a large number of repeats and/or are of near-perfect fidelity to the (CA)_n pattern. Both of these characteristics are known to increase the probability of polymorphism (Weber, 1990). These clones were sequenced directly from the vector to locate the repeat. We obtained a unique sequence on one side of the CA-repeat by using one of a set of possible primers complementary to the end of a CA-repeat, such as (GT)₁₀T. Based on this unique sequence, a primer was made to sequence back across the repeat in the other direction, yielding a unique sequence for design of a second primer flanking the CA-repeat. STRs were then screened for polymorphism on a small group of unrelated individuals and tested against the hybrid panel to confirm their physical localization. New markers which satisfied these criteria were then typed in a

set of 40 unrelated individuals from the Utah and CEPH families to obtain allele frequencies appropriate for the study population. Many of the other markers reported in this study were tested in a smaller group of CEPH unrelated individuals to obtain similarly appropriate allele frequencies.

Using the procedure described above, a total of eight polymorphic STRs were found from these YACS. Of the loci identified in this manner, four were both polymorphic and localized to the BRCA1 region. Four markers did not localize to chromosome 17, reflecting the chimeric nature of the YACs used. The four markers which were in the region were denoted AA1, ED2, 4-7, and YM29. AA1 and ED2 were developed from YACs positive for the RNU2 gene, 4-7 from an EPB3 YAC and YM29 from a cosmid which localized to the region by the hybrid panel. A description of the number of alleles, heterozygosity and source of these four and all other STR polymorphisms analyzed in the breast cancer kindreds is given below in Table 5.

TABLE 5

Polymorphic Short Tandem Repeat Markers Used
for Fine Structure Mapping of the BRCA1 Locus

Clone	Gene	Na**	Hetero- zygosity	Allele* Frequency (%)					
				1	2	3	4	5	6
Mfd15	D17S250	10	0.82	26	22	15	7	7	23
THRA1	THRA1	5							
Mfd191	D17S776	7	0.55	48	20	11	7	7	7
ED2	D17S1327	12	0.55	62	9	8	5	5	11
AA1	D17S1326	7	0.83	28	28	25	8	6	5
CA375	D17S184	10	0.75	26	15	11	9	9	20
4-7	D17S1183	9	0.50	63	15	8	6	4	4
YM29	--	9	0.62	42	24	12	7	7	8
Mfd188	D17S579	12	0.92	33	18	8	8	8	25
SCG40	D17S181	14	0.90	20	18	18	10	8	35
42D6	D17S588	11	0.86	21	17	11	10	9	32
6C1	D17S293	7	0.75	30	30	11	11	9	9
Z109	D17S750	9	0.70	33	27	7	7	7	19
tdj1475	D17S1321	13	0.84	21	16	11	11	8	33
CF4	D17S1320	6	0.63	50	27	9	7	4	3
tdj1239	D17S1328	10	0.80	86	10	9	7	4	14
U5	D17S1325	13	0.83	19	16	12	10	9	34

* Allele codes 1-5 are listed in decreasing frequency; allele numbers do not correspond to fragment sizes. Allele 6 frequency is the joint frequency of all other alleles for each locus.

** Number of alleles seen in the genetically independent DNA samples used for calculating allele frequencies.

5

SECRET

TABLE 6

Key Recombinants Used for Genetic
Ordering of New STR Loci Developed in
Our Laboratory Within the BRCA1 Region of 17q

CEPH Family	ID	M/P	Mfd15	THRA1	Mfd191	ED2	AA1	Z109	4-7	YM29	Mfd188	SCG40	42D6
13292	4	M	1	1	1	1	1	0	0	0	0	0	0
13294	4	M	1	0	0	-	0	-	-	-	0	-	-
13294	6	M	0	0	1	-	1	-	-	-	1	-	-
1333	4	M	1	1	1	-	0	-	-	0	0	-	0
1333	6	M	0	0	1	-	1	-	-	1	1	-	1
1333	3	M	0	0	1	-	-	-	1	1	1	-	1

TABLE 7

TABLE 7

Kindred 2082 Recombinants

<u>ID</u>	<u>M/P</u>	<u>Mfd15</u>	<u>Mfd191</u>	<u>ED2</u>	<u>AA1</u>	<u>4-7</u>	<u>YM29</u>	<u>Mfd188</u>	<u>SCG40</u>	<u>42D6</u>
63	M	0	0	1	-	1	1	1	1	1
125	M	1	1	1	-	1	1	1	0	0
40	M	1	1	0	-	0	-	0	0	0
22	P	0	0	1	1	1	1	1	1	1

-59-

From CEPH 1333-04, we see that AA1 and YM29 must lie distal to Mfd191. From 13292, it can be inferred that both AA1 and ED2 are proximal to 4-7, YM29, and Mfd188. The recombinants found in K2082 provide some additional ordering information. Three independent observations (individual numbers 22, 40, & 63) place AA1, ED2, 4-7, and YM29, and Mfd188 distal to Mfd191, while ID 125 places 4-7, YM29, and Mfd188 proximal to SCG40. No genetic information on the relative ordering within the two clusters of markers AA1/ED2 and 4-7/YM29/Mfd188 was obtained from the genetic recombinant analysis. Although ordering loci with respect to hybrids which are known to contain "holes" in which small pieces of interstitial human DNA may be missing is problematic, the hybrid patterns indicate that 4-7 lies above both YM29 and Mfd188.

EXAMPLE 5

Genetic Analyses of Breast Cancer
Kindreds with Markers AA1, 4-7, ED2, and YM29

In addition to the three kindreds containing key recombinants which have been discussed previously, kindred K2039 was shown through analysis of the newly developed STR markers to be linked to the region and to contain a useful recombinant.

Table 8 defines the haplotypes (shown in coded form) of the kindreds in terms of specific marker alleles at each locus and their respective frequencies. In Table 8, alleles are listed in descending order of frequency; frequencies of alleles 1-5 for each locus are given in Table 5. Haplotypes coded H are BRCA1 associated haplotypes, P designates a partial H haplotype, and an R indicates an observable recombinant haplotype. As evident in Table 8, not all kindreds were typed for all markers; moreover, not all individuals within a kindred were typed for an identical set of markers, especially in K2082. With one exception, only haplotypes inherited from affected or at-risk kindred members are shown; haplotypes from spouses marrying into the kindred are not described. Thus in a given sibship, the appearance of haplotypes X and Y indicates that both haplotypes from the affected/at-risk individual were seen and neither was a breast cancer associated haplotype.

TABLE 8

Breast Cancer Linked Haplotypes Found in the Three Kindreds

Kin.	HAP	Mfd 15	THRA1	Mfd 191	tdj 1475	ED2	AA1	Z109	CA375	4-7	YM29	Mfd 188	SCG40	6C1	42D6
1901	H1	1	5	5	3	1	4	NI	NI	1	1	3	NI	NI	1
	R2	9	2	5	6	1	4	NI	NI	1	1	3	NI	NI	1
2082	H1	3	NI	4	6	6	1	NI	NI	2	1	4	2	NI	1
	P1	3	NI	4	NI	NI	NI	NI	NI	NI	NI	4	2	NI	1
	P2	3	NI	NI	NI	NI	NI	NI	NI	NI	NI	4	NI	NI	NI
	R1	6	NI	1	5	6	1	NI	NI	2	1	4	2	NI	1
	R2	6	NI	4	6	6	1	NI	NI	2	1	4	2	NI	1
	R3	3	NI	4	NI	6	1	NI	NI	2	1	4	1	NI	7
	R4	7	NI	1	NI	1	5	NI	NI	4	6	1	2	NI	1
	R5	3	NI	4	NI	NI	NI	NI	NI	NI	2	1	NI	NI	NI
	R6	3	NI	4	3	1	2	NI	NI	1	2	2	6	NI	6
R7	3	NI	4	3	7	1	NI	NI	1	1	3	7	NI	4	
2035	H1	8	2	1	NI	5	1	1	4	3	1	6	8	2	4
	H2	8	2	1	NI	5	1	1	2	1	1	2	3	1	4
	R2	8	2	1	NI	5	1	1	2	1	1	2	3	6	1

In kindred K1901, the new markers showed no observable recombination with breast cancer susceptibility, indicating that the recombination event in this kindred most likely took place between THRA1 and ED2. Thus, no new BRCA1 localization information was obtained based upon studying the four new markers in this kindred. In kindred 2082 the key recombinant individual has inherited the linked alleles for ED2, 4-7, AA1, and YM29, and was recombinant for tdj1474 indicating that the recombination event occurred in this individual between tdj1474 and ED2/AA1.

There are three haplotypes of interest in kindred K2035, H1, H2, and R2 shown in Table 8. H1 is present in the four cases and one obligate male carrier descendant from individual 17 while H2 is present or inferred in two cases and two obligate male carriers in descendants of Individual 10. R2 is identical to H2 for loci between and including Mfd15 and SCG40, but has recombined between SCG40 and 42D6. Since we have established that BRCA1 is proximal to 42D6, this H2/R2 difference adds no further localization information. H1 and R2 share an identical allele at Mfd15, THRA1, AA1, and ED2 but differ for loci presumed distal to ED2, i.e., 4-7, Mfd188, SCG40, and 6C1. Although the two haplotypes are concordant for the 5th allele for marker YM29, a marker which maps physically between 4-7 and Mfd188, it is likely that the alleles are shared identical by state rather than identical by descent since this allele is the most common allele at this locus with a frequency estimated in CEPH parents of 0.42. By contrast, the linked alleles shared at the Mfd15 and ED2 loci have frequencies of 0.04 and 0.09, respectively. They also share more common alleles at Mfd191 (frequency = 0.52), THRA1, and AA1 (frequency = 0.28). This is the key recombinant in the set as it is the sole recombinant in which breast cancer segregated with the proximal portion of the haplotype, thus setting the distal boundary. The evidence from this kindred therefore places the BRCA1 locus proximal to 4-7.

The recombination event in kindred 2082 which places BRCA1 distal to tdj1474 is the only one of the four events described which can be directly inferred; that is, the affected mother's genotype can be inferred from her spouse and offspring, and the recombinant haplotype can be seen in her affected daughter. In this family the odds in favor of affected individuals carrying BRCA1 susceptibility alleles are extremely high; the only possible interpretations of the data are that BRCA1 is distal to Mfd191 or alternatively that the purported recombinant is a sporadic case of ovarian cancer at age 44. Rather than a directly observable or inferred recombinant, interpretation of kindred 2035 depends on the observation of distinct 17q-haplotypes segregating in different and sometimes distantly related branches of the kindred. The observation that portions of these haplotypes have alleles in common for some markers while they differ at other markers places the BRCA1 locus in

the shared region. The confidence in this placement depends on several factors: the relationship between the individuals carrying the respective haplotypes, the frequency of the shared allele, the certainty with which the haplotypes can be shown to segregate with the BRCA1 locus, and the density of the markers in the region which define the haplotype. In the case of kindred 2035, the two branches are closely related, and each branch has a number of early onset cases which carry the respective haplotype. While two of the shared alleles are common, (Mfd191, THRA1), the estimated frequencies of the shared alleles at Mfd15, AA1, and ED2 are 0.04, 0.28, and 0.09, respectively. It is therefore highly likely that these alleles are identical by descent (derived from a common ancestor) rather than identical by state (the same allele but derived from the general population).

EXAMPLE 6

Refined Physical Mapping Studies Place the BRCA1 Gene in a Region Flanked by tdj11474 and U5R

Since its initial localization to chromosome 17q in 1990 (Hall *et al.* 1990) a great deal of effort has gone into localizing the BRCA1 gene to a region small enough to allow implementation of effective positional cloning strategies to isolate the gene. The BRCA1 locus was first localized to the interval Mfd15 (D17S250) - 42D6 (D17S588) by multipoint linkage analysis (Easton *et al.* 1993) in the collaborative Breast Cancer Linkage Consortium dataset consisting of 214 families collected worldwide. Subsequent refinements of the localization have been based upon individual recombinant events in specific families. The region THRA1 - D17S183 was defined by Bowcock *et al.*, 1993; and the region THRA1 - D17S78 was defined by Simard *et al.*, 1993.

We further showed that the BRCA1 locus must lie distal to the marker Mfd191 (D17S776) (Goldgar *et al.* 1994). This marker is known to lie distal to THRA1 and RARA. The smallest published region for the BRCA1 locus is thus between D17S776 and D17S78. This region still contains approximately 1.5 million bases of DNA, making the isolation and testing of all genes in the region a very difficult task. We have therefore undertaken the tasks of constructing a physical map of the region, isolating a set of polymorphic STR markers located in the region, and analyzing these new markers in a set of informative families to refine the location of the BRCA1 gene to a manageable interval.

Four families provide important genetic evidence for localization of BRCA1 to a sufficiently small region for the application of positional cloning strategies. Two families (K2082, K1901)

provide data relating to the proximal boundary for BRCA1 and the other two (K2035, K1813) fix the distal boundary. These families are discussed in detail below. A total of 15 Short Tandem Repeat markers assayable by PCR were used to refine this localization in the families studied. These markers include DS17S7654, DS17S975, tdj1474, and tdj1239. Primer sequences for these markers are provided in SEQ ID NO:3 and SEQ ID NO:4 for DS17S754; in SEQ ID NO:5 and SEQ ID NO:6 for DS17S975; in SEQ ID NO:7 and SEQ ID NO:8 for tdj1474; and, in SEQ ID NO:9 and SEQ ID NO:10 for tdj1239.

Kindred 2082

Kindred 2082 is the largest BRCA1-linked breast/ovarian cancer family studied to date. It has a Lod score of 8.6, providing unequivocal evidence for 17q linkage. This family has been previously described and shown to contain a critical recombinant placing BRCA1 distal to MFD191 (D17S776). This recombinant occurred in a woman diagnosed with ovarian cancer at age 45 whose mother had ovarian cancer at age 63. The affected mother was deceased; however, from her children, she could be inferred to have the linked haplotype present in the 30 other linked cases in the family in the region between Mfd15 and Mfd188. Her affected daughter received the linked allele at the loci ED2, 4-7, and Mfd188, but received the allele on the non-BRCA1 chromosome at Mfd15 and Mfd191. In order to further localize this recombination breakpoint, we tested the key members of this family for the following markers derived from physical mapping resources: tdj1474, tdj1239, CF4, D17S855. For the markers tdj1474 and CF4, the affected daughter did not receive the linked allele. For the STR locus tdj1239, however, the mother could be inferred to be informative and her daughter did receive the BRCA1-associated allele. D17S855 was not informative in this family. Based on this analysis, the order is 17q centomere - Mfd191 - 17HSD - CF4 - tdj1474 - tdj1239 - D17S855 - ED2 - 4-7 - Mfd188 - 17q telomere. The recombinant described above therefore places BRCA1 distal to tdj1474, and the breakpoint is localized to the interval between tdj1474 and tdj1239. The only alternative explanation for the data in this family other than that of BRCA1 being located distal to tdj1474, is that the ovarian cancer present in the recombinant individual is caused by reasons independent of the BRCA1 gene. Given that ovarian cancer diagnosed before age 50 is rare, this alternate explanation is exceedingly unlikely.

Kindred 1901

Kindred 1901 is an early-onset breast cancer family with 7 cases of breast cancer diagnosed before 50, 4 of which were diagnosed before age 40. In addition, there were three cases of breast cancer diagnosed between the ages of 50 and 70. One case of breast cancer also had ovarian cancer at age 61. This family currently has a Lod score of 1.5 with D17S855. Given this linkage evidence and the presence of at least one ovarian cancer case, this family has a posterior probability of being due to BRCA1 of over 0.99. In this family, the recombination comes from the fact that an individual who is the brother of the ovarian cancer case from which the majority of the other cases descend, only shares a portion of the haplotype which is cosegregating with the other cases in the family. However, he passed this partial haplotype to his daughter who developed breast cancer at age 44. If this case is due to the BRCA1 gene, then only the part of the haplotype shared between this brother and his sister can contain the BRCA1 gene. The difficulty in interpretation of this kind of information is that while one can be sure of the markers which are not shared and therefore recombinant, markers which are concordant can either be shared because they are non-recombinant, or because their parent was homozygous. Without the parental genotypic data it is impossible to discriminate between these alternatives. Inspection of the haplotype in K1901, shows that he does not share the linked allele at Mfd15 (D17S250), THRA1, CF4 (D17S1320), and tdj1474 (17DS1321). He does share the linked allele at Mfd191 (D17S776), ED2 (D17S1327), tdj1239 (D17S1328), and Mfd188 (D17S579). Although the allele shared at Mfd191 is relatively rare (0.07), we would presume that the parent was homozygous since they are recombinant with markers located nearby on either side, and a double recombination event in this region would be extremely unlikely. Thus the evidence in this family would also place the BRCA1 locus distal to tdj1474. However, the lower limit of this breakpoint is impossible to determine without parental genotype information. It is intriguing that the key recombinant breakpoint in this family confirms the result in Kindred 2082. As before, the localization information in this family is only meaningful if the breast cancer was due to the BRCA1 gene. However, her relatively early age at diagnosis (44) makes this seem very likely since the risk of breast cancer before age 45 in the general population is low (approximately 1%).

Kindred 2035

This family is similar to K1901 in that the information on the critical recombinant events is not directly observed but is inferred from the observation that the two haplotypes which are cosegregating with the early onset breast cancer in the two branches of the family appear identical

for markers located in the proximal portion of the 17q BRCA1 region but differ at more distal loci. Each of these two haplotypes occurs in at least four cases of early-onset or bilateral breast cancer. The overall LOD score with ED2 in this family is 2.2, and considering that there is a case of ovarian cancer in the family (indicating a prior probability of BRCA1 linkage of 80%), the resulting posterior probability that this family is linked to BRCA1 is 0.998. The haplotypes are identical for the markers Mfd15, THRA1, Mfd191, ED2, AA1, D17S858 and D17S902. The common allele at Mfd15 and ED2 are both quite rare, indicating that this haplotype is shared identical by descent. The haplotypes are discordant, however, for CA375, 4-7, and Mfd188, and several more distal markers. This indicates that the BRCA1 locus must lie above the marker CA-375. This marker is located approximately 50 kb below D17S78, so it serves primarily as additional confirmation of this previous lower boundary as reported in Narod *et al.* (1993).

Kindred 1813

Kindred 1813 is a small family with four cases of breast cancer diagnosed under the age of 40 whose mother had breast cancer diagnosed at age 45 and ovarian cancer at age 61. This situation is somewhat complicated by the fact the four cases appear to have three different fathers, only one of whom has been genotyped. However, by typing a number of different markers in the BRCA1 region as well as highly polymorphic markers elsewhere in the genome, the paternity of all children in the family has been determined with a high degree of certainty. This family yields a maximum multipoint lod score of 0.60 with 17q markers and, given that there is at least one case of ovarian cancer, results in a posterior probability of being a BRCA1 linked family of 0.93. This family contains a directly observable recombination event in individual 18 (see Figure 5 in Smard *et al.*, *Human Mol. Genet.* 2:1193-1199 (1993)), who developed breast cancer at age 34. The genotype of her affected mother at the relevant 17q loci can be inferred from her genotypes, her affected sister's genotypes, and the genotypes of three other unaffected siblings. Individual 18 inherits the BRCA1-linked alleles for the following loci: Mfd15, THRA1, D17S800, D17S855, AA1, and D17S931. However, for markers below D17S931, i.e., U5R, vrs31, D17S858, and D17S579, she has inherited the alleles located on the non-disease bearing chromosome. The evidence from this family therefore would place the BRCA1 locus proximal to the marker U5R. Because of her early age at diagnosis (34) it is extremely unlikely that the recombinant individual's cancer is not due to the gene responsible for the other cases of breast/ovarian cancer in this family; the uncertainty in this family

comes from our somewhat smaller amount of evidence that breast cancer in this family is due to BRCA1 rather than a second, as yet unmapped, breast cancer susceptibility locus.

Size of the region containing BRCA1

Based on the genetic data described in detail above, the BRCA1 locus must lie in the interval between the markers tdj1474 and U5R, both of which were isolated in our laboratory. Based upon the physical maps shown in Figure 2 and 3, we can try to estimate the physical distance between these two loci. It takes approximately 14 P1 clones with an average insert size of approximately 80 kb to span the region. However, because all of these P1s overlap to some unknown degree, the physical region is most likely much smaller than 14 times 80 kb.. Based on restriction maps of the clones covering the region, we estimate the size of the region containing BRCA1 to be approximately 650 kb.

EXAMPLE 7

Identification of Candidate cDNA Clones for the BRCA1 Locus by Genomic Analysis of the Contig Region

Complete screen of the plausible region. The first method to identify candidate cDNAs, although labor intensive, used known techniques. The method comprised the screening of cosmid and P1 and BAC clones in the contig to identify putative coding sequences. The clones containing putative coding sequences were then used as probes on filters of cDNA libraries to identify candidate cDNA clones for future analysis. The clones were screened for putative coding sequences by either of two methods.

Zoo blots. The first method for identifying putative coding sequences was by screening the cosmid and P1 clones for sequences conserved through evolution across several species. This technique is referred to as "zoo blot analysis" and is described by Monaco, 1986. Specifically, DNA from cow, chicken, pig, mouse and rat were digested with the restriction enzymes EcoRI and HindIII (8 µg of DNA per enzyme). The digested DNAs were separated overnight on an 0.7% gel at 20 volts for 16 hours (14 cm gel), and the DNA transferred to Nylon membranes using standard Southern blot techniques. For example, the zoo blot filter was treated at 65°C in 0.1 x SSC, 0.5% SDS, and 0.2M Tris, pH 8.0, for 30 minutes and then blocked overnight at 42°C in 5x SSC, 10%

PEG 8000, 20 mM NaPO₄ pH 6.8, 100 µg/ml Salmon Sperm DNA, 1x Denhardt's, 50% formamide, 0.1% SDS, and 2 µg/ml cot-1 DNA.

The cosmid and Pl clones to be analyzed were digested with a restriction enzyme to release the human DNA from the vector DNA. The DNA was separated on a 14 cm, 0.5% agarose gel run overnight at 20 volts for 16 hours. The human DNA bands were cut out of the gel and electroeluted from the gel wedge at 100 volts for at least two hours in 0.5x Tris Acetate buffer (Maniatis *et al.*, 1989). The eluted Not I digested DNA (~15 kb to 25 kb) was then digested with EcoRI restriction enzyme to give smaller fragments (~0.5 kb to 5.0 kb) which melt apart more easily for the next step of labeling the DNA with radionucleotides. The DNA fragments were labeled by means of the hexamer random prime labeling method (Boehringer-Mannheim, Cat. #1004760). The labeled DNA was spermine precipitated (add 100 µl TE, 5 µl 0.1 M spermine, and 5 µl of 10 mg/ml salmon sperm DNA) to remove unincorporated radionucleotides. The labeled DNA was then resuspended in 100 µl TE, 0.5 M NaCl at 65°C for 5 minutes and then blocked with Human Cot-1 DNA for 2-4 hrs. as per the manufacturer's instructions (Gibco/BRL, Cat. #5279SA). The Cot-1 blocked probe was incubated on the zoo blot filters in the blocking solution overnight at 42°C. The filters were washed for 30 minutes at room temperature in 2 x SSC, 0.1% SDS, and then in the same buffer for 30 minutes at 55°C. The filters were then exposed 1 to 3 days at -70°C to Kodak XAR-5 film with an intensifying screen. Thus, the zoo blots were hybridized with either the pool of Eco-R1 fragments from the insert, or each of the fragments individually. .

HTF island analysis. The second method for identifying cosmids to use as probes on the cDNA libraries was HTF island analysis. Since the pulsed-field map can reveal HTF islands, cosmids that map to these HTF island regions were analyzed with priority. HTF islands are segments of DNA which contain a very high frequency of unmethylated CpG dinucleotides (Tonolio *et al.*, 1990) and are revealed by the clustering of restriction sites of enzymes whose recognition sequences include CpG dinucleotides. Enzymes known to be useful in HTF-island analysis are AscI, NotI, BssHII, EagI, SacII, NaeI, NarI, SmaI, and MluI (Anand, 1992). A pulsed-field map was created using the enzymes NotI, NruI, EagI, SacII, and SalI, and two HTF islands were found. These islands are located in the distal end of the region, one being distal to the GP2B locus, and the other being proximal to the same locus, both outside the BRCA1 region. The cosmids derived from the YACs that cover these two locations were analyzed to identify those that contain these restriction sites, and thus the HTF islands.

cDNA screening. Those clones that contain HTF islands or show hybridization to other species DNA besides human are likely to contain coding sequences. The human DNA from these clones was isolated as whole insert or as EcoRI fragments and labeled as described above. The labeled DNA was used to screen filters of various cDNA libraries under the same conditions as the zoo blots except that the cDNA filters undergo a more stringent wash of 0.1 x SSC, 0.1% SDS at 65°C for 30 minutes twice.

Most of the cDNA libraries used to date in our studies (libraries from normal breast tissue, breast tissue from a woman in her eighth month of pregnancy and a breast malignancy) were prepared at Clontech, Inc. The cDNA library generated from breast tissue of an 8 month pregnant woman is available from Clontech (Cat. #HL1037a) in the Lambda gt-10 vector, and is grown in C600Hfl bacterial host cells. Normal breast tissue and malignant breast tissue samples were isolated from a 37 year old Caucasian female and one-gram of each tissue was sent to Clontech for mRNA processing and cDNA library construction. The later two libraries were generated using both random and oligo-dT priming, with size selection of the final products which were then cloned into the Lambda Zap II vector, and grown in XL1-blue strain of bacteria as described by the manufacturer. Additional tissue-specific cDNA libraries include human fetal rain (Stratagene, Cat. 936206), human testing (Clontech Cat. HL3024), human thymus (Clontech Cat. HL1127n), human brain (Clontech Cat. HL11810), human placenta (Clontech Cat 1075b), and human skeletal muscle (Clontech Cat. HL1124b).

The cDNA libraries were plated with their host cells on NZCYM plates, and filter lifts are made in duplicate from each plate as per Maniatis *et al.*(1989). Insert (human) DNA from the candidate genomic clones was purified and radioactively labeled to high specific activity. The radioactive DNA was then hybridized to the cDNA filters to identify those cDNAs which correspond to genes located within the candidate cosmid clone. cDNAs identified by this method were picked, replated, and screened again with the labeled clone insert or its derived EcoRI fragment DNA to verify their positive status. Clones that were positive after this second round of screening were then grown up and their DNA purified for Southern blot analysis and sequencing. Clones were either purified as plasmid through *in vivo* excision of the plasmid from the Lambda vector as described in the protocols from the manufacturers, or isolated from the Lambda vector as a restriction fragment and subcloned into plasmid vector.

The Southern blot analysis was performed in duplicate, one using the original genomic insert DNA as a probe to verify that cDNA insert contains hybridizing sequences. The second blot was

hybridized with cDNA insert DNA from the largest cDNA clone to identify which clones represent the same gene. All cDNAs which hybridize with the genomic clone and are unique were sequenced and the DNA analyzed to determine if the sequences represent known or unique genes. All cDNA clones which appear to be unique were further analyzed as candidate BRCA1 loci. Specifically, the clones are hybridized to Northern blots to look for breast specific expression and differential expression in normal versus breast tumor RNAs. They are also analyzed by PCR on clones in the BRCA1 region to verify their location. To map the extent of the locus, full length cDNAs are isolated and their sequences used as PCR probes on the YACs and the clones surrounding and including the original identifying clones. Intron-exon boundaries are then further defined through sequence analysis.

We have screened the normal breast, 8 month pregnant breast and fetal brain cDNA libraries with zoo blot-positive Eco RI fragments from cosmid BAC and P1 clones in the region. Potential BRCA1 cDNA clones were identified among the three libraries. Clones were picked, replated, and screened again with the original probe to verify that they were positive.

Analysis of hybrid-selected cDNA. cDNA fragments obtained from direct selection were checked by Southern blot hybridization against the probe DNA to verify that they originated from the contig. Those that passed this test were sequenced in their entirety. The set of DNA sequences obtained in this way were then checked against each other to find independent clones that overlapped. For example, the clones 694-65, 1240-1 and 1240-33 were obtained independently and subsequently shown to derive from the same contiguous cDNA sequence which has been named EST:489:1.

Analysis of candidate clones. One or more of the candidate genes generated from above were sequenced and the information used for identification and classification of each expressed gene. The DNA sequences were compared to known genes by nucleotide sequence comparisons and by translation in all frames followed by a comparison with known amino acid sequences. This was accomplished using Genetic Data Environment (GDE) version 2.2 software and the Basic Local Alignment Search Tool (Blast) series of client/server software packages (e.g., BLASTN 1.3.13MP), for sequence comparison against both local and remote sequence databases (e.g., GenBank), running on Sun SPARC workstations. Sequences reconstructed from collections of cDNA clones identified with the cosmids and PIs have been generated. All candidate genes that represented new sequences were analyzed further to test their candidacy for the putative BRCA1 locus.

Mutation screening. To screen for mutations in the affected pedigrees, two different approaches were followed. First, genomic DNA isolated from family members known to carry the susceptibility

allele of BRCA1 was used as a template for amplification of candidate gene sequences by PCR. If the PCR primers flank or overlap an intron/exon boundary, the amplified fragment will be larger than predicted from the cDNA sequence or will not be present in the amplified mixture. By a combination of such amplification experiments and sequencing of P1, BAC or cosmid clones using the set of designed primers it is possible to establish the intron/exon structure and ultimately obtain the DNA sequences of genomic DNA from the pedigrees.

A second approach that is much more rapid if the intron/exon structure of the candidate gene is complex involves sequencing fragments amplified from pedigree lymphocyte cDNA. cDNA synthesized from lymphocyte mRNA extracted from pedigree blood was used as a substrate for PCR amplification using the set of designed primers. If the candidate gene is expressed to a significant extent in lymphocytes, such experiments usually produce amplified fragments that can be sequenced directly without knowledge of intron/exon junctions.

The products of such sequencing reactions were analyzed by gel electrophoresis to determine positions in the sequence that contain either mutations such as deletions or insertions, or base pair substitutions that cause amino acid changes or other detrimental effects.

Any sequence within the BRCA1 region that is expressed in breast is considered to be a candidate gene for BRCA1. Compelling evidence that a given candidate gene corresponds to BRCA1 comes from a demonstration that pedigree families contain defective alleles of the candidate.

EXAMPLE 8

Identification of BRCA1

Identification of BRCA1. Using several strategies, a detailed map of transcripts was developed for the 600 kb region of 17q21 between D17S1321 and D17S1324. Candidate expressed sequences were defined as DNA sequences obtained from: 1) direct screening of breast, fetal brain, or lymphocyte cDNA libraries, 2) hybrid selection of breast, lymphocyte or ovary cDNAs, or 3) random sequencing of genomic DNA and prediction of coding exons by XPOUND (Thomas and Skolnick, 1994). These expressed sequences in many cases were assembled into contigs composed of several independently identified sequences. Candidate genes may comprise more than one of these candidate expressed sequences. Sixty-five candidate expressed sequences within this region were identified by hybrid selection, by direct screening of cDNA libraries, and by random sequencing of P1 subclones. Expressed sequences were characterized by transcript size, DNA sequence, database comparison,

expression pattern, genomic structure, and, most importantly, DNA sequence analysis in individuals from kindreds segregating 17q-linked breast and ovarian cancer susceptibility.

Three independent contigs of expressed sequence, 1141:1 (649 bp), 694:5 (213 bp) and 754:2 (1079 bp) were isolated and eventually shown to represent portions of BRCA1. When ESTs for these contigs were used as hybridization probes for northern analysis, a single transcript of approximately 7.8 kb was observed in normal breast mRNA, suggesting that they encode different portions of a single gene. Screens of breast, fetal brain, thymus, testes, lymphocyte and placental cDNA libraries and PCR experiments with breast mRNA linked the 1141:1, 694:5 and 754:2 contigs. 5' RACE experiments with thymus, testes, and breast mRNA extended the contig to the putative 5' end, yielding a composite full length sequence. PCR and direct sequencing of P1s and BACs in the region were used to identify the location of introns and allowed the determination of splice donor and acceptor sites. These three expressed sequences were merged into a single transcription unit that proved in the final analysis to be BRCA1. This transcription unit is located adjacent to D17S855 in the center of the 600 kb region (Fig. 4).

Combination of sequences obtained from cDNA clones, hybrid selection sequences, and amplified PCR products allowed construction of a composite full length BRCA1 cDNA (SEQ ID NO:1). The sequence of the BRCA1 gene has also been deposited with GenBank and assigned accession number U-14680. This deposited sequence is incorporated herein by reference. The cDNA clone extending farthest in the 3' direction contains a poly(A) tract preceded by a polyadenylation signal. Conceptual translation of the cDNA revealed a single long open reading frame of 208 kilodaltons (amino acid sequence: SEQ ID NO:2) with a potential initiation codon flanked by sequences resembling the Kozak consensus sequence (Kozak, 1987). Smith-Waterman (1981) and BLAST (Altschul *et al.*, 1990) searches identified a sequence near the amino terminus with considerable homology to zinc-finger domains (Fig. 5). This sequence contains cystine and histidine residues present in the consensus C3HC4 zinc-finger motif and shares multiple other residues with zinc-finger proteins in the databases. The BRCA1 gene is composed of 23 coding exons arrayed over more than 100 kb of genomic DNA (Fig. 6). Northern blots using fragments of the BRCA1 cDNA as probes identified a single transcript of about 7.8 kb, present most abundantly in breast, thymus and testis, and also present in ovary (Fig. 7). Four alternatively spliced products were observed as independent cDNA clones; 3 of these were detected in breast and 2 in ovary mRNA (Fig. 6). A PCR survey from tissue cDNAs further supports the idea that there is considerable heterogeneity near the 5' end of transcripts from this gene; the molecular basis for the heterogeneity

involves differential choice of the first splice donor site, and the changes detected all alter the transcript in the region 5' of the identified start codon. We have detected five potential alternate splice donors, the longest of which would add 1,050 bp to the sequence in SEQ ID NO:2. The predominate form of the BRCA1 protein in breast and ovary lacks exon 4. The nucleotide sequence for BRCA1 exon 4 is shown in SEQ ID NO:11, with the predicted amino acid sequence shown in SEQ ID NO:12.

Additional 5' sequence of BRCA1 genomic DNA is set forth in SEQ ID NO:13. The G at position 1 represents the potential start site in testis. The A in position 148 represents the potential start site in somatic tissue. The G at position 363 represents the canonical first splice donor site. The G at position 452 represents the first splice donor site in two clones (testis 1 and testis 2). The G at position 897 represents the first splice donor site in thymus 3. The T at position 1519 represents the first splice acceptor site. The A at position 1538 is the canonical start site, which appears at position 64 of SEQ ID NO:1. Partial genomic DNA sequences determined for BRCA1 are set forth in SEQ ID NO:14. The lower case letters denote intron sequence while the upper case letters denote exon sequence. Uncertainty about sequence at an intron/exon boundary is indicated by vvvvv. Indefinite intervals (between introns) are designated with vvvvvvvvvvvvvv. The intron/exon junctions are shown in Table 9.

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

GGKCTTACTC TGTTGTCCCA GCTGGAGTAC AGWGTGCGAT CATGAGGCTT ACTGTTGCTT	60
GACTCCTAGG CTCAAGCGAT CCTATCACCT CAGTCTCCAA GTAGCTGGAC T	111

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens

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

Xaa	Leu	Leu	Cys	Cys	Pro	Ser	Trp	Ser	Thr	Xaa	Cys	Asp	His	Glu	Ala
1			5					10						15	
Tyr	Cys	Cys	Leu	Thr	Pro	Arg	Leu	Lys	Arg	Ser	Tyr	His	Leu	Ser	Leu
			20				25					30			
Gln	Val	Ala	Gly	Leu											
			35												

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1540 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens

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

GAGGCTAGAG GGCAGGCACT TTATGGCAAA CTCAGGTAGA ATTCTTCCTC TTCCGTCTCT	60
TTCTTTTAC GTCATCCGGG GGCAGACTGG GTGGCCAATC CAGAGCCCCG AGAGACGCTT	120
GGCTCTTCT GTCCCTCCCA TCCTCTGATT GTACCTTGAT TTCGTATTCT GAGAGGCTGC	180
TGCTTAGCGG TAGCCCCTTG GTTCCCGTGG CAACGGAAAA GCGCGGGAAT TACAGATAAA	240
TTAAAACTGC GACTGCGCGG CGTGAGCTCG CTGAGACTTC CTGGACGGGG GACAGGCTGT	300

Exon No.	Base Position		Length	Intron Borders	
	5'	3'		5'	3'
e17	5110	5198	88	ATGATAATGGAATATTTGATTTAATTTTCAG ^{47*}	TATACCCAAGAACCTTTACAGAATACCTTK ⁴⁸
e18	5199	5277	78	TAATCCTTTGAGTGKTCTCATTCTGCAGA ⁴⁹	KAAGKATAATACTATWTCTCCCCYCCTCCC ⁵⁰
e19	5278	5316	38	TAACCTGTCTTTTCTATGATCTCTTTAGGG ⁵¹	GTAAGTACTTGATGTTACAACTAACCAGA ⁵²
e20	5317	5400	83	TTATCCTGATGGGTGTGNTGTTTCTTTTCAG ⁵³	GTAAAGCTCCCTCCCTCAAGTTGACNNAAA ⁵⁴
e21	5401	5455	54	TAATTTGGATTCCCGTCTCGGGTTCAGATC ⁵⁵	GTAAGAGCCTCGGGAGAANCCAGAGTTCCA ⁵⁶
e22	5456	5530	74	GTGATTTTAGATCCTAAATGTCCATTTTAG ⁵⁷	...CAGAGCAAGACCCTGTCTCAAAAACAAACA ⁵⁸
e23	5531	5590	59	TTGAATGCTCTTTCCTTCCTGGGGANCCAG ⁵⁹	TTTATCACTCATTACCTGGTGCTTGAGTAG ⁶⁰
e24	5591	5918	327	CTAATCTCTGCTTGTGTTCTCTGTCTCCAG ⁶¹	

* Numbers in superscript refer to SEQ ID NOS.

Low stringency blots in which genomic DNA from organisms of diverse phylogenetic background were probed with BRCA1 sequences that lack the zinc-finger region revealed strongly hybridizing fragments in human, monkey, sheep and pig, and very weak hybridization signals in rodents. This result indicates that, apart from the zinc-finger domain, BRCA1 is conserved only at a moderate level through evolution.

Germline BRCA1 mutations in 17q-linked kindreds. The most rigorous test for BRCA1 candidate genes is to search for potentially disruptive mutations in carrier individuals from kindreds that segregate 17q-linked susceptibility to breast and ovarian cancer. Such individuals must contain BRCA1 alleles that differ from the wildtype sequence. The set of DNA samples used in this analysis consisted of DNA from individuals representing 8 different BRCA1 kindreds (Table 10).

TABLE 10
KINDRED DESCRIPTIONS AND ASSOCIATED LOD SCORES

Kindred	Cases (n)			Sporadic Cases ¹ (n)	LOD Score	Marker(s)
	Br	Br	<50 Ov			
2082	31	20	22	7	9.49	D17S1327
2099	22	14	2*	0	2.36	D17S800/D17S855 ²
2035	10	8	1*	0	2.25	D17S1327
1901	10	7	1*	0	1.50	D17S855
1925	4	3	0	0	0.55	D17S579
1910	5	4	0	0	0.36	D17S579/D17S250 ²
1927	5	4	0	1	-0.44	D17S250
1911	8	5	0	2	-0.20	D17S250

¹ Number of women with breast cancer (diagnosed under age 50) or ovarian cancer (diagnosed at any age) who do not share the BRCA1-linked haplotype segregating in the remainder of the cases in the kindred.

² Multipoint LOD score calculated using both markers

* kindred contains one individual who had both breast and ovarian cancer; this individual is counted as a breast cancer case and as an ovarian cancer case.

The logarithm of the odds (LOD) scores in these kindreds range from 9.49 to -0.44 for a set of markers in 17q21. Four of the families have convincing LOD scores for linkage, and 4 have low positive or negative LOD scores. The latter kindreds were included because they demonstrate haplotype sharing at chromosome 17q21 for at least 3 affected members. Furthermore, all kindreds in the set display early age of breast cancer onset and 4 of the kindreds include at least one case of ovarian cancer, both hallmarks of BRCA1 kindreds. One kindred, 2082, has nearly equal incidence of breast and ovarian cancer, an unusual occurrence given the relative rarity of ovarian cancer in the population. All of the kindreds except two were ascertained in Utah. K2035 is from the midwest. K2099 is an African-American kindred from the southern USA.

In the initial screen for predisposing mutations in BRCA1, DNA from one individual who carries the predisposing haplotype in each kindred was tested. The 23 coding exons and associated splice junctions were amplified either from genomic DNA samples or from cDNA prepared from lymphocyte mRNA. When the amplified DNA sequences were compared to the wildtype sequence, 3 of the 8 kindred samples were found to contain sequence variants (Table 11).

TABLE 11
PREDISPOSING MUTATIONS

Kindred Number	Mutation	Coding Effect	Location*
2082	C→T	Gln→Stop	4000
1910	extra C	frameshift	5329
2099	T→G	Met→Arg	5387
2035	?	loss of transcript	
1901	11 bp deletion	frameshift	133

* In Sequence ID NO:1

All three sequence variants are heterozygous and each appears in only one of the kindreds. Kindred 2082 contains a nonsense mutation in coding exon 10 (Fig. 8A), Kindred 1910 contains a single nucleotide insertion in coding exon 19 (Fig. 8B), and Kindred 2099 contains a missense mutation in coding exon 20, resulting in a Met->Arg substitution (Fig. 8C). The frameshift and nonsense mutations are likely disruptive to the function of the BRCA1 product. The peptide encoded by the frameshift allele in Kindred 1910 would contain an altered amino acid sequence beginning 107

residues from the wildtype C-terminus. The peptide encoded by the frameshift allele in Kindred 1901 would contain an altered amino acid sequence beginning with the 24th residue from the wildtype N-terminus. The mutant allele in Kindred 2082 would encode a protein missing 548 residues from the C-terminus. The missense substitution observed in Kindred 2099 is potentially disruptive as it causes the replacement of a small hydrophobic amino acid (Met), by a large charged residue (Arg). Eleven common polymorphisms were also identified, 8 in coding sequence and 3 in introns.

The individual studied in Kindred 2035 evidently contains a regulatory mutation in BRCA1. In her cDNA, a polymorphic site (A->G at base 3921) appeared homozygous, whereas her genomic DNA revealed heterozygosity at this position (Fig. 8C). A possible explanation for this observation is that mRNA from her mutated BRCA1 allele is absent due to a mutation that affects its production or stability. This possibility was explored further by examining 5 polymorphic sites in the BRCA1 coding region, which are separated by as much as 3.5 kb in the BRCA1 transcript. In all cases where her genomic DNA appeared heterozygous for a polymorphism, cDNA appeared homozygous. In individuals from other kindreds and in non-haplotype carriers in Kindred 2035, these polymorphic sites could be observed as heterozygous in cDNA, implying that amplification from cDNA was not biased in favor of one allele. This analysis indicates that a BRCA1 mutation in Kindred 2035 either prevents transcription or causes instability or aberrant splicing of the BRCA1 transcript.

Cosegregation of BRCA1 mutations with BRCA1 haplotypes and population frequency analysis.

In addition to potentially disrupting protein function, two criteria must be met for a sequence variant to qualify as a candidate predisposing mutation. The variant must: 1) be present in individuals from the kindred who carry the predisposing BRCA1 haplotype and absent in other members of the kindred, and 2) be rare in the general population.

Each mutation was tested for cosegregation with BRCA1. For the frameshift mutation in Kindred 1910, two other haplotype carriers and one non-carrier were sequenced (Fig. 8B). Only the carriers exhibited the frameshift mutation. The C to T change in Kindred 2082 created a new AvrII restriction site. Other carriers and non-carriers in the kindred were tested for the presence of the restriction site (Fig. 8A). An allele-specific oligonucleotide (ASO) was designed to detect the presence of the sequence variant in Kindred 2099. Several individuals from the kindred, some known to carry the haplotype associated with the predisposing allele, and others known not to carry the associated haplotype, were screened by ASO for the mutation previously detected in the kindred.

In each kindred, the corresponding mutant allele was detected in individuals carrying the BRCA1-associated haplotype, and was not detected in noncarriers. In the case of the potential regulatory mutation observed in the individual from Kindred 2035, cDNA and genomic DNA from carriers in the kindred were compared for heterozygosity at polymorphic sites. In every instance, the
5 extinguished allele in the cDNA sample was shown to lie on the chromosome that carries the BRCA1 predisposing allele (Fig. 8C).

To exclude the possibility that the mutations were simply common polymorphisms in the population, ASOs for each mutation were used to screen a set of normal DNA samples. Gene frequency estimates in caucasians were based on random samples from the Utah population. Gene
10 frequency estimates in African-Americans were based on 39 samples provided by M. Peracek-Vance which originate from African-Americans used in her linkage studies and 20 newborn Utah African-Americans. None of the 3 potential predisposing mutations was found in the appropriate control population, indicating that they are rare in the general population. Thus, two important requirements for BRCA1 susceptibility alleles were fulfilled by the candidate predisposing mutations: 1)
15 cosegregation of the mutant allele with disease, and 2) absence of the mutant allele in controls, indicating a low gene frequency in the general population.

Phenotypic Expression of BRCA1 Mutations. The effect of the mutations on the BRCA1 protein correlated with differences in the observed phenotypic expression in the BRCA1 kindreds. Most BRCA1 kindreds have a moderately increased ovarian cancer risk, and a smaller subset have
20 high risks of ovarian cancer, comparable to those for breast cancer (Easton *et al.*, 1993). Three of the four kindreds in which BRCA1 mutations were detected fall into the former category, while the fourth (K2082) falls into the high ovarian cancer risk group. Since the BRCA1 nonsense mutation found in K2082 lies closer to the amino terminus than the other mutations detected, it might be expected to have a different phenotype. In fact, Kindred K2082 mutation has a high incidence of
25 ovarian cancer, and a later mean age at diagnosis of breast cancer cases than the other kindreds (Goldgar *et al.*, 1994). This difference in age of onset could be due to an ascertainment bias in the smaller, more highly penetrant families, or it could reflect tissue-specific differences in the behavior of BRCA1 mutations. The other 3 kindreds that segregate known BRCA1 mutations have, on average, one ovarian cancer for every 10 cases of breast cancer, but have a high proportion of breast
30 cancer cases diagnosed in their late 20's or early 30's. Kindred 1910, which has a frameshift mutation, is noteworthy because three of the four affected individuals had bilateral breast cancer, and

in each case the second tumor was diagnosed within a year of the first occurrence. Kindred 2035, which segregates a potential regulatory BRCA1 mutation, might also be expected to have a dramatic phenotype. Eighty percent of breast cancer cases in this kindred occur under age 50. This figure is as high as any in the set, suggesting a BRCA1 mutant allele of high penetrance (Table 10).

Although the mutations described above clearly are deleterious, causing breast cancer in women at very young ages, each of the four kindreds with mutations includes at least one woman who carries the mutation who lived until age 80 without developing a malignancy. It will be of utmost importance in the studies that follow to identify other genetic or environmental factors that may ameliorate the effects of BRCA1 mutations.

In four of the eight putative BRCA1-linked kindreds, potential predisposing mutations were not found. Three of these four have LOD scores for BRCA1-linked markers of less than 0.55. Thus, these kindreds may not in reality segregate BRCA1 predisposing alleles. Alternatively, the mutations in these four kindreds may lie in regions of BRCA1 that, for example, affect the level of transcript and therefore have thus far escaped detection. In particular, no mutation has been found in Kindred 1901, which has a phenotypic pattern similar to Kindred 2035 and is highly unlikely to be due to BRCA1. It may segregate a similar regulatory mutation, but is uninformative for all the polymorphisms examined to date.

Role of BRCA1 in Cancer. Most tumor suppressor genes identified to date give rise to protein products that are absent, nonfunctional, or reduced in function. The majority of TP53 mutations are missense; some of these have been shown to produce abnormal p53 molecules that interfere with the function of the wildtype product (Shaulian *et al.* 1992; Srivastava *et al.*, 1993). A similar dominant negative mechanism of action has been proposed for some adenomatous polyposis coli (APC) alleles that produce truncated molecules (Su *et al.*, 1993), and for point mutations in the Wilms' tumor gene (WT1) that alter DNA binding of the protein (Little *et al.*, 1993). The nature of the mutations observed in the BRCA1 coding sequence is consistent with production of either dominant negative proteins or nonfunctional proteins. The mutations identified thus far are all located in the C-terminal half of the peptide. The regulatory mutation inferred in Kindred 2035 cannot be a dominant negative; rather, this mutation likely causes reduction or complete loss of BRCA1 expression from the affected allele.

The BRCA1 protein contains a C₃HC₄ zinc-finger domain, similar to those found in numerous DNA binding proteins and implicated in zinc-dependent binding to nucleic acids. The first 180

amino acids of BRCA1 contain five more basic residues than acidic residues. In contrast, the remainder of the molecule is very acidic, with a net excess of 70 acidic residues. The excess negative charge is particularly concentrated near the C-terminus. Thus, one possibility is that BRCA1 encodes a transcription factor with an N-terminal DNA binding domain and a C-terminal transactivational "acidic blob" domain. Interestingly, another familial tumor suppressor gene, WT1, also contains a zinc-finger motif (Haber *et al.*, 1990). Many cancer predisposing mutations in WT1 alter zinc-finger domains (Little *et al.*, 1993; Haber *et al.*, 1990; Little *et al.*, 1992). WT1 encodes a transcription factor, and alternative splicing of exons that encode parts of the zinc-finger domain alter the DNA binding properties of WT1 (Bickmore *et al.*, 1992). Some alternatively spliced forms of WT1 mRNA generate molecules that act as transcriptional repressors (Drummond *et al.*, 1994). Some BRCA1 splicing variants may alter the zinc-finger motif, raising the possibility that a regulatory mechanism similar to that which occurs in WT1 may apply to BRCA1.

EXAMPLE 9

Analysis of Tumors for BRCA1 Mutations

To focus the analysis on tumors most likely to contain BRCA1 mutations, primary breast and ovarian carcinomas were typed for LOH in the BRCA1 region. Three highly polymorphic, simple tandem repeat markers were used to assess LOH: D17S1323 and D17S855, which are intragenic to BRCA1, and D17S1327, which lies approximately 100 kb distal to BRCA1. The combined LOH frequency in informative cases (i.e., where the germline was heterozygous) was 32/72 (44%) for the breast carcinomas and 12/21 (57%) for the ovarian carcinomas, consistent with previous measurements of LOH in the region (Futreal *et al.*, 1992; Jacobs *et al.*, 1993; Sato *et al.*, 1990; Eccles *et al.*, 1990; Cropp *et al.*, 1994). The analysis thus defined a panel of 32 breast tumors and 12 ovarian tumors of mixed race and age of onset to be examined for BRCA mutations. The complete 5,589 bp coding region and intron/exon boundary sequences of the gene were screened in this tumor set by direct sequencing alone or by a combination of single-strand conformation analysis (SSCA) and direct sequencing.

A total of six mutations were found, one in an ovarian tumor, four in breast tumors and one in a male unaffected haplotype carrier (Table 12). One mutation, Glu1548Ter, introduced a stop codon that would create a truncated protein missing 273 amino acids at the carboxy terminus. In addition, three missense mutations were identified. Two cases, Ala1715Glu and Met1782Arg,

involve substitutions of small, hydrophobic residues by charged residues, and a third, Pro1644Leu, found in an ovarian tumor, involves a more conservative substitution. The missense mutation in the male carrier involve a substitution of a charged residue by a hydrophobic residue.

TABLE 12
Predisposing Mutations

<u>Patient</u>	<u>Codon</u>	<u>Nucleotide Change</u>	<u>Amino Acid Change</u>	<u>Age of Onset</u>	<u>Family History</u>
BT098	1548	GAG → TAG	Glu → Stop	39	-
OV24	1644	CCA → CTA	Pro → Leu	44	-
BT106	1715	GCG → GAG	Ala → Glu	24	+
MC44	1782	ATG → AGG	Met → Arg	42	+
17764	954/5	2bp deletion	frameshift	31	+
19964	841	CGG → TGG	Arg → Trp		+

* Unaffected haplotype carrier, male

Several lines of evidence suggest that all four mutations represent BRCA1 susceptibility alleles:

- (i) all mutations are present in the germline;
- (ii) all are absent in appropriate control populations, suggesting they are not common polymorphisms;

- (iii) each mutant allele is retained in the tumor, as is the case in tumors from patients belonging to kindreds that segregate BRCA1 susceptibility alleles (Smith *et al.*, 1992; Kelsell *et al.*, 1993) (if the mutations represented neutral polymorphisms, they should be retained in only 50% of the cases);

- (iv) the age of onset in the four breast cancer cases with mutations varied between 24 and 42 years of age, consistent with the early age of onset of breast cancer in individuals with BRCA1 susceptibility; similarly, the ovarian cancer case was diagnosed at 44, an age that falls in the youngest 13% of all ovarian cancer cases; and finally,

- (v) three of the five cases have positive family histories of breast or ovarian cancer found retrospectively in their medical records, although the tumor set was not selected with regard to this criterion.

BT106 was diagnosed at age 24 with breast cancer. Her mother had ovarian cancer, her father had melanoma, and her paternal grandmother also had breast cancer. Patient MC44, an African-American, had bilateral breast cancer at age 42. This patient had a sister who died of breast cancer at age 34, another sister who died of lymphoma, and a brother who died of lung cancer. Her mutation (Met1782Arg) had been detected previously in Kindred 2099, an African-American family that segregates a BRCA1 susceptibility allele, and was absent in African-American and Caucasian controls. Patient MC44, to our knowledge, is unrelated to Kindred 2099. The detection of a rare mutant allele, once in a BRCA1 kindred and once in the germline of an apparently unrelated early-onset breast cancer case, suggests that the Met1782Arg change may be a common predisposing mutation in African-Americans. Collectively, these observations indicate that all four BRCA1 mutations in tumors represent susceptibility alleles; no somatic mutations were detected in the samples analyzed.

The paucity of somatic BRCA1 mutations is unexpected, given the frequency of LOH on 17q, and the usual role of susceptibility genes as tumor suppressors in cancer progression. There are three possible explanations for this result: (i) some BRCA1 mutations in coding sequences were missed by our screening procedure; (ii) BRCA1 somatic mutations fall primarily outside the coding exons; and (iii) LOH events in 17q do not reflect BRCA1 somatic mutations.

If somatic BRCA1 mutations truly are rare in breast and ovary carcinomas, this would have strong implications for the biology of BRCA1. The apparent lack of somatic BRCA1 mutations implies that there may be some fundamental difference in the genesis of tumors in genetically predisposed BRCA1 carriers, compared with tumors in the general population. For example, mutations in BRCA1 may have an effect only on tumor formation at a specific stage early in breast and ovarian development. This possibility is consistent with a primary function for BRCA1 in premenopausal breast cancer. Such a model for the role of BRCA1 in breast and ovarian cancer predicts an interaction between reproductive hormones and BRCA1 function. However, no clinical or pathological differences in familial versus sporadic breast and ovary tumors, other than age of onset, have been described (Lynch *et al.*, 1990). On the other hand, the recent finding of increased TP53 mutation and microsatellite instability in breast tumors from patients with a family history of breast cancer (Glebov *et al.*, 1994) may reflect some difference in tumors that arise in genetically predisposed persons. The involvement of BRCA1 in this phenomenon can now be addressed directly. Alternatively, the lack of somatic BRCA1 mutations may result from the existence of multiple genes that function in the same pathway of tumor suppression as BRCA1, but which collectively represent

a more favored target for mutation in sporadic tumors. Since mutation of a single element in a genetic pathway is generally sufficient to disrupt the pathway, BRCA1 might mutate at a rate that is far lower than the sum of the mutational rates of the other elements.

EXAMPLE 10

Analysis of the BRCA1 Gene

The structure and function of BRCA 1 gene are determined according to the following methods.

Biological Studies. Mammalian expression vectors containing BRCA1 cDNA are constructed and transfected into appropriate breast carcinoma cells with lesions in the gene. Wild-type BRCA1 cDNA as well as altered BRCA1 cDNA are utilized. The altered BRCA1 cDNA can be obtained from altered BRCA1 alleles or produced as described below. Phenotypic reversion in cultures (e.g., cell morphology, doubling time, anchorage-independent growth) and in animals (e.g., tumorigenicity) is examined. The studies will employ both wild-type and mutant forms (Section B) of the gene.

Molecular Genetics Studies. *In vitro* mutagenesis is performed to construct deletion mutants and missense mutants (by single base-pair substitutions in individual codons and cluster charged → alanine scanning mutagenesis). The mutants are used in biological, biochemical and biophysical studies.

Mechanism Studies. The ability of BRCA1 protein to bind to known and unknown DNA sequences is examined. Its ability to transactivate promoters is analyzed by transient reporter expression systems in mammalian cells. Conventional procedures such as particle-capture and yeast two-hybrid system are used to discover and identify any functional partners. The nature and functions of the partners are characterized. These partners in turn are targets for drug discovery.

Structural Studies. Recombinant proteins are produced in *E. coli*, yeast, insect and/or mammalian cells and are used in crystallographical and NMR studies. Molecular modeling of the proteins is also employed. These studies facilitate structure-driven drug design.

Industrial Utility

As previously described above, the present invention provides materials and methods for use in testing BRCA1 alleles of an individual and an interpretation of the normal or predisposing nature of the alleles. Individuals at higher than normal risk might modify their lifestyles appropriately. In the case of BRCA1, the most significant non-genetic risk factor is the protective effect of an early,

full term pregnancy. Therefore, women at risk could consider early childbearing or a therapy designed to simulate the hormonal effects of an early full-term pregnancy. Women at high risk would also strive for early detection and would be more highly motivated to learn and practice breast self examination. Such women would also be highly motivated to have regular mammograms, perhaps starting at an earlier age than the general population. Ovarian screening could also be undertaken at greater frequency. Diagnostic methods based on sequence analysis of the BRCA1 locus could also be applied to tumor detection and classification. Sequence analysis could be used to diagnose precursor lesions. With the evolution of the method and the accumulation of information about BRCA1 and other causative loci, it could become possible to separate cancers into benign and malignant.

Women with breast cancers may follow different surgical procedures if they are predisposed, and therefore likely to have additional cancers, than if they are not predisposed. Other therapies may be developed, using either peptides or small molecules (rational drug design). Peptides could be the missing gene product itself or a portion of the missing gene product. Alternatively, the therapeutic agent could be another molecule that mimics the deleterious gene's function, either a peptide or a nonpeptidic molecule that seeks to counteract the deleterious effect of the inherited locus. The therapy could also be gene based, through introduction of a normal BRCA1 allele into individuals to make a protein which will counteract the effect of the deleterious allele. These gene therapies may take many forms and may be directed either toward preventing the tumor from forming, curing a cancer once it has occurred, or stopping a cancer from metastasizing.

It will be appreciated that the methods and compositions of the instant invention can be incorporated in the form of a variety of embodiments, only a few of which are disclosed herein. It will be apparent to the artisan that other embodiments exist and do not depart from the spirit of the invention. Thus, the described embodiments are illustrative and should not be construed as restrictive.

LIST OF REFERENCES

- Altschul, S.F. *et al.* (1990). *J. Mol. Biol.* **215**, 195-197.
- American Cancer Society, Cancer Facts & Figures - 1992. (American Cancer Society, Atlanta, GA,).
- Anand, R. (1992). *Techniques for the Analysis of Complex Genomes*, (Academic Press).
- Anderson, *et al.* (1980). *Proc. Nat. Acad. Sci. USA* **77**:5399-5403.
- Anderson, J.A., *et al.* (1992). *J. Ontology* **21**:321.
- Anderson, D.E. (1992). *J. Natl. Cancer Inst.* **48**:1029-1034.
- Ausebel, F.M., *et al.* (1992). *Current Protocols in Molecular Biology*, (J. Wiley and Sons, N.Y.)
- Beaucage & Carruthers (1981). *Tetra. Letts.* **22**:1859-1862.
- Berkner, *et al.* (1988). *BioTechniques* **6**:616-629.
- Berkner (1992). *Curr. Top. Microbiol. Immunol.* **158**:39-61.
- Bickmore, W.A., *et al.* (1992). *Science* **257**, 235-7.
- Bishop, D.T., *et al.* (1988). *Genet. Epidemiol.* **5**:151-169.
- Bishop, D.T. and Gardner, E.J. (1980). In: Banbury Report 4: Cancer Incidence in Defined Populations (J. Cairns, J.L. Lyon, M. Skolnick, eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 309-408.
- Botstein, *et al.* (1980). *Am. J. Hum. Genet.* **32**:314-331.
- Bradyopadhyay and Temin (1984). *Mol. Cell. Biol.* **4**:749-754.
- Breakfield and Geller (1987). *Mol. Neurobiol.* **1**:337-371.
- Brinster, *et al.* (1981). *Cell* **27**:223-231.
- Buchschacher and Panganiban (1992). *J. Virol.* **66**:2731-2739.
- Buckler, *et al.* (1991). *Proc. Nat. Acad. Sci. USA* **88**:4005-4009.
- Cannon-Albright, L., *et al.* (1994). *Cancer Research* **54**:2378-2385.
- Capecci, M.R. (1989). *Science* **244**:1288.

- Cariello (1988). *Human Genetics* 42:726.
- Claus, E., et al. (1991). *Am. J. Hum. Genet.* 48:232-242.
- Conner, B.J., et al. (1983). *Proc. Nat. Acad. Sci. USA.* _____
- Constantini and Lacy (1981). *Nature* 294:92-94.
- Cotten, et al. (1990). *Proc. Nat. Acad. Sci. USA* 87:4033-4037.
- Cotton, et al. (1988). *Proc. Nat. Acad. Sci. USA* 85:4397.
- Cropp, C.S., et al. (1994). *Cancer Res.* 54:2548-2551.
- Culver, et al. (1992). *Science* 256:1550-1552.
- Curiel, et al. (1991). *Proc. Nat. Acad. Sci. USA* 88:8850-8854.
- Curiel, et al. (1991). *Hum. Gene Ther.* 3:147-154.
- Deutscher, M. (1990). *Meth. Enzymology* 182, (Academic Press, San Diego, Cal.).
- Donehower, L.A., et al. (1992). *Nature* 356:215.
- Drummond, I. A., et al. (1994). *Mol. Cell Biol.* 14, 3800-9.
- Easton, D., et al. (1993). *Am. J. Hum. Genet.* 52:678-701.
- Eccles, D.M., et al. (1990). *Oncogene* 5:1599-1601.
- Erickson, J. et al., (1990). *Science* 249:527-533.
- Fain, P.R. (1992). *Cytogen. Cell Genet.* 60:178.
- Felgner, et al. (1987). *Proc. Nat. Acad. Sci. USA* 84:7413-7417.
- Fiers, et al. (1978). *Nature* 273:113.
- Fink, et al. (1992). *Hum. Gene Ther.* 3:11-19.
- Finkelstein, J., et al. (1990). *Genomics* 7:167-172.
- Freese, et al. (1990). *Biochem. Pharmacol.* 40:2189-2199.
- Friedman, T. (1991). In *Therapy for Genetic Diseases*, T. Friedman, ed., Oxford University Press, pp. 105-121.
- Futreal (1993). Ph.D. Thesis, University of North Carolina, Chapel Hill.

- Futreal, A., *et al.* (1992). *Hum. Molec. Genet.* 1:66.
- Futreal, P.A., *et al.* (1992). *Cancer Res.* 52:2624-2627.
- Glebov, O.K., *et al.* (1994). *Cancer Res.* 54:3703-3709.
- Glover, D. (1985). *DNA Cloning, I and II*, (Oxford Press).
- Go, R.C.P., *et al.* (1983). *J. Natl. Cancer Inst.* 71:455-461.
- Goding (1986). *Monoclonal Antibodies: Principles and Practice*, 2d ed. (Academic Press, N.Y.).
- Godowski, *et al.* (1988). *Science* 241:812-816.
- Goldgar, D.E., *et al.* (1994). *J. Natl. Can. Inst.* 86:3:200-209.
- Gordon, *et al.* (1980). *Proc. Nat. Acad. Sci. USA* 77:7380-7384.
- Gorziglia and Kapikian (1992). *J. Virol.* 66:4407-4412.
- Graham and van der Eb (1973). *Virology* 52:456-467.
- Guthrie, G. & Fink, G.R. (1991). *Guide to Yeast Genetics and Molecular Biology*, (Academic Press).
- Haber, D. A., *et al.* (1990). *Cell* 61:1257-69.
- Hall, J.M., *et al.* (1992). *Am. J. Hum. Genet.* 50:1235-1241.
- Hall, J.M., *et al.* (1990). *Science* 250:1684-1689.
- Harlow & Lane (1988). Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).
- Hasty, P., K., *et al.* (1991). *Nature* 350:243.
- Helseth, *et al.* (1990). *J. Virol.* 64:2416-2420.
- Hodgson, J. (1991). *Bio/Technology* 9:19-21.
- Huse, *et al.* (1989). *Science* 246:1275-1281.
- Innis *et al.* (1990). *PCR Protocols: A Guide to Methods and Applications*, (Academic Press, San Diego, Cal.).
- Jacobs, I.J., *et al.* (1993). *Cancer Res.* 53:1218-1221.
- Jeffreys, *et al.* (1985). *Nature* 314:67-73.

- Johnson, J.L., *et al.* (1994). *Mol. Cell Biol.* 14:1956-63.
- Johnson, *et al.* (1992). *J. Virol.* 66:2952-2965.
- Kamb, A. *et al.* (1994). *Science* 264:436-440.
- Kandpal, *et al.* (1990). *Nucl. Acids Res.* 18:1789-1795.
- Kaneda, *et al.* (1989). *J. Biol. Chem.* 264:12126-12129.
- Kanehisa (1984). *Nucl. Acids Res.* 12:203-213.
- Kelsell, D.P., *et al.* (1993). *Human Mol. Genet.* 2:1823-1828.
- Kinszler, K.W., *et al.* (1991). *Science* 251:1366-1370.
- Knudson, A.G. (1993). *Nature Genet.* 5:103.
- Kozak, M. (1987). *Nucleic Acids Res.* 15:8125-8148.
- Kubo, T., *et al.* (1988). *FEBS Letts.* 241:119.
- Landegren, *et al.* (1988). *Science* 242:229.
- Lim, *et al.* (1992). *Circulation* 83:2007-2011.
- Lindsay, S., *et al.* (1987). *Nature* 327:336-368.
- Litt, *et al.* (1989). *Am. J. Hum. Genet.* 44:397-401.
- Lovett, *et al.* (1991). *Proc. Nat. Acad. Sci. USA* 88:9628-9632.
- Lynch, H.T., *et al.* (1990). *Gynecol. Oncol.* 36:48-55.
- Madzak, *et al.* (1992). *J. Gen. Virol.* 73:1533-1536.
- Malkin, D., *et al.* (1990). *Science* 250:1233-1238.
- Maniatis, T., *et al.* (1989). *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Mann and Baltimore (1985). *J. Virol.* 54:401-407.
- Margaritte, *et al.* (1992). *Am. J. Hum. Genet.* 50:1231-1234.
- Margolskee (1992). *Curr. Top. Microbiol. Immunol.* 158:67-90.
- Matthews & Kricka (1988). *Anal. Biochem.* 169:1.

- Merrifield (1963). *J. Am. Chem. Soc.* **85**:2149-2156.
- Mettlin, C., et al. (1990). *American Journal of Epidemiology* **131**:973-983.
- Metzger, et al. (1988). *Nature* **334**:31-36.
- Miller (1992). *Curr. Top. Microbiol. Immunol.* **158**:1-24.
- Miller, et al. (1985). *Mol. Cell. Biol.* **5**:431-437.
- Miller, et al. (1988). *J. Virol.* **62**:4337-4345.
- Mittlin (1989). *Clinical Chem.* **35**:1819.
- Modrich, P. (1991). *Ann. Rev. Genet.* **25**:229-253.
- Mombaerts, P., et al. (1992). *Cell* **68**:869.
- Monaco, et al. (1986). *Nature* **323**:646.
- Moss (1992). *Curr. Top. Microbiol. Immunol.* **158**:25-38.
- Muzyczka (1992). *Curr. Top. Microbiol. Immunol.* **158**:97-123.
- Nabel, et al. (1990). *Science* **249**:1285-1288.
- Nabel (1992). *Hum. Gene Ther.* **3**:399-410.
- Nakamura, et al. (1987). *Science* **235**:1616-1622.
- Narod, S.A., et al. (1991). *The Lancet* **338**:82-83.
- Newman, B., et al. (1988). *Proc. Nat. Acad. Sci. USA* **85**:3044-3048.
- Novack, et al. (1986). *Proc. Nat. Acad. Sci. USA* **83**:586.
- Oh, J. (1985). *Analysis of Human Genetic Linkage*, Johns Hopkins University Press, Baltimore, Md, pp. 1-216.
- Ohi, et al. (1990). *Gene* **89**:279-282.
- Oliphant, A., et al. (1991). *Nucleic Acid Res.* **19**:4794.
- Oliphant, A., et al. (1991). *Nucleic Acid Res.* **19**:4795.
- Orita, et al. (1989). *Proc. Nat. Acad. Sci. USA* **86**:276-2770.
- Orita, et al. (1989). *Genomics* **5**:874-879.

- Page, *et al.* (1990). *J. Virol.* **64**:5370-5276.
- Pellicer, *et al.* (1980). *Science* **209**:1414-1422.
- Petropoulos, *et al.* (1992). *J. Virol.* **66**:3391-3397.
- Philpott, K.L., *et al.* (1992). *Science* **256**:1448.
- Pierce, *et al.* (1992). *Proc. Nat. Acad. Sci. USA* **89**:2056-2060.
- Quantin, *et al.* (1992). *Proc. Natl. Acad. Sci. USA* **89**:2581-2584.
- Rano & Kidd (1989). *Nucl. Acids Res.* **17**:8392.
- Rosenfeld, *et al.* (1992). *Cell* **68**:143-155.
- Sambrook, J., *et al.* (1989). *Molecular Cloning: A Laboratory Manual*, 2nd Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).
- Sato, T., *et al.* (1990). *Cancer Res.* **50**:7184-7189.
- Scharf, (1986). *Science* **233**:1076.
- Scopes, R, (1982). *Protein Purification: Principles and Practice*, (Springer-Verlag, N.Y.).
- Shaulian, E., *et al.* (1992). *Mol. Cell Biol.* **12**, 5581-92.
- Sheffield, V.C., *et al.* (1989). *Proc. Nat. Acad. Sci. USA* **86**: 232-236.
- Smith, T.F., *et al.* (1981). *J. Mol. Biol.* **147**:195-197.
- Srivastava, S., *et al.* (1993). *Cancer Res.* **53**, 4452-5.
- Swift, M., *et al.* (1991). *N. Engl. J. Med.* **325**:1831-1836.
- Swift, M., *et al.* (1976). *Cancer Res.* **36**:209-215.
- Su, L. K., *et al.* (1993). *Cancer Res* **53**, 2728-31.
- Shenk, *et al.* (1975). *Proc. Nat. Acad. Sci. USA* **72**:989.
- Shimada, *et al.* (1991). *J. Clin. Invest.* **88**:1043-1047.
- Shinkai, Y., *et al.* (1992). *Cell* **68**:855.
- Skolnick, M.H. (1980). Banbury Report No 4: Cancer Incidence in Defined Populations (J. Cairns, J. Lyon, M. Skolnick, eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., pp. 285-297.

- Skolnick, M.H. and Wallace, B.R. (1988). *Genomics* 2:273-279.
- Skolnick, M.H. (1977). *Genet. Hum. Cancer*, (J.J. Mulvill, R.W. Miller and J.F. Freeman, Jr., eds.), Raven Press, N.Y.
- Skolnick, M.H., *et al.* (1990). *Science* 250:1715-1720.
- Smith, S.A., *et al.* (1992). *Nature Genetics* 2: 128-131.
- Snouwaert, J.N., *et al.* (1992). *Science* 257:1083.
- Sorge, *et al.* (1984). *Mol. Cell. Biol.* 4:1730-1737.
- Sternberg (1990). *Proc. Nat. Acad. Sci. USA* 87:103-107.
- Sternberg, *et al.* (1990). *The New Biologist* 2:151-162.
- Stewart, *et al.* (1992). *Hum. Gene Ther.* 3:267-275.
- Stratford-Perricaudet, *et al.* (1990). *Hum. Gene Ther.* 1:241-256.
- Thomas, A., *et al.*, *IMA Journal of Mathematics Applied in Medicine and Biology*, (in press).
- Tonolio, D., *et al.* (1990). Cold Spring Harbor Conference.
- Valancius, V. & Smithies, O. (1991). *Mol. Cell Biol.* 11:1402.
- van Dilla, *et al.* (1986). *Biotechnology* 4:537-552.
- Wagner, *et al.* (1991). *Proc. Nat. Acad. Sci. USA* 88:4255-4259.
- Wagner, *et al.* (1990). *Proc. Nat. Acad. Sci. USA* 87:3410-3414.
- Wang and Huang (1989). *Biochemistry* 28:9508-9514.
- Wartell, R.M., *et al.* (1990). *Nucl. Acids Res.* 18:2699-2705.
- Weber and May (1989). *Am. J. Hum. Genet.* 44:388-396.
- Weber, J.L. (1990). *Genomics* 7:524-530.
- Weber, J.L., *et al.* (1990). *Nucleic Acid Res.* 18:4640.
- Wells, J.A. (1991). *Methods in Enzymol.* 202:390-411.
- Wetmur & Davidson (1968). *J. Mol. Biol.* 31:349-370.
- White & Lalouel (1988). *Ann. Rev. Genet.* 22:259-279.

- Wilkinson, *et al.* (1992). *Nucleic Acids Res.* **20**:2233-2239.
- Willams and Anderson (1984). *Genet. Epidemiol.* **1**:7-20.
- Wolff, *et al.* (1990). *Science* **247**:1465-1468.
- Wolff, *et al.* (1991). *BioTechniques* **11**:474-485.
- Wooster, R., *et al.* (1994). *Science*, (In press).
- Wu, *et al.* (1989). *J. Biol. Chem.* **264**:16985-16987.
- Wu, *et al.* (1991). *J. Biol. Chem.* **266**:14338-14342.
- Wu, *et al.* (1989). *Genomics* **4**:560-569.
- Zenke, *et al.* (1990). *Proc. Nat. Acad. Sci. USA* **87**:3655-3659.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Skolnick, Mark H.
Goldgar, David E.
Miki, Yoshio
Swenson, Jeff
Kamb, Alexander
Harshman, Keith D.
Shattuck-Eidens, Donna M.
Tavtigian, Sean V.
- (ii) TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
Susceptability Gene
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Venable, Baetjer, Howard & Civiletti
(B) STREET: 1201 New York Avenue, N.W., Suite 1000
(C) CITY: Washington
(D) STATE: DC
(F) ZIP: 20005
- (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 #1.25
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE: 02-SEP-1994
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/300,266
(B) FILING DATE: 02-SEP-1994
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/289,221
(B) FILING DATE: 12-AUG-1994
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/246,602
(B) FILING DATE: 20-MAY-1994
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Ihnen, Jeffrey L.
(B) REGISTRATION NUMBER: 28,957
(C) REFERENCE/DOCKET NUMBER: 24884-109347-2
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 202-962-4810
(B) TELEFAX: 202-962-8300

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5857 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 64..5769

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

GGCCCCTGCG CTCAGGAGGC CTTCAACCTC TGCTCTGGGT AAAGTTCATT GGAACAGAAA	60
GAAATGGATT TATCTGCTCT TCGCGTTGAA GAAGTACAAA ATGTCATTAA TGCTATGCAG	120
AAAATCTTAG AGTGTCCCAT CTGTCTGGAG TTGATCAAGG AACCTGTCTC CACAAAGTGT	180
GACCACATAT TTTGCAAATT TTGCATGCTG AAACCTCTCA ACCAGAAGAA AGGGCCTTCA	240
CAGTGTCCCTT TATGTAAGAA TGATATAACC AAAAGGAGCC TACAAGAAAAG TACGAGATTT	300
AGTCAACTTG TTGAAGAGCT ATTGAAAATC ATTTGTGCTT TTCAGCTTGA CACAGGTTTG	360
GAGTATGCAA ACAGCTATAA TTTTGCAAAA AAGGAAAATA ACTCTCCTGA ACATCTAAAA	420
GATGAAGTTT CTATCATCCA AAGTATGGGC TACAGAAAACC GTGCCAAAAG ACTTCTACAG	480
AGTGAACCCG AAAATCCTTC CTTGCAGGAA ACCAGTCTCA GTGTCCAACCT CTCTAACCTT	540
GGAACGTGTA GAACTCTGAG GACAAAGCAG CGGATACAAC CTCAAAAGAC GTCTGTCTAC	600
ATTGAATTGG GATCTGATTC TTCTGAAGAT ACCGTTAATA AGGCAACTTA TTGCAGTGTG	660
GGAGATCAAG AATTGTTACA AATCACCCCT CAAGGAACCA GGGATGAAAT CAGTTTGGAT	720
TCTGCAAAAA AGGCTGCTTG TGAATTTTCT GAGACGGATG TAACAAATAC TGAACATCAT	780
CAACCCAGTA ATAATGATTT GAACACCACT GAGAAGCGTG CAGCTGAGAG GCATCCAGAA	840
AAGTATCAGG GTAGTTCTGT TTCAAACCTG CATGTGGAGC CATGTGGCAC AAATACTCAT	900
GCCAGCTCAT TACAGCATGA GAACAGCAGT TTATTACTCA CTAAAGACAG AATGAATGTA	960
GAAAAGGCTG AATTCTGTAA TAAAAGCAAA CAGCCTGGCT TAGCAAGGAG CCAACATAAC	1020
AGATGGGCTG GAAGTAAGGA AACATGTAAT GATAGGCGGA CTCCCAGCAC AGAAAAAAG	1080
GTAGATCTGA ATGCTGATCC CCTGTGTGAG AGAAAAGAAT GGAATAAGCA GAAACTGCCA	1140
TGCTCAGAGA ATCCTAGAGA TACTGAAGAT GTTCCTTGGA TAACACTAAA TAGCAGCATT	1200
CAGAAAGTTA ATGAGTGGTT TTCCAGAAGT GATGAACTGT TAGGTTCTGA TGAATCACAT	1260
GATGGGGAGT CTGAATCAAA TGCCAAAGTA GCTGATGTAT TGGACGTCCT AATTGAGGTA	1320
GATGAATATT CTGGTTCTTC AGAGAAAATA GACTTACTGG CCAGTGATCC TCATGAGGCT	1380
TTAATATGTA AAAGTGACAG AGTTCACCTC AAATCAGTAG AGAGTAATAT TGAAGGCCAA	1440
ATATTTGGGA AAACCTATCG GAAGAAGGCA AGCCTCCCCA ACTTAAGCCA TGTAAGTGAA	1500
AATCTAATTA TAGGNGCATT TGTTTCTGAG CCACAGATAA TACAAGAGCG TCCCCTCACA	1560

AATAAATTAA	AGCGTAAAAG	GAGACCTACA	TCAGGCCTTC	ATCCTGAGGA	TTTTATCAAG	1620
AAAGCAGATT	TGGCAGTTCA	AAAGACTCCT	GAAATGATAA	ATCAGGGAAC	TAACCAAACG	1680
GAGCAGAATG	GTCAAGTGAT	GAATATTACT	AATAGTGGTC	ATGAGAATAA	AACAAAAGGT	1740
GATTCTATT	AGAATGAGAA	AAATCCTAAC	CCAATAGAAT	CACTCGAAAA	AGAATCTGCT	1800
TTCAAAACGA	AAGCTGAACC	TATAAGCAGC	AGTATAAGCA	ATATGGAAC	CGAATTAAAT	1860
ATCCACAATT	CAAAAGCACC	TAAAAAGAAT	AGGCTGAGGA	GGAAGTCTTC	TACCAGGCAT	1920
ATTCATGCGC	TTGAACTAGT	AGTCAGTAGA	AATCTAAGCC	CACCTAATTG	TACTGAATTG	1980
CAAATTGATA	GTTGTTCTAG	CAGTGAAGAG	ATAAAGAAAA	AAAAGTACAA	CCAAATGCCA	2040
GTCAGGCACA	GCAGAAACCT	ACAACTCATG	GAAGGTAAAG	AACCTGCAAC	TGGAGCCAAG	2100
AAGAGTAACA	AGCCAAATGA	ACAGACAAGT	AAAAGACATG	ACAGTGATAC	TTTCCCAGAG	2160
CTGAAGTTAA	CAAATGCACC	TGGTTCTTTT	ACTAAGTGTT	CAAATACCAG	TGAACTTAAA	2220
GAATTTGTCA	ATCCTAGCCT	TCCAAGAGAA	GAAAAAGAAG	AGAACTAGA	AACAGTTAAA	2280
GTGTCTAATA	ATGCTGAAGA	CCCCAAAGAT	CTCATGTTAA	GTGGAGAAAG	GGTTTTGCAA	2340
ACTGAAAGAT	CTGTAGAGAG	TAGCAGTATT	TCCTGGTAC	CTGGTACTGA	TTATGGCACT	2400
CAGGAAAGTA	TCTCGTTACT	GGAAGTTAGC	ACTCTAGGGA	AGGCAAAAAC	AGAACCAAAT	2460
AAATGTGTGA	GTCAGTGTGC	AGCATTGTGA	AACCCCAAGG	GACTAATTCA	TGGTTGTTCC	2520
AAAGATAATA	GAAATGACAC	AGAAGGCTTT	AAGTATCCAT	TGGGACATGA	AGTTAACCAC	2580
AGTCGGGAAA	CAAGCATAGA	AATGGAAGAA	AGTGAACCTG	ATGCTCAGTA	TTTGCAGAAT	2640
ACATTCAAGG	TTTCAAAGCG	CCAGTCATTT	GCTCTGTTTT	CAAATCCAGG	AAATGCAGAA	2700
GAGGAATGTG	CAACATTCTC	TGCCCACTCT	GGGTCCTTAA	AGAAACAAAG	TCCAAAAGTC	2760
ACTTTTGAAT	GTGAACAAAA	GGAAGAAAAT	CAAGGAAAGA	ATGAGTCTAA	TATCAAGCCT	2820
GTACAGACAG	TTAATATCAC	TGCAGGCTTT	CCTGTGGTTG	GTCAGAAAGA	TAAGCCAGTT	2880
GATAATGCCA	AATGTAGTAT	CAAAGGAGGC	TCTAGGTTTT	GTCTATCATC	TCAGTTCAGA	2940
GGCAACGAAA	CTGGACTCAT	TACTCCAAAT	AAACATGGAC	TTTTACAAAA	CCCATATCGT	3000
ATACCACCAC	TTTTTCCCAT	CAAGTCATTT	GTTAAAACTA	AATGTAAGAA	AAATCTGCTA	3060
GAGGAAAAC	TTGAGGAACA	TTCAATGTCA	CCTGAAAGAG	AAATGGGAAA	TGAGAACATT	3120
CCAAGTACAG	TGAGCACAAT	TAGCCGTAAT	AACATTAGAG	AAAATGTTTT	TAAAGAAGCC	3180
AGCTCAAGCA	ATATTAATGA	AGTAGGTTCC	AGTACTAATG	AAGTGGGCTC	CAGTATTAAT	3240
GAAATAGGTT	CCAGTGATGA	AAACATTCAA	GCAGAACTAG	GTAGAAACAG	AGGGCCAAAA	3300
TTGAATGCTA	TGCTTAGATT	AGGGGTTTTG	CAACCTGAGG	TCTATAAACA	AAGTCTTCCT	3360
GGAAGTAATT	GTAAGCATCC	TGAAATAAAA	AAGCAAGAAT	ATGAAGAAGT	AGTTCAGACT	3420
GTTAATACAG	ATTTCTCTCC	ATCTCTGATT	TCAGATAACT	TAGAACAGCC	TATGGGAAGT	3480
AGTCATGCAT	CTCAGGTTTG	TTCTGAGACA	CCTGATGACC	TGTTAGATGA	TGGTGAAATA	3540
AAGGAAGATA	CTAGTTTTGC	TGAAAATGAC	ATTAAGGAAA	GTTCTGCTGT	TTTTAGCAAA	3600

AGCGTCCAGA	AAGGAGAGCT	TAGCAGGAGT	CCTAGCCCTT	TCACCCATAC	ACATTTGGCT	3660
CAGGGTTACC	GAAGAGGGGC	CAAGAAATTA	GAGTCCTCAG	AAGAGAACTT	ATCTAGTGAG	3720
GATGAAGAGC	TTCCCTGCTT	CCAACACTTG	TTATTTGGTA	AAGTAAACAA	TATACCTTCT	3780
CAGTCTACTA	GGCATAGCAC	CGTTGCTACC	GAGTGTCTGT	CTAAGAACAC	AGAGGAGAAT	3840
TTATTATCAT	TGAAGAATAG	CTTAAATGAC	TGCAGTAACC	AGGTAATATT	GGCAAAGGCA	3900
TCTCAGGAAC	ATCACCTTAG	TGAGGAAACA	AAATGTTCTG	CTAGCTTGTT	TTCTTCACAG	3960
TGCAGTGAAT	TGGAAGACTT	GACTGCAAAT	ACAAACACCC	AGGCTCCTTT	CTTGATTGGT	4020
TCTTCCAAAC	AAATGAGGCA	TCAGTCTGAA	AGCCAGGGAG	TTGGTCTGAG	TGACAAGGAA	4080
TTGGTTTCAG	ATGATGAAGA	AAGAGGAACG	GGCTTGGAAG	AAAATAATCA	AGAAGAGCAA	4140
AGCATGGATT	CAAACCTTAG	TGAAGCAGCA	TCTGGGTGTG	AGAGTGAAAC	AAGCGTCTCT	4200
GAAGACTGCT	CAGGGCTATC	CTCTCAGAGT	GACATTTTAA	CCACTCAGCA	GAGGGATACC	4260
ATGCAACATA	ACCTGATAAA	GCTCCAGCAG	GAAATGGCTG	AACTAGAAGC	TGTGTTAGAA	4320
CAGCATGGGA	GCCAGCCTTC	TAACAGCTAC	CCTTCCATCA	TAAGTGACTC	TTCTGCCCTT	4380
GAGGACCTGC	GAAATCCAGA	ACAAAGCACA	TCAGAAAAAG	CAGTATTAAC	TTACAGAAAA	4440
AGTAGTGAAT	ACCCTATAAG	CCAGAATCCA	GAAGGCCTNT	CTGCTGACAA	GTTTGAGGTG	4500
TCTGCAGATA	GTTCTACCAG	TAAAAATAAA	GAACCAGGAG	TGGAAAGGTC	ATCCCCTTCT	4560
AAATGCCCAT	CATTAGATGA	TAGGTGGTAC	ATGCACAGTT	GCTCTGGGAG	TCTTCAGAAAT	4620
AGAAACTACC	CACCTCAAGA	GGAGCTCATT	AAGGTTGTTG	ATGTGGAGGA	GCAACAGCTG	4680
GAAGAGTCTG	GGCCACACGA	TTTGACGGAA	ACATCTTACT	TGCCAAGGCA	AGATCTAGAG	4740
GGAACCCCTT	ACCTGGAATC	TGGAATCAGC	CTCTTCTCTG	ATGACCCTGA	ATCTGATCCT	4800
TCTGAAGACA	GAGCCCCAGA	GTCAGCTCGT	GTTGGCAACA	TACCATCTTC	AACCTCTGCA	4860
TTGAAAGTTC	CCCAATTGAA	AGTTGCAGAA	TCTGCCCAGA	GTCCAGCTGC	TGCTCATACT	4920
ACTGATACTG	CTGGGTATAA	TGCAATGGAA	GAAAGTGTGA	GCAGGGAGAA	GCCAGAATTG	4980
ACAGCTTCAA	CAGAAAGGGT	CAACAAAAGA	ATGTCCATGG	TGGTGTCTGG	CCTGACCCCCA	5040
GAAGAATTTA	TGCTCGTGTA	CAAGTTTGCC	AGAAAACACC	ACATCACTTT	AACTAATCTA	5100
ATTACTGAAG	AGACTACTCA	TGTTGTTATG	AAAACAGATG	CTGAGTTTGT	GTGTGAACGG	5160
ACACTGAAAT	ATTTTCTAGG	AATTGCGGGA	GGAAAATGGG	TAGTTAGCTA	TTTCTGGGTG	5220
ACCCAGTCTA	TTAAAGAAAAG	AAAAATGCTG	AATGAGCATG	ATTTTGAAGT	CAGAGGAGAT	5280
GTGGTCAATG	GAAGAAAACCA	CCAAGGTCCA	AAGCGAGCAA	GAGAAATCCCA	GGACAGAAAAG	5340
ATCTTCAGGG	GGCTAGAAAT	CTGTTGCTAT	GGGCCCTTCA	CCAACATGCC	CACAGATCAA	5400
CTGGAATGGA	TGGTACAGCT	GTGTGGTGCT	TCTGTGGTGA	AGGAGCTTTC	ATCATTCACC	5460
CTTGGCACAG	GTGTCCACCC	AATTGTGGTT	GTGCAGCCAG	ATGCCTGGAC	AGAGGACAAT	5520
GGCTTCCATG	CAATTGGGCA	GATGTGTGAG	GCGCCTGTGG	TGACCCGAGA	GTGGGTGTTG	5580
GACAGTGTAG	CACTCTACCA	GTGCCAGGAG	CTGGACACCT	ACCTGATACC	CCAGATCCCC	5640

CACAGCCACT ACTGACTGCA GCCAGCCACA GGTACAGAGC CACAGGACCC CAAGAATGAG 5700
 CTTACAAAGT GGCCTTTTCCA GGCCCTGGGA GCTCCTCTCA CTCTTCAGTC CTTCTACTGT 5760
 CCTGGCTACT AAATATTTTA TGTACATCAG CCTGAAAAGG ACTTCTGGCT ATGCAAGGGT 5820
 CCCTTAAAGA TTTTCTGCTT GAAGTCTCCC TTGGAAA 5857

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1863 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(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	
Ile	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	Asp	Arg	Val	His
		435					440					445			
Ser	Lys	Ser	Val	Glu	Ser	Asn	Ile	Glu	Gly	Gln	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	Xaa	Ala	Phe	Val	Ser	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
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
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
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
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	Ser	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 Ala 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
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
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 Xaa		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
Asn Tyr Pro Pro 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
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:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: s754 A

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

CTAGCCTGGG CAACAAACGA

20

- (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: s754 B

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

GCAGGAAGCA GGAATGGAAC

20

- (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: s975 A

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

TAGGAGATGG ATTATTGGTG

20

- (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: s975 B

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

AGGCAACTTT GCAATGAGTG

20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: tdj1474 A

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

CAGAGTGAGA CCTTGTCTCA AA

22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: tdj1474 B

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

TTCTGCAAAC ACCTTAAACT CAG

23

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: tdj1239 A

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

AACCTGGAAG GCAGAGGTTG

20

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: tdj1239 B

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

TCTGTACCTG CTAAGCAGTG G

21

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 111 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

GGKCTTACTC TGTTGTCCCA GCTGGAGTAC AGWGTGCGAT CATGAGGCTT ACTGTTGCTT	60
GACTCCTAGG CTCAAGCGAT CCTATCACCT CAGTCTCCAA GTAGCTGGAC T	111

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

Xaa	Leu	Leu	Cys	Cys	Pro	Ser	Trp	Ser	Thr	Xaa	Cys	Asp	His	Glu	Ala
1				5					10					15	
Tyr	Cys	Cys	Leu	Thr	Pro	Arg	Leu	Lys	Arg	Ser	Tyr	His	Leu	Ser	Leu
			20					25					30		
Gln	Val	Ala	Gly	Leu											
			35												

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1540 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

1 GAGGCTAGAG GGCAGGCACT TTATGGCAAA CTCAGGTAGA ATTCTTCCTC

51 TTCCGTCTCT TTCCTTTTAC GTCATCCGGG GGCAGACTGG GTGGCCAATC
101 CAGAGCCCCG AGAGACGCTT GGCTCTTTCT GTCCCTCCCA TCCTCTGATT
151 GTACCTTGAT TTCGTATTCT GAGAGGCTGC TGCTTAGCGG TAGCCCCCTG
201 GTTTCCGTGG CAACGGAAAA GCGCGGGAAT TACAGATAAA TTAAACTGC
251 GACTGCGCGG CGTGAGCTCG CTGAGACTTC CTGGACGGGG GACAGGCTGT
301 GGGGTTTCTC AGATAACTGG GCCCCTGCGC TCAGGAGGCC TTCACCCTCT
351 GCTCTGGGTA AAGGTAGTAG AGTCCCGGGA AAGGGACAGG GGGCCCAAGT
401 GATGCTCTGG GGTACTGGCG TGGGAGAGTG GATTTCCGAA GCTGACAGAT
451 GGGTATTCTT TGACGGGGGG TAGGGGCGGA ACCTGAGAGG CGTAAGGCGT
501 TGTGAACCCT GGGGAGGGGG GCAGTTTGTA GGTGCGGAGG GAAGCGCTGA
551 GGATCAGGAA GGGGGCACTG AGTGTCCGTG GGGGAATCCT CGTGATAGGA
601 ACTGGAATAT GCCTTGAGGG GGACACTATG TCTTTAAAAA CGTCGGCTGG
651 TCATGAGGTC AGGAGTTCCA GACCAGCCTG ACCAACGTGG TGAAACTCCG
701 TCTCTACTAA AAATACnAAA ATTAGCCGGG CGTGGTGCCG CTCCAGCTAC
751 TCAGGAGGCT GAGGCAGGAG AATCGCTAGA ACCCGGGAGG CGGAGGTTGC
801 AGTGAGCCGA GATCGCGCCA TTGCACTCCA GCCTGGGCGA CAGAGCGAGA
851 CTGTCTCAAA ACAAAACAAA ACAAAACAAA ACAAAAACA CCGGCTGGTA
901 TGTATGAGAG GATGGGACCT TGTGGAAGAA GAGGTGCCAG GAATATGTCT
951 GGGGAAGGGGA GGAGACAGGA TTTTGTGGGA GGGAGAACTT AAGAACTGGA
1001 TCCATTTGCG CCATTGAGAA AGCGCAAGAG GGAAGTAGAG GAGCGTCAGT
1051 AGTAACAGAT GCTGCCGGCA GGGATGTGCT TGAGGAGGAT CCAGAGATGA
1101 GAGCAGGTCA CTGGGAAAGG TTAGGGGCGG GGAGGCCTTG ATTGGTGTTG
1151 GTTTGGTTCG TGTGATTTT GGTTTTATGC AAGAAAAAGA AAACAACCAG
1201 AAACATTGGR GAAAGCTAAG GCTACCACCA CCTACCCGGT CAGTCACTCC
1251 TCTGTAGCTT TCTCTTYCTT GGAGAAAGGA AAAGACCCAA GGGGTTGGCA
1301 GCRATATGTG AAAAAATTCA GAATTTATGT TGTCTAATTA CAAAAAGCAA
1351 CTTCTAGAAT CTTTAAAAAT AAAGGACGTT GTCATTAGTT CTTTGGTTTG
1401 TATYATTCTA AAACCTTCCA AATCTTAAAT TTACTTTATT TTAAATGAT
1451 AAAATGAAGT TGTCATTTTA TAAACCTTTT AAAAAGATAT ATATATATGT
1501 TTTTCTAATG TGTTAnAGTT CATTGGAACA GAAAGAAATG

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21781 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

1 ccttttacgt catcggggag actgggtggc aatcgagcc cgagagacgc atggctcttt
61 ctgccctcca tcctctgatg taccttgatt tcgtattctg agaggetgct gcttagcggt
121 agccccttgg ttccgtggc aacggaaaag cgcgggaatt acagataaat taaaactgcg
181 actgcgcggc gtgAGCTCGC TGAGACTTCC TGGACCCCGC ACCAGGCTGT GGGGTTTCTC
241 AGATAACTGG GCCCCTGCGC TCAGGAGGCC TTCACCCTCT GCTCTGGGTA AAGvvvvvaa
301 aaataaagga cgctgtcgay tagttcttct ggtttgatt attctmaaac cttocaaatc
361 ttcaaattta cttyatttta aaatgatgaa aatgaagttg tcattttata aaccttttaa
421 aaagatatat atatatgttt ttctaagtgt tTAAAGTTCA TTGGAACAGA AAGAAATGGA
481 TTTATCTGCT CTTGCGTTG AAGAAGTACA AAATGTCATT AATGCTATGC AGAAAATCTT
541 AGAGTGTCCT ATCTGgtaag tcagcacaag agtgtattaa tttgggattc ctatgattat
601 ctcttatgca aatgaacaga attgacctta catactaggg aagaaaagac atgtctagta
661 agattaggct attgtaattg ctgattttct taactgaaga actttaaaaa tatagaaaat
721 gattccttgt tctccatcca ctctgcctct cccactcctc tccttttcaa cacatcctgt
781 ggtccgggaa agcagggctc tgtcttgatt ggttctgcac tgggcaggat ctgtagata
841 ctgcatttgc tttctccagc tctaaavvvv vvvvvvvvaa atgctgatga tagtatagag
901 tattgaaggg atcaatataa ttctgttttg atatctgaaa gctcactgaa ggtaaggatc
961 gtattctctg ctgtattctc agttcctgac acagcagaca tttaataaat attgaacgaa
1021 cttgaggcct tatgttgact cagtcataac agctcamagt tgaacttatt cactaagaat
1081 agctttatth ytaataaat tattgagcct catttattht ctttttctcv vvvvTCTGGA
1141 GTTGATCAAG GAACCTGTCT CCACAAAGTG TGACCACATA TTTTGCAAvv vvatgtggc
1201 tccattatta gcttttgtht ttgtccttca taaccagga aacacctaac tttatagaag
1261 ctttacttht ttcaattaag tgagaacgaa aatccaactc catttcattc tttctcagag
1321 agtatatagt tatcaaaagt tgggtgtaat catagttcct ggtaaagtht tgacatatat
1381 tatctthttht tttttttgag acaagtctcg ctctgtcgcc caggctggag tgcagtgga
1441 tgaggcttgc tcaactgcacc tccgcccccg agttcagoga ctctvvvvv vvvvvvtga
1501 gatctagacc acatgggtcaa agagatagaa tgtgagcaat aaatgaacct taaatthttht
1561 aacagctact tttttthttht ttttttgaga cagGGkCTTA CTCTGTTGTC CCAGCTGGAG

1621 TACAGw-G-T GCGATCATGA GGCTTACTGT TGC-TTGAC- TCCTA-GGCT CAAGCGATCC
1681 TATCACCTCA GTCTC-CAAG -TAGCT-GGA CTgtaagtgc acaccaccat atccagctaa
1741 attttgtggt ttctgtagag acgggggttc gccatgtttc ccaggctggt cttgaacttt
1801 gggcttaacc cgtctgccc cctaggcatc ccaaagtgt aggattacag gtgtgagtca
1861 tcatgcctgg ccagtatttt agttagctct gtcttttcaa gtcataata agttcatttt
1921 ctttaagtta gttacaacc ttatatcatg tattcttttc tagcavvvv vvvvvvvtg
1981 tgatcataac agtaagccat atgcatgtaa gttcagtttc atagatcatt gcttatgtag
2041 tttagggttt tgcttatgca gcatccaaa acaattagga aactattgct tgtaattcac
2101 ctgccattac tttttaaatg gctcttaagg gcagttgtga gattatcttt tcatggctat
2161 ttgccttttg agtattcttt ctacaaaagg aagtaaatta aattgttctt tctttcttta
2221 taatttatag atvvvvvATT TTGCATGCTG AAACCTTCTCA ACCAGAAGAA AGGGCCTTCA
2281 CAGTGTCTT TATGTAAGAA TGATATAACC AAAAGgatat aannngaant gatgctaggt
2341 tggaagcaac cacagtagga aaagtagaaa ttatttaata acatagcgtt cctataaaac
2401 cattcatcag aaaaatttat aaaagagttt ttagcacaca gtaaattatt tccaaagtta
2461 ttttctgaa agttttatgg gcatctgctt tatacaggtt ttgvvvvvv vvvvvvggta
2521 ggcttaaatg aatgacaaaa agttactaaa tcaactgcat cacacgggta tatacagatg
2581 tcaatgatgt attgatttag aggttttcta ctgttgctgc atcttatttt tatttgttta
2641 catgtctttt cttatttttag tgccttaaaa aggttgataa tcaactgctg agtgtgtttc
2701 tcaaavvvv GAGCCTACAA GAAAGTACGA GATTTAGTCA ACTTGTTGAA GAGCTATTGA
2761 AAATCATTTG TGCTTTTCAG CTTGACACAG GTTTGGAGTA TGCAAACAGC TATAATTTTG
2821 CAAAAAGGA AAATAACTCT CCTGAACATC TAAAAGATGA AGTTTCTATC ATCCAAAGTA
2881 TGGGCTACAG AAACCGTGCC AAAAGACTTC TACAGAGTGA ACCCGAAAAT CCTTCCTTGC
2941 AGaaaccatt tgttttcttc ttcttctctt tcttcttttc tttttttttt cttttttttt
3001 tcwtgagagt ggagtcyttg gctcgtssyg gcgcgaggo taggaagcag ctctctctgc
3061 gctcgtacgc cacgctagta ggtagcgtgg gcakacacgg cacgcgcacc atggccaggo
3121 taatttttgt attgttvvv vvvvvvvvva tgatggagat cttaaaaagt aatcattctg
3181 gggctgggog tagtagcttg cacctgtaat cccagcactt cgggaggctg aggcaggcag
3241 ataatttgag gtcaggagtt tgagaccagc ctggccaaca tggtgaaacc catctctact
3301 aaaaatacaa aaattagctg ggtgtggtgg cacgtacctg taatccagc tactcgggag
3361 gcggaggcac aagaattgct tgaacctagg acgcggaggt tgcagcgagc caagatogcg
3421 ccaactgcact ccagcctggg ccgtagagtg agactctgtc tcaaaaaaga aaaaaagta
3481 attgttctag ctgggcgcag tggtcttgc ctgtaatcc agcactttgg gaggccaagg
3541 cgggtggatc tcgagtcta gagttcaaga ccagcctagg caatgtggtg aaaccccatc
3601 gctacaaaa atacaaaat tagccaggca tggtggcgtg cgcagttagt cccagctct

U6A10000

3661 tgggaggctg aggtgggagg atcacttgaa cccaggagac agaggttgca gtgaaccgag
3721 atcacgccac cacgctccag cctgggcaac agaacaagac tctgtctaaa aaaatacaaaa
3781 taaaataaaa gtagttctca cagtaccagc attcattttt caaaagatat agagctaaaa
3841 aggaaggaaa aaaaaagtaa tgttgggctt ttaaatactc gttcctatac taaatgttct
3901 taggagtgct ggggttttat tgtcatcatt tacccttttt aaaaatgtta ttggccaggc
3961 acggtggctc atggctgtaa tcccagcact ttgggaggcc gaggcaggca gatcacctga
4021 ggtcaggagt gtgagaccag cctggccaac atggcgaaac ctgtctctac taaaaataca
4081 aaaattaact aggcgtgggtg gtgtacgcct gtagtcccaG ctactcggga ggctgaggca
4141 ggagaatcaa ctgaaccagg gaggtggagg ttgcagtgtg ccgagatcac gccactgcac
4201 tctagcctgg caacagagca agattctgtc tcaaaaaaaaa aaaacatata tacacatata
4261 tcccaaagtg ctgggattac atatatatat atatatatat atTatatata tatatatata
4321 tatatgtgat atatatgtga tatatatata acatatatat atgtAatata tatgtgatat
4381 atatatataa tatatatgtA atatatatgt gatatatata tatacacaca cacacacata
4441 tatatgtAtg tgttgtgtAca Cacacacaca caaattagcc aggcatagtt gcacacgctt
4501 ggtagaccCa gctactcagg aggCtgaggG aggagaATct cttgaActta ggaggcggag
4561 gttgcagtga gctgagattg cggcactgca ctccagcctg ggtgacagag caggactctg
4621 tacaccccc aaacaaaaa aaaaagttat cagatgtgat tggaatgtat atcaagtatc
4681 agcttcaaaa tatgtctatat taatacttca aaaattacac aaataatata taatcagggt
4741 tgaaaaattt aagacaacms aaaaaaaaw ycmaatcaca matatccac acattttatt
4801 attmctmctm cwattatttt gwagagmctg ggtctcacyc ykttgctwat gctgggtctt
4861 gaacyccykg ccycarcar tctstctcca bcctcccaar gtgctgggga twataggcat
4921 garctaccg ctcccagccc magacaattt tagtgtgwgc aaaattcctg ggcatttttt
4981 cmaggchdnm wayayvtwds mtgcwhrtga hbggtgatg wgtcaattta ttttgyccat
5041 ggtgtcaakt ytctcttcag gaggaaaagc acamgaactg gccayaatt gcttgactgt
5101 tctttaccaw actgttwarc agGAAACCAr TCTCAGTGTC CCAACTCTCT AACCTTGGAA
5161 CTTGTGAGAA CTCTGAGGAC AAAGCAGCGG ATACAACCTC AAAAGACGTC TGTCTACATT
5221 GAATTGggta agggctctcag gttttttaag tatttaataa taattgctgg attccttato
5281 ttatagtttt gccamaaato ttggtcataa tttgtatttg tggtaggcag cwttgggag
5341 tgaattttat gagccctatg gtgagttata aaaaatgtaa aagacgcagt tcccaccttg
5401 aagaatctta ctttaaaaag ggagcaaaag aggccaggca tgggtggctca cacctgtaat
5461 cccagcactt tgggaggcca aagtgggtgg atcacctgag gtcgggagtt cgagaccagc
5521 ctagccaaca tggagaaact ctgtctgtac caaaaaataa aaaattagcc aggtgtgggtg
5581 gcacataact gtaatccag ctactcggga ggctgaggca ggagaatcac ttgaaCccgg
5641 gaggtggagg ttgcggtgaa ccgagatcgc accattgcac tccagcctgg gcaaaaaatG

5701 cgaaactcca tctaaaaaaaa aaaaagagag caaaagaaag amtmtctggt tttamtmtg
5761 tgtaaataatg ttttTggaaa gatggagagt agcaataaga aaaaacatga tggattgcta
5821 cagtatttag ttccaagata aattgtacta gatgaggaag ccttttaaga agagctgaat
5881 tgccaggcgc agtgggtcac gctgtaatc ccagcacttt gggaggccga ggtgggcgga
5941 tcacctgagg tcgggagttc aagaaccagcc tgaccaacat ggagaaaccc catctctact
6001 aaaaaaaaaa aaaaaaaaaat tagccgggggt ggtggcttat gctgtaatc ccagctactc
6061 aggaggctga ggcaggagaa tcgcttgaac ccaggaagca gaggttgagc tgagccaaga
6121 tcgcaccatt gcactccagc ctaggcaaca agagtgaac tccatctcaa aaaaaaaaaa
6181 aAAgagctga Atcttggctg ggcaggatgg ctctgctg taatcctaac gctttggaag
6241 accgaggcag aaggattggt tgagtccacg agtttaagac cagcctggcc aacatagggg
6301 aacctgtct ctatttttaa aataataata ctttttggc cgggtgcggtg gctcatgct
6361 gtaatcccaa tactttggga ggctgaggca ggtagatcac ctgaggtcag agttcgagac
6421 cagcctggat aacctgggtga aacccctctt tactaaaaat acaaaaaaaaa aaaaaatta
6481 gctgggtgtg gtagcacatg cttgtaatcc cagctacttg ggaggctgag gcaggagaat
6541 cgcttgaacc agggaggcgg aggttacaat gagccaacac tacaccactg cactccagcc
6601 tgggcaatag agtgagactg catctcaaaa aaataataat ttttaaaaat aataaatttt
6661 ttttagctta taaaaagaaa agttgaggcc agcatagtag ctacatctg taatctcagc
6721 agtggcagag gattgcttga agccaggagt ttgagaccag cctgggcaac atagcaagac
6781 ctcatctcta caaaaaaatt tcttttttaa attagctggg tgtggtggtg tgcatctgta
6841 gtcccagcta ctcaggaggc agaggtgagt ggatacattg aaccaggag tttgaggctg
6901 tagtgagcta tgatcatgcc actgcactcc aacctgggtg acagagcaag acctccaaaa
6961 aaaaaaaaaa aagagctgct gagctcagaa ttcaaactgg gctctcaaat tggattttct
7021 tttagaatat atttataatt aaaaaggata gccatctttt gagctccag gcaccaccat
7081 ctatttatca taacacttac tgttttcccc ccttatgatc ataaattcct agacaacagg
7141 cattgtaaaa atagtatat tagttgatat ttaggagcac ttaactatat tccaggcact
7201 attgtgcttt tcttgataa ctcatagat gcttgtcaga cctctgagat tgttcctatt
7261 atacttattt tacagatgag aaaattaagg cacagagaag ttatgaaatt tttccaagg
7321 attaaacctt gtaagtggct gagccatgat tcaaacctag gaagttagat gtcagagcct
7381 gtgctttttt tttgttttg tttttgttt cagTagaaac gggggtctca ctttgttggc
7441 caggctggtc ttgaactcct aacctcaaat aatccacca tctcggcctc ctcaagtgt
7501 gggattacag gtgagagcca ctgtgcctgg cgaagcccat gcctttaacc acttctctgt
7561 attacatact agcttaacta gcattgtacc tgccacagta gatgtcagt aaatatttct
7621 agttgaatat ctgtttttca acaagtacat tttttaacc cttttaatta agaaaacttt
7681 tattgattta tttttgggg ggaaattttt taGGATCTGA TTCTTCTGAA GATACCGTTA

GeneDX 1021, pg. 120

9781 ctgtagatag aacagagtat taagtagact taatggccag taatcttttag agtacatcag
9841 aaccagtttt ctgatggcca atctgctttt aattcactct tagacgttag agaaataggt
9901 gtggtttctg catagggaaa attctgaaat taavvvvvvv vvvvvvgatc ctaagtggaa
9961 ataatctagg taaataggaa ttaaatgaaa gagtatgagc tacatottca gtatacttgg
10021 tagtttatga ggttagtttc tctaatatag ccagttggTt gatttccacc tccaaggtgt
10081 atgaagtatg ttttttttta atgacaattc agtttktgag taccttggtta tttttgtata
10141 ttttcaGCTG CTTGTGAATT TTCTGAGACG GATGTAACAA ATACTGAACA TCATCAACCC
10201 AGTAATAA-T GATTTGAACA CCACTGAGAA GCGTGCAGCT GAGAGGCATC CAGAAAAGTA
10261 TCAGGGTAGT TCTGTTTCAA ACTTGCATGT GGAGCCATGT GGCACAAATT ACTCATGCCA
10321 GCTCATTACA GCATGAGAAC AGCAGTTTAT TACTCACTAA AGACAGAATG AATGTAGAAA
10381 AGGCTGAATT CTGTAATAAA AGCAAACAGC CTGGCTTAGC AAGGAGCCAA CATAACAGAT
10441 GGGCTGGAAG TAAGGAAACA TGTAATGATA GGCGGACTCC CAGCACAGAA AAAAAGGTAG
10501 ATCTGAATGC TGATCCCCTG TGTGAGAGAA AAGAATGGAA TAAGCAGAAA CTGCCATGCT
10561 CAGAGAATCC TAGAGATACT GAAGATGTTT CTTGGATAAC ACTAAATAGC AGCATTGAGA
10621 AAGTTAATGA GTGGTTTTCC AGAAGTGATG AACTGTTAGG TTCTGATGAC TCACATGATG
10681 GGGAGTCTGA ATCAAATGCC AAAGTAGCTG ATGTATTGGA CGTCCCTAATT GAGGTAGATG
10741 AATATTCTGG TTCTTCAGAG AAAATAGACT TACTGGCCAG TGATCCTCAT GAGGCTTTAA
10801 TATGTAAAAG TGACAGAGTT CACTCCAAAT CAGTAGAGAG TAATATTGAA GGCCAAA-TA
10861 TTTGGGAAAA CCTATCGGAA GAAGGCAAGC CTCCCCAACT TAAGCCATGT AACTGAAAAAT
10921 CTAATTATAG GnGCATTTGT TTCTGAGCCA CAGATAATAC AAGAGCGTCC CCTCACAAAT
10981 AAATTAAAGC GTAAAAGGAG ACCTACATCA GGCCTTCATC CTGAGGATTT TATCAAGAAA
11041 GCAGATTTGG CAGTTCAAAA GACTCCTGAA ATGATAAATC AGGGAACATA CCAAACGGAG
11101 CAGAAtggtc aagtgaTGAA TATTACTAAT AGTGG-TCA- TGAGAATAAAA ACAAAGGTG
11161 ATTCTATTCA GAATGAGAAA AATCCTAACC CAATAGAATC ACTCGAAAAA GAATCTGCTT
11221 TCAAAACGAA AGCTGAACCT ATAAGCAGCA GTATAAGCAA TATGGAACTC GAATTAAATA
11281 TCCACAATTC AAAAGCACCT AAAAAGAATA GGCTGAGGAG GAAGTCTTCT ACCAGGCATA
11341 TTCATGCGCT TGAAC TAGTA GTCAGTAGAA ATCTAAGCCC ACCTAATTGT ACTGAATTGC
11401 AAATTGATAG TTGTTCTAGC AGTGAAGAGA TAAAGAAAAA AAAGTACAAC CAAATGCCAG
11461 TCAGGCACAG CAGAAACCTA CAACTCATGG AAGGTAAAGA ACCTGCAACT GGAGCCAAGA
11521 AGAGTAACAA GCCAAATGAA CAGACAAGTA AAAGACATGA CAGTGATACT TTCCAGAGC
11581 TGAAGTTAAC AAATGCACCT GGTTC'TTTTA CTAAGTGTTT AAATACCAGT GAACTTAAAG
11641 AATTTGTCAA TCCTAGCCTT CCAAGAGAAG AAAAAGAAGA GAACTAGAA ACAGTTAAAG
11701 TGTCTAATAA TGCTGAAGAC CCCAAAGATC TCATGTTAAG TGGAGAAAGG GTTTTGCAAA
11761 CTGAAAGATC TGTAGAGAGT AGCAGTATTT CACTGGTACC TGGTACTGAT TATGGCACTC

11821 AGGAAAGTAT CTCGTTACTG GAAGTTAGCA CTCTAGGGAA GGCAAAAACA GAACCAAATA
11881 AATGTGTGAG TCAGTGTGCA GCATTTGAAA ACCCCAAGGG ACTAATTCAT GGTGTGTTCCA
11941 AAGATAATAG AAATGACACA GAAGGCTTTA AGTATCCATT GGGACATGAA GTTAACCACA
12001 GTCGGGAAAC AAGCATAGAA ATGGAAGAAA GTGAACTTGA TGCTCAGTAT TTGCAGAATA
12061 CATTCAAGGT TTCAAAGCGC CAGTCATTTG CTCTGTTTTT AAATCCAGGA AATGCAGAAG
12121 AGGAATGTGC AACATTCTCT GCCCACTCTG GGTCTTAAA GAAACAAAGT CCAAAAGTCA
12181 CTTTTGAATG TGAACAAAAG GAAGAAAATC AAGGAAAGAA TGAGTCTAAT ATCAAGCCTG
12241 TACAGACAGT TAATATCACT GCAGGCTTTC CTGTGGTTGG TCAGAAAGAT AAGCCAGTTG
12301 ATAATGCCAA ATGTAGTATC AAAGGAGGCT CTAGGTTTTG TCTATCATCT CAGTTCAGAG
12361 GCAACGAAAC TGGACTCATT ACTCCAAATA AACATGGACT TTTACAAAAC CCATATCGTA
12421 TACCACCACT TTTTCCCATC AAGTCATTTG TTAAACTAA ATGTAAGAAA AATCTGCTAG
12481 AGGAAAACCT TGAGGAACAT TCAATGTCAC CTGAAAGAGA AATGGGAAAT GAGAACATTC
12541 CAAGTACAGT GAGCACAATT AGCCGTAATA ACATTAGAGA AAATGTTTTT AAAGAAGCCA
12601 GCTCAAGCAA TATTAATGAA GTAGGTTCCA GTACTAATGA AGTGGGCTCC AGTATTAATG
12661 AAATAGGTTT CAGTGATGAA AACATTCAAG CAGAACTAGG TAGAAACAGA GGGCCAAAAT
12721 TGAATGCTAT GCTTAGATTA GGGGTTTTGC AACCTGAGGT CTATAAACAA AGTCTTCTCTG
12781 GAAGTAATTG TAAGCATCCT GAAATAAAAA AGCAAGAATA TGAAGAAGTA GTTCAGACTG
12841 TTAATACAGA TTTCTCTCCA TCTCTGATTT CAGATAACTT AGAACAGCCT ATGGGAAGTA
12901 GTCATGCATC TCAGGTTTGT TCTGAGACAC CTGATGACCT GTTAGATGAT GGTGAAATAA
12961 AGGAAGATAC TAGTTTTGCT GAAATGACA TTAAGGAAAG TTCTGCTGTT TTTAGCAAAA
13021 GCGTCCAGAG GGGAGAGCTT AGCAGGnGTC CTAGCCCTTT CACCCATACA CATTTGGCTC
13081 AGGGTTACCG AAGAGGGGCC AAGAAATTAG AGTCCTCAGA AGAGAACTTA TCTAGTGAGG
13141 ATGAAGAGCT TCCCTGCTTC CAACACTTGT TATTGGTAA AGTAAACAAT ATACCTTCTC
13201 AGTCTACTAG GCATAGCACC GTTGCTACCG AGTGTCTGTC TAAGAACACA GAGGAGAATT
13261 TATTATCATT GAAGAATAGC TTAAATGACT GCAGTAACCA GGTAATATTG GCAAAGGCAT
13321 CTCAGGAACA TCACCTTAGT GAGGAAACAA AATGTTCTGC TAGCTTGTTT TCTTCACAGT
13381 GCAGTGAATT GGAAGACTTG ACTGCAAATA CAAACACCCA GGCTCCTTTC TTGATTGGTT
13441 CTTCCAAACA AATGAGGCAT CAGTCTGAAA GCCAGGGAGT TGGTCTGAGT GACAAGGAAT
13501 TGGTTTCAGA TGATGAAGAA AGAGGAACGG GCTTGGAAGA AAATAATCAA GAAGAGCAAA
13561 GCATGGATTC AAACCTTAGgt attggaacca ggttt--gtg tgtgtggccc agtctattta
13621 tagaagtgag ctaaattgttt atgcttttgg ggagcacatt ttacaaattt ccaagtatag
13681 ttaaaggaac tgcttcttaa acttganaca tgttctctct aaggtggttn ncatagaaaa
13741 aagtccttca cacagctagg acgtcatott tggcttgaat gagctttaac atcttantta
13801 ctggtggact tacttctggt ttcattttat taaagaaatc caggtgtccc aagcaaggat

13861 ttatcactgg gtgaaagaaa gtaatccagc cccgcantga gaggaaagca nagaagtgga
13921 gngnttgget tgttnactg acccnctttg gnnttaggga gagactgtcc vvvvvvvvvv
13981 vvvvggtgtcc caaagcaagg aatttaatca ttttgtgtga catgaaagta natccagtcc
14041 tgccaatgag aagaaaaagc acagcaagnn ngcagcgttt atagtctgct tttacatctg
14101 aacctctggt tttgttattt aagGT-AAGG TGAAGCAGCA TCTGGGTGTG AGAGTGAAAC
14161 AAGCGTCTCT -GAAGACTGC TCAGGGCTAT CCTCTCAGAG TGACATTTTA -ACCACTCAG
14221 CAGAGvvvvv vvvvvvvvng ngaatgtaat cctaataatt cncncna-c tta-aaagaa
14281 -tacc-actc caanggcato ncaatacatc aatcaattgg gg-aattggg atttt--ccc
14341 tcnctaacat ca-ntggaat aatttcatgg c-attaattg catgaatgtg gttagattaa
14401 aaggtgttca tgcta-gaac ttgtagtctc atacta-ggt -atttcaat- cnttgttcoo
14461 caaan-aatt tgtatgatat attntcattt aatggaaa-g cttotcaaag tatttc-att
14521 tt-cttggtta ccatnnatcg tttttgaatn ag-GG-AT-A CCATGCAACA TAACCTGATA
14581 AAGCTCCAGC AGG-AAATGG CTGAACTAGA AGCTGTGTTA GAACAGCATG GGAGCCAGCC
14641 TTCTAACAGC TACCCTTCCA TCATAAGTGA CTCTTCTGCC CTTGAGGACC TGCAGAAATCC
14701 AGAACA-A-A GCACATCAGA -AAAAGGtgt gtattgttgg ccaaactg atatottaan
14761 caaaattctt tcttccctt ttatctcctt ctgaagagta aggaoctagc tocaacattt
14821 tatgatcctt gctcagcaca tgggtaatta tggagccttg gttcttgtcc ctgctcacia
14881 ctaatatacc agtcagaggg acccaaggca gtcattcatg ttgtcatctg agavvvvvvv
14941 vvvvvvaat tctgectgat atacttgtn caaaaccaat ttgtgtatca tagattgatg
15001 cttttgaana aaatcagtat tctaacctga attatcacta tcagaacaaa gcagtaaagt
15061 agatttggtt tctcattcca tttaaagcag TATTAAC TTC ACAGA-AAAG TAGTGAATAC
15121 CCTATAAGCC AGAATCCAGA AGGCCtntCT GCTGACAAGT TTGAGGTGTC TGCAGATAGT
15181 TCTACCAGTA AAAATAAAGA ACCAGGAGTG GAAgtaagaa acatcaatgt aaagatgctg
15241 tggatatctga catctttatt tatattgaac tetgattgtt aatttttttc accatacttt
15301 ctccagtttt tttgcataca ggcatttata cacttttatt gctctaggat acttcttttg
15361 tttaatccta tataggvvvv vvvvvvvvg gataagntca agagatatat tgatagggtga
15421 tgcagtgatn aattngaaa atttntgccc tgcttttaac ctccccctg tctttcttcc
15481 tncctccctc ccttctncc tcccgctctt ncctttctt tccctccctt cnccttctt
15541 tccntctntc tttctttct tctctgtcta cctttcttct cttctccct tcttttctt
15601 ttctttcttt cctttcttt tctttcttt ctttcttct cttctttct tgacagagtc
15661 ttgctctgtc actcaggtg gagtgcagt gctgatctc gntc-actg caacctct--
15721 gtc-tc-cca ggt-tcaagc aattttctg cctcagctc ccgagtagct gagatta-ca
15781 ggccgagcc -ac-cac-ac cc-ag-ctac tgacctgctt ttvvvaaca gctgggagat
15841 atggtgctc agaccatccc catgttatat gtaaacctg acatattggc aggcaacatg

ttatcactgg

15901 aatccagact totaggetgt catgegggt ctttctttgc cagtcattnc tgatctctct
 15961 gacatgaact gtntcaggta tgctttggct gcccagcaag tatgaaatgt cctntcacia
 16021 ggggtggcga tggttttctc ctt-ccattt atctttctAG GTCATCCCCCT TCTAAATGCC
 16081 CATCA-TTA- GATGATAGGT GGTACATG-C ACAGTTGCTC TGGGAGTCTT CAGAA-TAGA
 16141 AACTACCCA- CCTCAAGAGG AGCT-CATTA AGGTTGTTGA TGTGGAGGAG CAACAGCTGG
 16201 AAGAGTCTGG GCCACACGAT TTGACGGAAA CATCTTACTT GCCAAGGCAA GATCTAGgta
 16261 atatttcacg tgctgtattg garcamacac tytgatttta ctctgaatcc tacataaaga
 16321 tatt-ctggg tarccaactt -ttagat-gt rctagtctat -catggacac ttttgt-tat
 16381 acttaattaa gccacttta gaaaaatagc tcaagtgtta atcwaggttt acttgwaaat
 16441 tattgaaact gttaatccat ctatatttta attaatgggt taactaatga ttttgaggat
 16501 gwgggagtok tgggtgtactc tamatgtatt atttcaggcc aggcatagtg gctcacgct
 16561 ggtaatccca gtayyemrga gcccaggcca ggtggagcca gctgaggtca ggagttcaag
 16621 acctgtcttg gccaacatgg gngaaaccct gtcttcttct taaaaaanac aaaaaaatt
 16681 aactgggttg tgcttaggtg natgccccgn atcctagttn ttcttgnggg ttgagggagg
 16741 agatcacntt ggacccccga gggngnggtg gggngagca ggncaaaaca cngaccagc
 16801 tgggggtgaa ggggaagccca ctcaaaaaa nnttnvvvv vvvvvvvvAG GGAACCCCTT
 16861 ACCTGGAATC TGGAATCAGC CTCTT-CTCT GATGACCC-T GAATCTGATC C-TTCTGAAG
 16921 ACA-GAGCCC C-A-GAGTCA -GCTCGTGT GGC-AACA-T ACCATCTTC- AACCTCTGCA
 16981 TTGAAAGTT- CCCCATTGA AA-GTTGCAG AATCTGCCCA GAGTCCAGCT GCTGCTCA-T
 17041 ACTACTGATA CTGCTGGGTA TAATGCAATG GAAGAAAGTG TGAGCAGGGA GAAGCCAGAA
 17101 TTGACAGCTT CAACAGAAAG GGTCAACAAA AGAATGTCCA TGGTGG-TGT CTGGCCTGAC
 17161 CCCAGAAGAA TTTgtgagtg tatccatatg tatctcccta atgactaaga cttaacaaca
 17221 ttctggaaag agttttatgt aggtattgtc aattaataac ctagaggaag aaatctagaa
 17281 ancaatcaca gttctgtgta atttaatttc gattactaat ttctgaaaat ttagaayvvv
 17341 vvvvvvvvvv ncccncccc cnaatctgaa atgggggtaa cccccccca accganacnt
 17401 gggtn-gcnt agaga-nttt -aa-tggccc nt-tctgagg -nacanaagc ttaagc-cag
 17461 gng-acgtgg ancnat-gng ttgtttnttg tt-tggttac ctccagcctg gtgacaganc
 17521 aagctctgtc tngnaaaant agngaggggn naagattcac nntaaatagt tccaggacac
 17581 gtgtagaacg tgcaggattg ctacgtaggt aaacatatgc catggtggga taactagtat
 17641 tctgagctgt gtgctagagg taactcatga taatggaata tttgatttaa tttcagATGC
 17701 TCGTGTACAA GTTTGCCAGA AAACACCACA TCACTTTAAC TAATCTAATT ACTGAAGAGA
 17761 CTAATCATGT TGTATGAAA ACAGatatac ccaagaacct ttacagaata ccttkcatct
 17821 sctgcataaa accacatgag gcgaggcaca gtggcgcatg cctgtawtcg caRcactttg
 17881 ggaggccgag gcgggcagat accccgagRt tMSgagtcga gccatcctgg ccagcatggt

17941 gaaacccccgt ctctactaNN aaatggNaaa attaNctggg tgtggtcgcg tgcNcctgta
18001 gtcccagcta ctcgtgaggc tgaggcagga gaatcacttg aaccggggaa atggaggttt
18061 cagtgcagcag agatcatNcc cctNcattcc agcctggcga cagagcaagg ctccgtcNcc
18121 Naaaaaataa aaaaaaacgt gaacaaataa gaatatttgt tgagcatagc atggatgata
18181 gtcttctaata agtcaatcaa ttactttatg aaagacaaat aatagttttg ctgcttcctt
18241 acctcctttt gttttgggtt aagatttggg gtgtggggcca ggcacvvvvv vvvvvvvvga
18301 tctat-agct agccttggCg tctagaagat ggggtgttgag aagagggagt ggaaagatat
18361 ttctcttggt cttaacttca tatcagcctc ccctagactt ccaaatatcc atacctgctg
18421 gttataatta gtgggtgtttt cagcctctga ttctgtcacc aggggtttta gaatcataaa
18481 tccagattga tcttgggagt gtaaaaaact gaggtctttt agcttcttag gacAgcactt
18541 cctgattttg ttttcaactt ctaatccttt gagtgTktct cattctgcag aTGCTGAGTT
18601 TGTGTGTGAA CGGACACTGA AATATTTTCT AGGAATTGCG GGAGGAAAAT GGGTAGTTAG
18661 CTATTTCTGG kaagkataat actatwtctc ccycctccc CtAacamTc agaattgsat
18721 tttacaccta acatttacam ctaAgnyyyt kgctgatgct tgagtctgag tcmcmmaaagk
18781 cctttaatTg Taatactaac tactTttatc ttTaatatca cttTgTtcag AtaagCtggg
18841 gaTGCTggga aaatgggtctc ttttataact aataggacct aAtctgctcc tagcaaTgtt
18901 agca-tatga gctagggatt tatttaatag tcggcaggaa tccatgtgca rcagncaaac
18961 ttataatgtt taaattaaac atcaactctg tctccagaag gaaactgctg ctacaagcct
19021 tattaaaggg ctgtggcttt agagggaagg acctctctc tgtcattctt cctgtgctct
19081 tttgtgaatc gotgacctct ctatctccgt gaaaagagca cgctctctg ctgtatgtaa
19141 cctgtctttt ctatgatctc tvvvvvvvvv vvvvnaaaaa cggggnnngg antgggcctt
19201 aanccaaag ggcnaactcc ccaaccattn aaaaantgac nggggattat taaaancggc
19261 gggaaacatt tcaongccca actaatattg ttaaattaa accaccacn ctgcnccaag
19321 gagggacctt ctccnncaaa ncttattaa gggcctgngc cttaaagagg gaaggacctt
19381 ctctctgtc aatcttctg tgotctnttt gtgaatcgct gacctctcta tctccgtgaa
19441 aaagagcacg ttcttctgct gtatgtaacc tgtcttttct atgatctctt tagggGTGAC
19501 CCAGTCTATT AAAGAAAGAA AAATGCTGAA TGAGgtaagt acttgatgtt acaaactaac
19561 cagagatatt cattcagtca tatagttaaa aatgtatttg ctctcttcca tcaatgcacc
19621 actttcctta acaatgcaca aattttccat gataatgagg atcatcaaga attatgcagg
19681 cctgcactgt ggctcatacc tataatocca gcgctttggg aggctgaggc gcttggatcv
19741 vvvvvvvvvv vvtgcacccc aagacggatg ggnnctcccc cncnctatt ttgatttaga
19801 aantgggttc ccatgngtaa gngtgcgac cctgcctca atgactnctc ctaatctccc
19861 aaattctagg atagngggtg anccc-ccn c-tgnc-tga -t-ccntaaa atg-gacgtg
19921 tctgctncnc cact-ccatt gaagga-gct tctctttctc ttatctgat gggt-gtgn

19981 -gtttctttc agCATGATTT TGAAGTCAGA GGAGATGTGG TCAATGGAAG AAACCACCAA
20041 GGTCCAAAGC GAGCAAGAGA ATCCCAGGAC AGAAAGgtaa agctccctcc ctcaagttga
20101 cnnaaatctc accccaccac tctgtattcc actcccttt gcagagatgg gccgcttcat
20161 tttgtaagac ttattacata catacacagt gctagatact ttcacacagg ttcttttttc
20221 actcttccat cccaaccaca taaataagta ttgtctctac tttatgaatg ataaaactaa
20281 gagattttaga gaggtgtgt aatttggatt ccggtctcgg gttcagatcv vvvvvvvvvv
20341 vvATCTTCAG GGGGCTAGAA ATCTGTTGCT ATGGGCCCTT CACCAACATG CCCACAGgta
20401 agagcctcgg gagaancag agttccagca ccagccttg tcttacatag tggagtatta
20461 taagcaaggt cccaogatgg gggttctca gattgctgaa atgttctaga ggtatttcta
20521 tttctctacc actctccaaa caaacagca cctaatggt atctatggc aaaaaaaaaac
20581 tataccttgt cccctctctc aagagcatga aggtggttaa tagttaggat tcagtatggt
20641 atgtgttcag atggcggtga gctgctgtta gtgccvvvvv vvvvvvvvtt tgagagacta
20701 tcaaacctta taccaagtgg ccttatggag actgataacc agagtacatg gcatatcagt
20761 ggcaaattga cttaaaatcc ataccctac tattttaaga ccattgtcct ttggagcaga
20821 gagacagact ctccattga gaggtottgc tataagcctt catccggaga gtgtagggtta
20881 gagggcctgg gttaagtatg cagattactg cagtgtttt aGatcctaaa tgtccatttt
20941 agATCAACTG GAATGGATGG TACAGCTGTG TGGTGCTTCT GTGGTGAAGG AGCTTTCATC
21001 ATTCACCCTT GGCACAGvvv vvvvvvvvvv cagagcaaga ccctgtctca aaaacaaaca
21061 aaaaaaatga tgaagtgaca gttccagtag tctactttg acactttgaa tgctctttcc
21121 ttcttgggga nccagGTGTC CACCCAATTG TGGTTGTGCA GCCAGATGCC TGGACAGAGG
21181 ACAATGGCTT CCATGtttat cactcattac ctgggtgctt agtagcacag ttcttggcac
21241 attttttaaat atttgttgaa tgaatggcta aaatgtcttt ttgatgtttt tattgttatt
21301 tgttttatat tgtaaaagta atacatgaac tgtttccatg ggggtgggagt aagatatgaa
21361 tgttcatcac vvvvvvvvvv vvvvcaat cctnagaact catacgaccg ggcccttggga
21421 gtcgnt-gnt tngagcctag tccngg-ag- aatgaattga cactaatctc tgcttgtggt
21481 ctctgtctcc agCAATTGGG CAGATGTGTG AGGCGCCTGT GGTGACCCGA GAGTGGGTGT
21541 TGGACAGTGT AGCACTCTAC CAGTGCCAGG AGCTGGACAC CTACCTGATA CCCCAGATCC
21601 CCCACAGCCA CTACTGACTG CAGCCAGCCA CAGGTACAGA GCCACAGGAC CCCAAGAATG
21661 AGCTTACAAA GTGGCCTTTC CAGGCCCTGG GAGCTCCTCT CACTCTTCAG TCCTTCTACT
21721 GTCCTGGCTA CTAAATATTT TATGTACATC AGCCTGAAAA GGAATTCTGG CTATGCAAGG
21781 GTCCCTTAAA GATTTTCTGC TTGAAGTCTC CCTTGGAAAT

WHAT IS CLAIMED IS:

1. An isolated DNA consisting essentially of DNA coding for a BRCA1 polypeptide, said DNA selected from the group consisting of cDNA and genomic DNA.
2. The isolated DNA of claim 1, wherein said BRCA1 polypeptide comprises the amino acid sequence set forth in SEQ ID No:2.
3. An isolated DNA consisting essentially of DNA having at least 15 nucleotides of the DNA of claim 1.
4. The isolated DNA of claim 3 which comprises regulatory sequences.
5. A pair of single-stranded DNA primers for determination of a nucleotide sequence of a BRCA1 gene by a polymerase chain reaction, the sequence of said primers being derived from human chromosome 17q, wherein the use of said primers in a polymerase chain reaction results in the synthesis of DNA having all or part of the sequence of the BRCA1 gene.
6. The pair of primers of claim 5 wherein said sequence of the BRCA1 gene is set forth in SEQ ID No:1.
7. An isolated DNA sequence consisting essentially of DNA coding for a mutant BRCA1 polypeptide, said DNA selected from the group consisting of cDNA and genomic DNA.
8. The isolated DNA of claim 7 selected from the group consisting of:
 - (a) SEQ ID NO:1 having T at nucleotide position 4000;
 - (b) SEQ ID NO:1 having extra C at nucleotide position 5329;
 - (c) SEQ ID NO:1 having G at nucleotide position 5387; and,
 - (d) SEQ ID NO:1 having 11 base pairs at nucleotide positions 133-143 deleted.
9. A nucleic acid probe complementary to human wild-type BRCA1 gene sequences.

10. A nucleic acid probe complementary to human altered BRCA1 gene sequences.
11. The nucleic acid probe which hybridizes to part of the altered BRCA1 gene, said part of the BRCA1 gene selected from the group consisting of:
 - (a) SEQ ID NO:1 having T at nucleotide position 4000;
 - (b) SEQ ID NO:1 having extra C at nucleotide position 5329;
 - (c) SEQ ID NO:1 having G at nucleotide position 5387; and
 - (d) SEQ ID NO:1 having 11 base pairs at nucleotide positions 133-143 deleted.
12. A replicative cloning vector which comprises the isolated DNA of claim 1 and a replicon operative in a host cell.
13. An expression system which comprises the isolated DNA of claim 1 operably linked to suitable control sequences.
14. Recombinant host cells transformed with the expression system of claim 13.
15. A method of producing recombinant BRCA1 polypeptide which comprises culturing the cells of claim 14 under conditions effective for the production of said BRCA1 polypeptide.
16. A preparation of human BRCA1 polypeptide substantially free of other human proteins, the amino acid sequence of said protein corresponding to that shown in SEQ ID NO:2.
17. A preparation of human BRCA1 polypeptide substantially free of other human proteins, the amino acid sequence of said polypeptide having substantial homology with the amino acid sequence set forth in SEQ ID NO:1, said polypeptide having substantially similar function as wild-type BRCA1 polypeptide.
18. An antibody immunoreactive with a human BRCA1 polypeptide and not substantially immunoreactive with other human polypeptides.
19. The antibody of claim 18 which is a monoclonal antibody.

20. A method for diagnosing a polymorphism associated with predisposition for breast and ovarian cancer in a human comprising: detecting a germline alteration of the wild-type BRCA1 gene or its expression products in a human sample, said alteration indicating a predisposition to said cancer.

5 21. The method of claim 20 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.

22. The method of claim 20 wherein said expression product is mRNA of BRCA1.

23. The method of claim 22 wherein the alteration of the wild-type BRCA1 gene mRNA is detected by hybridization of the mRNA from said tissue sample to a BRCA1 gene probe.

10 24. The method of claim 20 wherein an alteration is detected in the regulatory regions of the BRCA1 gene.

25. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by observing shifts in electrophoretic mobility of single-stranded DNA on non-denaturing polyacrylamide gels.

15 26. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by hybridization of a BRCA1 gene probe to genomic DNA isolated from said tissue.

27. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by hybridization with an allele-specific probe.

20 28. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplifying all or part of the BRCA1 gene in said tissue to produce an amplified sequence and sequencing the amplified sequence.

29. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplifying all or part of the BRCA1 gene for a specific BRCA1 allele.

30. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by molecularly cloning all or part of the BRCA1 gene in said tissue to produce a cloned sequence and sequencing the cloned sequence.
31. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by identifying a mismatch between molecules (1) BRCA1 gene genomic DNA or BRCA1 mRNA isolated from said tissue, and (2) a nucleic acid probe complementary to the human wild-type BRCA1 gene DNA, when molecules (1) and (2) are hybridized to each other to form a duplex.
32. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplification of BRCA1 gene sequences in said tissue and hybridization of the amplified sequences to nucleic acid probes which comprise wild-type BRCA1 gene sequences.
33. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplification of BRCA1 gene sequences in said tissue and hybridization of the amplified sequences to nucleic acid probes which comprise non-wild-type BRCA1 gene sequences.
34. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by screening for a deletion mutation.
35. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by screening for a point mutation.
36. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by screening for an insertion.
37. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by *in situ* hybridization of the BRCA1 gene with nucleic acid probes which comprise the BRCA1 gene.
38. The method of claim 20 wherein the expression products are protein molecules.

39. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by immunoblotting.
40. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by immunocytochemistry.
- 5 41. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by assaying for binding interactions between BRCA1 gene protein isolated from said tissue and its ligand.
42. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by assaying for the inhibition of biochemical activity of its binding partner.
- 43/ A method for confirming the lack of a predisposition for cancer at the BRCA1 locus in a human, which comprises detecting the wild-type BRCA1 gene or its expression products in a human sample, the presence of the wild-type gene or expression products indicating no predisposition to cancer at the BRCA1 locus.
44. The method of claim 43 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.
45. The method of claim 43 wherein said expression product is mRNA of BRCA1.
46. The method of claim 43 wherein the expression products are protein molecules.
- 47/ A method for diagnosing or prognosing tissue of a lesion in a human, comprising: detecting a somatic alteration of wild-type BRCA1 gene or its expression products in tissue isolated from said human, said alteration indicating mutation of said tissue.
- 20 48. The method of claim 47 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.

49. The method of claim 47 wherein said expression product is mRNA of BRCA1.
50. The method of claim 49 wherein the alteration of the wild-type BRCA1 gene mRNA is detected by hybridization of the mRNA from said tissue sample to a BRCA1 gene probe.
51. The method of claim 47 wherein the alteration is detected in the regulatory region of the
5 BRCA1 gene.
52. The method of claim 47 wherein the expression products are protein molecules.
53. A method for confirming the lack of a mutation at the BRCA1 locus in a human, which comprises detecting the wild-type BRCA1 gene or its expression products in a human sample, the presence of the wild-type gene or expression products indicating no mutation resulting from the BRCA1 locus.
54. The method of claim 53 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.
55. The method of claim 53 wherein said expression product is mRNA of BRCA1.
56. The method of claim 53 wherein the expression products are protein molecules.
- 15 57. A method for supplying a wild-type BRCA1 gene function or a BRCA1 function substantially similar to the wild-type to a cell which has lost said gene function or has altered gene function by virtue of a mutation in the BRCA1 gene, comprising: introducing a wild-type BRCA1 gene nucleic acid or nucleic acid substantially homologous to said wild-type BRCA1 gene nucleic acid and capable of suppression of a transformed state of said cell.
- 20 58. The method of claim 57 wherein said nucleic acid is a portion of wild-type BRCA1 gene, said portion encoding a part of the wild-type BRCA1 gene polypeptide which is required for suppression of a transformed state of said cell.

59. The method of claim 57 wherein said nucleic acid is a portion of wild-type BRCA1 gene, said portion encoding all of the wild-type BRCA1 gene polypeptide which is required for suppression of a transformed state of said cell.
60. The method of claim 57 wherein said nucleic acid is substantially homologous to a portion of the wild-type BRCA1 gene, said portion encoding a polypeptide which is substantially homologous to part of the wild-type BRCA1 polypeptide and has a substantially similar function to the part of the wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell.
61. The method of claim 57 wherein said nucleic acid is substantially homologous to a portion of the wild-type BRCA1 gene, said portion encoding a polypeptide which is substantially homologous to all of the wild-type BRCA1 polypeptide and has a substantially similar function to all of the wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell.
62. The method of claim 57 wherein said nucleic acid contains the BRCA1 gene regulatory sequences.
63. The method of claim 57 wherein said nucleic acid is incorporated into the genome of said cell.
64. A method for supplying a wild-type BRCA1 gene function or a BRCA1 function substantially similar to the wild-type to a cell which has lost said gene function or has altered gene function by virtue of a mutation in the BRCA1 gene, comprising: introducing into the cell all or a part of a wild-type BRCA1 polypeptide which is required for suppression of a transformed state of said cell.
65. A method for supplying a wild-type BRCA1 gene function or a BRCA1 function substantially similar to the wild-type to a cell which has lost said gene function or has altered gene function by virtue of a mutation in the BRCA1 gene, comprising: introducing into the cell a molecule which mimics the function of wild-type BRCA1 polypeptide.

66. The method of claim 65 wherein said molecule is a polypeptide substantially homologous to the complete wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell and has a substantially similar function as the wild-type BRCA1 polypeptide.
67. The method of claim 65 wherein said molecule is a polypeptide substantially homologous to a portion of the wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell and has a substantially similar function as the portion of the wild-type BRCA1 polypeptide.
68. A method for screening potential cancer therapeutics which comprises: combining a BRCA1 binding partner, BRCA1 polypeptide and a compound suspected of being a cancer therapeutic and determining the amount of binding of the BRCA1 polypeptide to its binding partner.
69. A method for screening potential cancer therapeutics which comprises: combining a BRCA1 binding partner and a compound suspected of being a cancer therapeutic and measuring the biological activity of the binding partner.
70. A method for screening potential cancer therapeutics which comprises: growing a transformed eukaryotic host cell containing an altered BRCA1 gene in the presence of a compound suspected of being a cancer therapeutic and determining the rate of growth of said host cell.
71. A transgenic animal which carries an altered BRCA1 allele from a second animal in its genome.
72. The transgenic animal of claim 71 wherein the altered BRCA1 allele is a deletion.
73. The transgenic animal of claim 71 wherein the altered BRCA1 allele is a disrupted allele.

74. / A method for identifying a mutant BRCA1 nucleotide sequence in a suspected mutant BRCA1 allele which comprises comparing the nucleotide sequence of the suspected mutant BRCA1 allele with the wild-type BRCA1 nucleotide sequence, wherein a difference between the suspected mutant and the wild-type sequences identifies a mutant BRCA1 nucleotide sequence.

5

11/23/2011 10:00:00 AM

DECLARATION AND POWER OF ATTORNEY

As 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

<u>Prior Foreign Application(s)</u>	<u>Priority Claimed</u>
<u>Appln. No.</u>	<u>Country</u> <u>Date Filed</u> <u>Yes</u> <u>No</u>
<u>Appln. No.</u>	<u>Country</u> <u>Date Filed</u> <u>Yes</u> <u>No</u>
<u>Appln. No.</u>	<u>Country</u> <u>Date Filed</u> <u>Yes</u> <u>No</u>

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

<u>08/300,266</u> (Appln. No.)	<u>02 September 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/289,221</u> (Appln. No.)	<u>12 August 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/249,602</u> (Appln. No.)	<u>20 May 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

Full Name of Sole or
First Inventor Mark H. SKOLNICK

Inventor's Signature _____

Date of Signature _____

Residence Address 460 Northmount Way
Salt Lake City, Utah 84103

Post Office Address SAME

Country of Citizenship USA

Full Name of Second
Inventor David E. GOLDBAR

Inventor's Signature David E. Goldgar

Date of Signature October 23 1995

Residence Address 2707 East Kelly Lane
Salt Lake City, Utah 84117

Post Office Address SAME

Country of Citizenship USA

Full Name of Third
Inventor Yoshio MIKI

Inventor's Signature _____

Date of Signature _____

Residence Address 40 South 900 East
Salt Lake City, Utah 84102

Post Office Address SAME

Country of Citizenship JAPAN

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature

Date of Signature

Residence Address

745 West Creek Bend, #2807
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

445960-0000

Full Name of Seventh
Inventor

Donna M. SHATTUCK-EIDENS

Inventor's Signature

Date of Signature

Residence Address

1874 Texas Street
Salt Lake City, Utah 84108

Post Office Address

SAME.

Country of Citizenship

USA.

Full Name of Eighth
Inventor

Sean V. TAVTIGIAN

Inventor's Signature

Date of Signature

Residence Address

557 East 1st Avenue, #3
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Ninth
Inventor

Roger W. WISEMAN

Inventor's Signature

Date of Signature

Residence Address

Post Office Address

SAME

Country of Citizenship

USA

P. Andrew FUTREAL

1. The first part of the document is a title page. It contains the title of the document, the author's name, and the date of the document. The title is "The First Part of the Document". The author's name is "John Doe". The date is "1/1/2020".

[illegible]

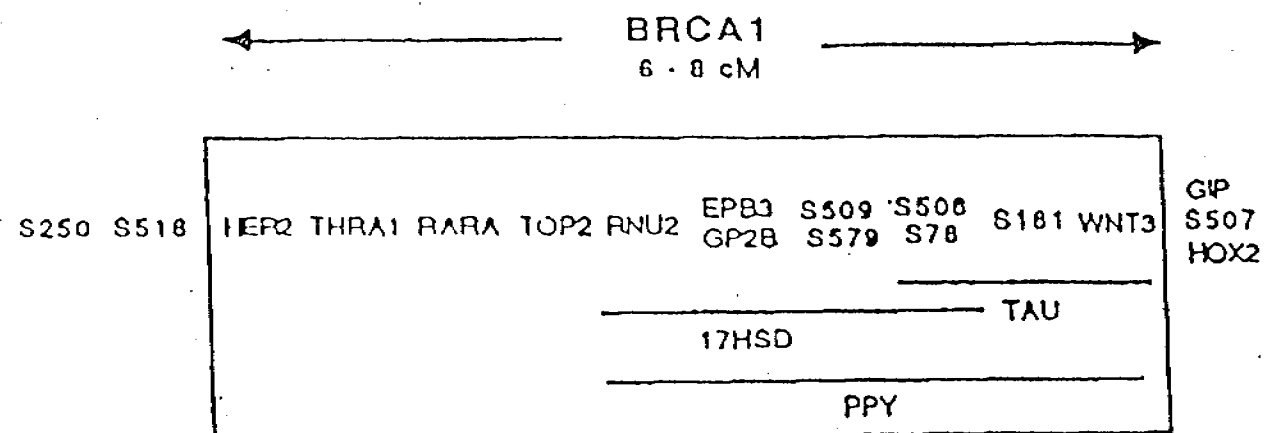
Downloaded from <http://www.sagepub.com> at NANYANG TECH UNIV LIBRARY on June 11, 2015

SAME

USA

1055100 100500000

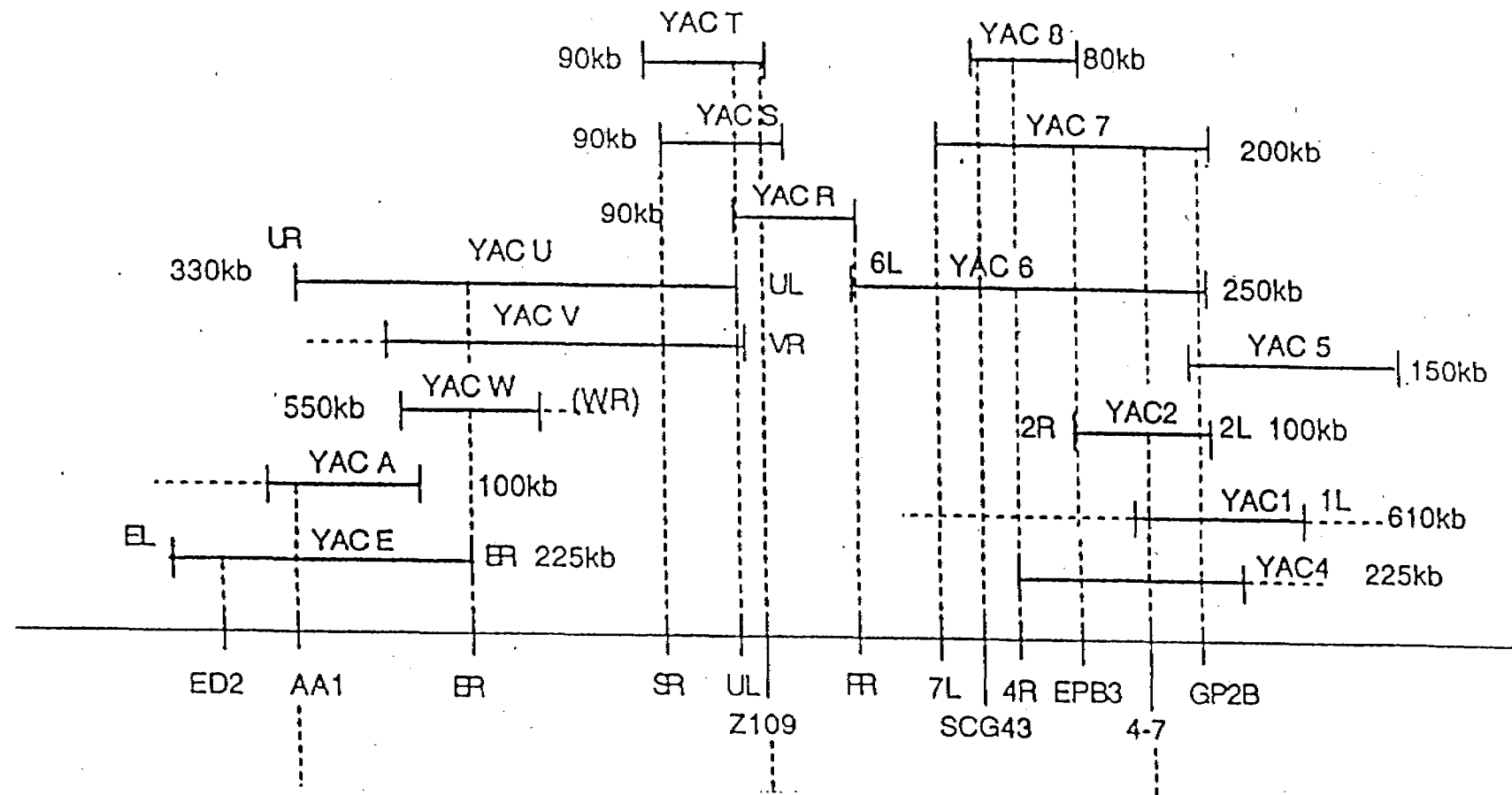
PRINT OF DRAWING
AS ORIGINALLY FILED



Map of the early onset breast and ovarian cancer region (BRCA1)

FIGURE 1

**PRINT OF DRAWING
AS ORIGINALLY FILED**



08/308104

000000 000000

Schematic Map of STS', Pls, and Bac's in the Region

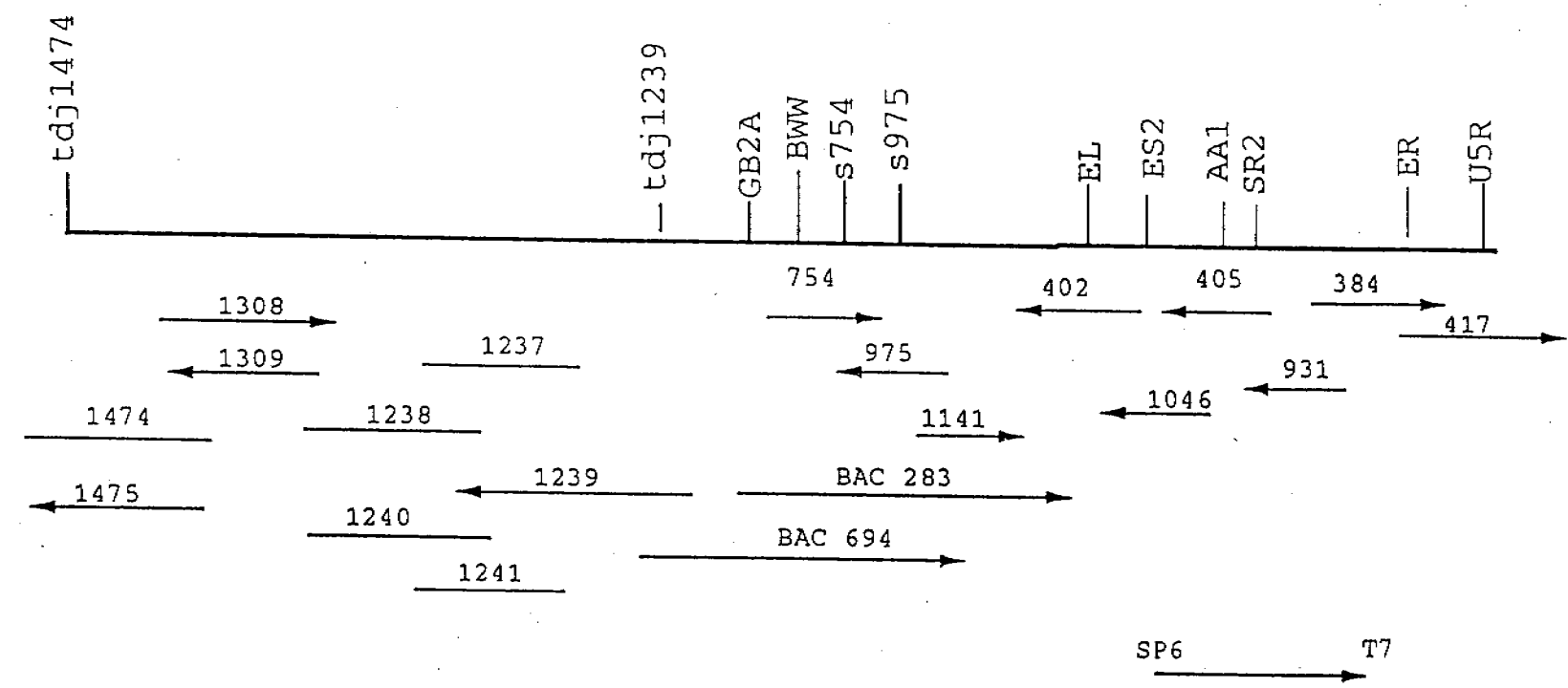


FIGURE 3

PRINT OF DRAWING
AS ORIGINALLY FILED

08/308104

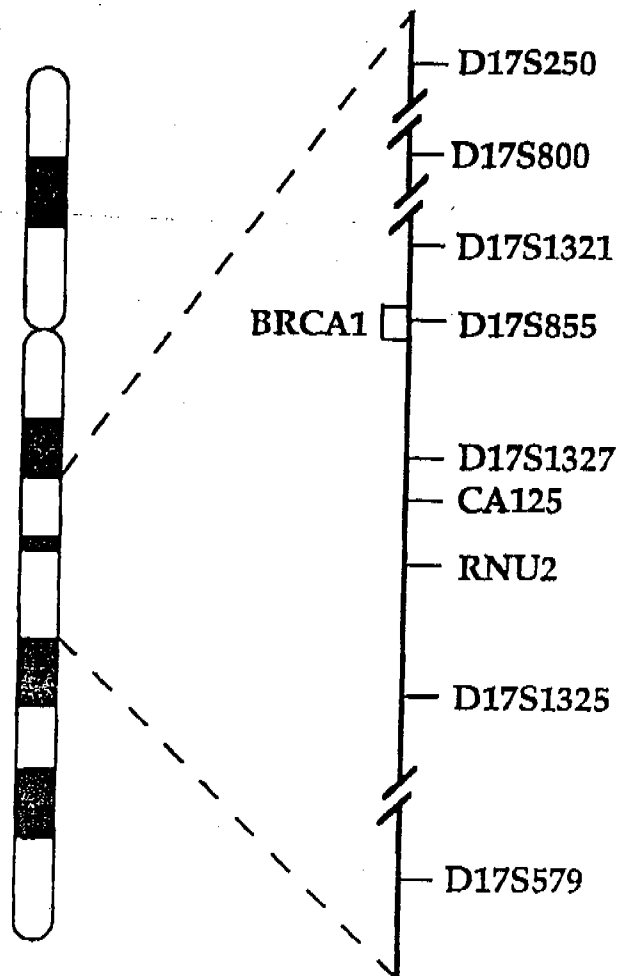


FIGURE 4

h0200* h020000

PRINT OF DRAWING
AS ORIGINALLY FILED

BRCA1
RPT1
RIN1
RFP
C3HC4 motif

CPICLELIKEPVSTK-CDHIFCKFCMLKLLNQKK---GPSQCPLCK
CPICLELLKEPVSA-D-CNHSFCRACITLNYESNRNTDGKGNCVP
CPICLDMLKNTMTTKECLHRFCSDCIVTALRS-----GNKECPTCR
CPVCLQYFAEPMMLD-CGHNICCACLARCWGTAC---TNVSCPQCR
C--C-----C-H--C--C-----C--C

FIGURE 5

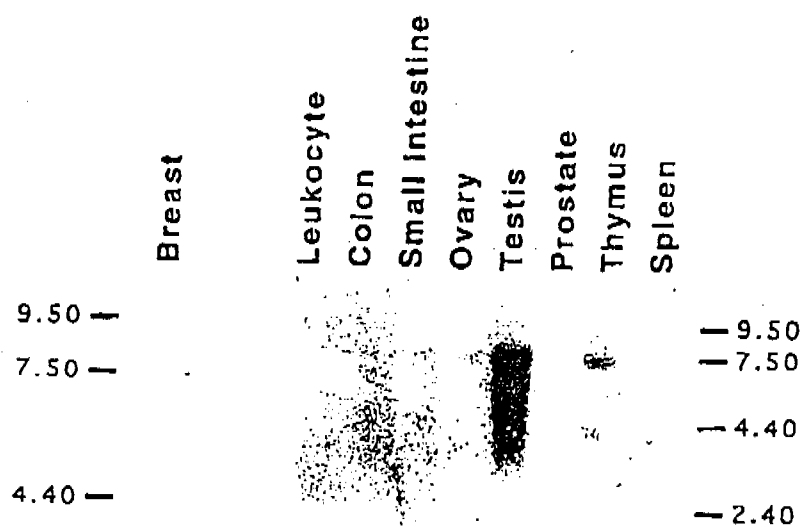


FIGURE 7

PRINT OF DRAWING
AS ORIGINALLY FI

08/308104

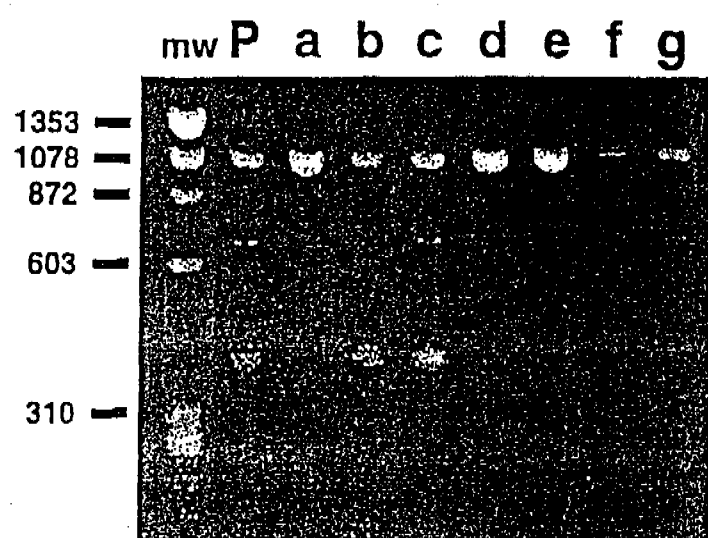


FIGURE 8A

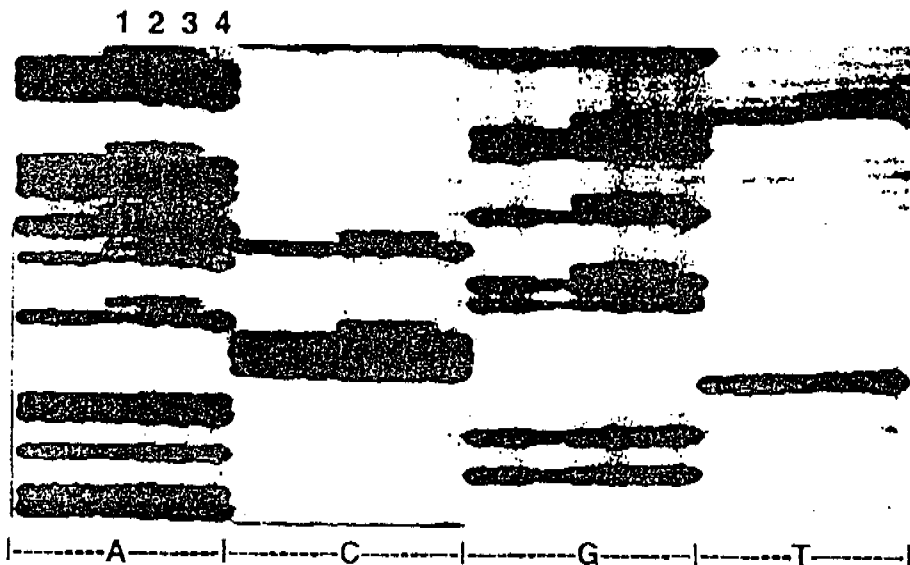
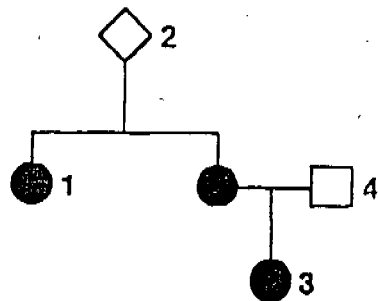
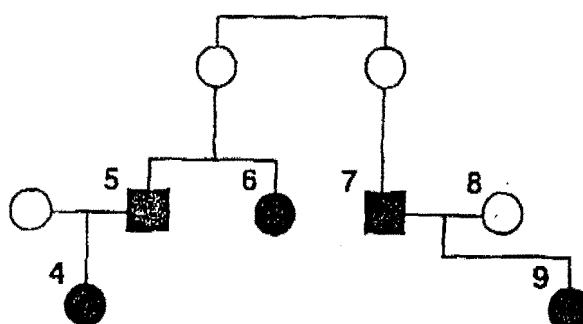


FIGURE 8B



PM1

	4	5	6	7	8	9	
							A
							G
							A
							G

PM7

								A
								G
								A
								G

FIGURE 8C



UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NUMBER	FILING DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
--------------------	-------------	-----------------------	------------------------

08/308,104 09/16/94 SKOLNICK

M 248841093473

03D1/1206

VENABLE, BAETJER, HOWARD & CIVILETTI
SUITE 1000
1201 NEW YORK AVENUE, N.W.
WASHINGTON, DC 20005

0000

DATE MAILED:

12/06/94

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 \$ 130.60 for large entities or \$ 65.00 for small entities who have filed a verified statement claiming such status. The surcharge is set forth in 37 CFR 1.16(e).

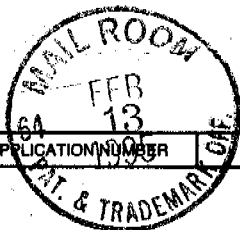
If all required items on this form are filed within the period set below, the total amount owed by applicant as a ☒ large entity, ☐ small entity (verified statement filed), is \$ 3416.00

Applicant is given **ONE MONTH FROM THE DATE OF THIS LETTER, OR TWO MONTHS FROM THE FILING DATE** of this application, **WHICHEVER IS LATER**, 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).

- ☒ The statutory basic filing fee is: ☒ missing ☐ insufficient. Applicant as a ☒ large entity ☐ small entity, must submit \$ 730.00 to complete the basic filing fee.
 - ☒ Additional claim fees of \$ 2556.00 as a ☒ large entity, ☐ small entity, including any required multiple dependent claim fee, are required. Applicant must submit the additional claim fees or cancel the additional claims for which fees are due.
 - ☒ The oath or declaration:
☒ is missing.
☐ does not cover items omitted at time of execution.
- An oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date is required.
- ☐ The oath or declaration does not identify the application to which it applies. An oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date, is required.
 - ☐ The signature to the oath or declaration is: ☐ missing; ☐ a reproduction; ☐ by a person other than the inventor or a 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.
 - ☐ 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.
 - ☐ The application was filed in a language other than English. Applicant must file a verified English translation of the application and a fee of \$ _____ under 37 CFR 1.17(k), unless this fee has already been paid.
 - ☐ A \$ _____ processing fee is required for returned checks. (37 CFR 1.21(m)).
 - ☐ Your filing receipt was mailed in error because check was returned without payment.
 - ☐ The application does not comply with the Sequence Rules. See attached Notice to Comply with Sequence Rules 37 CFR 1.821-1.825.
 - ☒ Other.

Direct the response and any questions about this notice to C. M. Hayward Application Processing Division, Special Processing and Correspondence Branch (703) 308-1202.

A copy of this notice MUST be returned with the response.



7 1-105 2556-103 A/n
UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

APPLICATION NUMBER 08/308,104	FILING DATE 09/16/94	FIRST NAMED APPLICANT SKOLNICK	ATTY. DOCKET NO./TITLE 248841093470
----------------------------------	-------------------------	-----------------------------------	--

VENABLE, BAETJER, HOWARD & CIVILETTI
SUITE 1000
1201 NEW YORK AVENUE, N.W.
WASHINGTON, DC 20005

DATE MAILED: 12/06/94

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 \$_____ for large entities or \$_____ for small entities who have filed a verified statement claiming such status. The surcharge is set forth in 37 CFR 1.16(e).

If all required items on this form are filed within the period set below, the total amount owed by applicant as a ☒ large entity, ☐ small entity (verified statement filed), is \$_____.

Applicant is given **ONE MONTH FROM THE DATE OF THIS LETTER, OR TWO MONTHS FROM THE FILING DATE** of this application, **WHICHEVER IS LATER**, 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).

1. ☒ The statutory basic filing fee is: ☐ missing ☐ insufficient. Applicant as a ☒ large entity ☐ small entity, must submit \$_____ to complete the basic filing fee.
2. ☒ Additional claim fees of \$_____ as a ☐ large entity, ☐ small entity, including any required multiple dependent claim fee, are required. Applicant must submit the additional claim fees or cancel the additional claims for which fees are due.
3. ☒ The oath or declaration:
☒ is missing.
☐ does not cover items omitted at time of execution.

An oath or declaration in compliance with 37 CFR 1.63, identifying the application by the above Application Number and Filing Date is required.
4. ☐ The oath or declaration does not identify the application to which it applies. An 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 to the oath or declaration is: ☐ missing; ☐ a reproduction; ☐ by a person other than the inventor or a 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.
6. ☐ 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.
7. ☐ The application was filed in a language other than English. Applicant must file a verified English translation of the application and a fee of \$_____ under 37 CFR 1.17(k), unless this fee has already been paid.
8. ☐ A \$_____ processing fee is required for returned checks. (37 CFR 1.21(m)).
9. ☐ Your filing receipt was mailed in error because check was returned without payment.
10. ☐ The application does not comply with the Sequence Rules. See attached Notice to Comply with Sequence Rules 37 CFR 1.821-1.825.
11. ☐ Other

Direct the response and any questions about this notice to _____, Application Processing Division, Special Processing and Correspondence Branch (703) 308-1202.

A copy of this notice MUST be returned with the response.



OSC 0
12/6/94

Sequence
#3

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE RECEIVED

FEB 21 1995

APPLICATION DIVISION

In re. Application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

RECEIVED

APR 7 1995

OFFICE OF PETITIONS

TRANSMITTAL LETTER

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

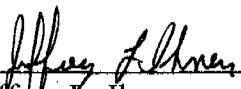
Transmitted herewith in response to the Notice To File Missing Parts mailed 6 December 1994 for the above-identified application are the following papers:

1. A copy of the signed Declarations and Powers of Attorney, the originals of which are attached to Applicants' Petition Under 37 CFR §1.48 to Add Joint Inventors;
2. Petition Under 37 CFR §1.48 to Add Joint Inventors, including attached Verified Statements of Facts, original Declarations and Powers of Attorney and check for \$130.00 for the petition fee, with authorization to charge any discrepancies to our deposit account;
3. A check for \$3,536.00, covering the filing fee (\$3,286.00), surcharge (\$130.00), and a portion of the extension fee (\$120.00);
4. A Petition for Extension of Two (2) Months, in triplicate, with authorization to charge Deposit Account 22-0261 for the remainder of the extension fee (\$250.00);
5. Copies of Form PTO-1533 and its attached Notice to Comply;
6. A Preliminary Amendment, including substitute pages 94-145 of specification, which contain the required Sequence Listing in accordance with the Notice to Comply (identified in 5 above); and
7. A signed Statement Pursuant to 37 CFR §1.821(f), with computer disk of sequence listing.

-2-

If any discrepancies in fees are noted, the same are to be charged or credited to Deposit Account No. 22-0261. This transmittal letter is submitted in triplicate.

Respectfully submitted,



Jeffrey L. Ihnen
Registration No. 28,957

VENABLE, BAETJER, HOWARD & CIVILETTI
1201 New York Avenue, N.W., Suite 1000
Washington, D.C. 20005
(202) 962-4810

Atty. Docket No. 24884-109347-3
Dated: 13 February 1995



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

PETITION UNDER 37 CFR §1.48 TO ADD JOINT INVENTORS

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

It is respectfully petitioned that the above-identified application, filed under 37 CFR §1.53(d), be amended to add as joint inventors the following individuals:

Roger W. WISEMAN
P. Andrew FUTREAL

Through error and without any deceptive intention, the application was made in the names of Mark H. Skolnick, David E. Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith D. Harshman, Donna M. Shattuck-Eidens and Sean V. Tavtigian as coinventors, which is apparent from the Verified Statement of Facts submitted concurrently herewith.

In support of this Petition, Applicants submit the following documents:

- (1) A Verified Statement of Facts by the inventors named in the application: Mark H. Skolnick, David E. Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith D. Harshman, Donna M. Shattuck-Eidens and Sean V. Tavtigian;
- (2) A Declaration and Power of Attorney under 37 CFR §1.63 by all of the *actual* joint inventors;
- (4) A check in the amount of \$130.00 for the fee required under 37 CFR §1.17(h). If any discrepancies in fee are noted, the same

should be charged or credited to Deposit Account No. 22-0261.
Three copies of this Petition are enclosed.

As is evident from the Verified Statement of Facts, the present invention is related to three applications, in which Mark Skolnick, David Goldgar, Yoshio Miki and Jeff Swenson, all of the University of Utah are co-inventors and a fourth application, in which Mark Skolnick, David Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith Harshman and Donna Shattuck-Eidens (the latter three of Myriad Genetics) are co-inventors. Roger Wiseman and Andrew Futreal were *not* coinventors on any of these applications.

It is clear from the Verified Statement that the decision to prepare and file the grand-parent application ('221 application) of the present application was made on 10 August 1994 with the filing to be made on 12 August 1994. Inventorship of the breast cancer gene was briefly discussed during the 10 August 1994 meeting. Based on the available information, it was determined that Roger Wiseman and Andrew Futreal were not joint inventors of the breast cancer gene, since they did not develop the entire gene sequence and had not discovered any mutations in the gene.

The parent application ('266 application) of the present application was filed as a CIP application of the '221 application on 2 September 1994 to coincide with the mailing of manuscript to Science. The '266 application contained the remainder of the DNA coding sequence which was not disclosed in the '221 application. Since the DNA coding sequence was determined at Myriad, it was determined that the inventorship of the '266 application was the same as the '221 application.

The present application was filed as a CIP application of the '266 application on 16 September 1994 after it was learned that Science had made copies of an early manuscript publically available in order to preserve rights in the event that Science released copies of the revised manuscript. This application contains additional 5' DNA sequence information and mutations not disclosed in the '266 application. Since the DNA sequence information and mutations were determined at Myriad, it was determined that the inventorship of the present application was the same as the '266 application.

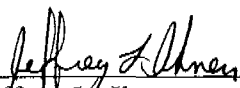
On 5 October 1994, following the filing of the present application, NIEHS claimed that Wiseman and Futreal were joint inventors of the application, since claim 3 in the application was directed to fragments of the breast cancer gene. On 28 November 1994 NIEHS provided information from its research records describing the identification of fragments of cod-

ing sequences by Wiseman and Futreal from materials obtained from Myriad. After a thorough analysis was made of all of the pertinent research records of the University of Utah and Myriad Genetics and of the information provided by NIEHS, it was determined that Wiseman and Futreal had been involved in the identification of two DNA fragments which are part of the breast cancer gene. After the facts were completely analyzed, it was recognized that Wiseman and Futreal are joint inventors as to a fragment claim and had been inadvertently omitted, without any deceptive intent, from the application as filed. As soon as it became apparent that Wiseman and Futreal were properly joint inventors, and in preparation for filing missing parts to the above application, this Petition was diligently made. See *Koret v. Hoskins*, 84 USPQ 461 (CCPA 1950). Accordingly, this Petition is timely.

Further, the statutes and rules governing inventorship conversion are remedial in nature and should be liberally construed in order that the error noted herein may be rectified. *Patterson & Lissner v. Hauck*, 144 USPQ 481 (CCPA 1965); *In re Schmidt*, 130 USPQ 404 (CCPA 1961); *Fritsch v. Lin*, 21 USPQ2d 1731 (BPAI 1991).

Thus, it is submitted that Wiseman and Futreal are properly joint inventors of the presently claimed subject matter, and that their omission from the application as filed was without deceptive intent. It is requested that the inventorship of this application be amended to add Roger W. Wiseman and P. Andre Futreal as coinventors.

Respectfully submitted,



Jeffrey L. Ihnen

Registration No. 28,957

JLI:ch

VENABLE, BAETJER, HOWARD & CIVILETTI
1201 New York Avenue, N.W., #1000
Washington, D.C. 20005
(202) 962-4810

Atty. Docket No. 24884-109347-3
Dated: 13 February 1995



DECLARATION AND POWER OF ATTORNEY

As 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

<u>Prior Foreign Application(s)</u>	<u>Priority Claimed</u>
<u>Appln. No.</u> <u>Country</u> <u>Date Filed</u>	<u>Yes</u> <u>No</u>
<u>Appln. No.</u> <u>Country</u> <u>Date Filed</u>	<u>Yes</u> <u>No</u>
<u>Appln. No.</u> <u>Country</u> <u>Date Filed</u>	<u>Yes</u> <u>No</u>

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

<u>08/300,266</u> (Appln. No.)	<u>02 September 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/289,221</u> (Appln. No.)	<u>12 August 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/249,602</u> (Appln. No.)	<u>20 May 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

Full Name of Sole or First Inventor	<u>Mark H. SKOLNICK</u>
Inventor's Signature	<u></u>
Date of Signature	<u></u>
Residence Address	<u>460 Northmount Way</u> <u>Salt Lake City, Utah 84103</u>
Post Office Address	<u>SAME</u>
Country of Citizenship	<u>USA</u>
Full Name of Second Inventor	<u>David E. GOLDGAR</u>
Inventor's Signature	<u>David E. Goldgar by Jeffrey L. Ihnen</u>
Date of Signature	<u>13 February 1995</u>
Residence Address	<u>2707 East Kelly Lane</u> <u>Salt Lake City, Utah 84117</u>
Post Office Address	<u>SAME</u>
Country of Citizenship	<u>USA</u>
Full Name of Third Inventor	<u>Yoshio MIKI</u>
Inventor's Signature	<u></u>
Date of Signature	<u></u>
Residence Address	<u>40 South 900 East</u> <u>Salt Lake City, Utah 84102</u>
Post Office Address	<u>SAME</u>
Country of Citizenship	<u>JAPAN</u>

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature

Date of Signature

Residence Address

745 West Creek Bend, #2807
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Seventh
Inventor

Donna M. SHATTUCK-EIDENS

Inventor's Signature

Date of Signature

Residence Address

1874 Texas Street
Salt Lake City, Utah 84108

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Eighth
Inventor

Sean V. TAVTIGIAN

Inventor's Signature

Date of Signature

Residence Address

557 East 1st Avenue, #3
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Ninth
Inventor

Roger W. WISEMAN

Inventor's Signature

Date of Signature

Residence Address

Post Office Address

SAME

Country of Citizenship

USA

-5-

Full Name of Tenth
Inventor

P. Andrew FUTREAL

Inventor's Signature

Date of Signature

Residence Address

Post Office Address

SAME

Country of Citizenship

USA

308,104



DECLARATION AND POWER OF ATTORNEY

As 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

<u>Prior Foreign Application(s)</u>			<u>Priority Claimed</u>	
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes</u>	<u>No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes</u>	<u>No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes</u>	<u>No</u>

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

<u>08/300,266</u> (Appln. No.)	<u>02 September 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/289,221</u> (Appln. No.)	<u>12 August 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/249,602</u> (Appln. No.)	<u>20 May 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

Full Name of Sole or
First Inventor

Mark H. SKOLNICK

Inventor's Signature

Mark Skolnick

Date of Signature

Feb 9, 1995

Residence Address

460 Northmount Way
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Second
Inventor

David E. GOLDGAR

Inventor's Signature

Date of Signature

Residence Address

2707 East Kelly Lane
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Third
Inventor

Yoshio MIKI

Inventor's Signature

Yoshio Miki

Date of Signature

Feb. 9, 1995

Residence Address

40 South 900 East
Salt Lake City, Utah 84102

Post Office Address

SAME

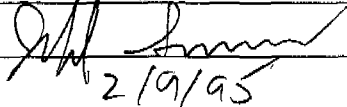
Country of Citizenship

JAPAN

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature


2/9/95

Date of Signature

Residence Address

~~745 West Creek Bend, #2007~~ 1210 E Lorraine Dr
~~Salt Lake City, Utah 84117~~ SLC, UT 84106 

Post Office Address

SAME

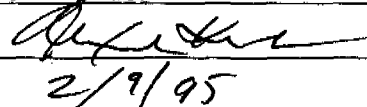
Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature


2/9/95

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Seventh
Inventor

Donna M. SHATTUCK-EIDENS

Inventor's Signature

Donna M. Shattuck Eids

Date of Signature

2/10/95

Residence Address

1874 Texas Street
Salt Lake City, Utah 84108

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Eighth
Inventor

Sean V. TAVTIGIAN

Inventor's Signature

SOT

Date of Signature

2/9/95

Residence Address

557 East 1st Avenue, #3
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Ninth
Inventor

Roger W. WISEMAN

Inventor's Signature

Date of Signature

Residence Address

Post Office Address

SAME

Country of Citizenship

USA

-5-

Full Name of Tenth
Inventor

P. Andrew FUTREAL

Inventor's Signature

Date of Signature

Residence Address

Post Office Address

SAME

Country of Citizenship

USA

10/02/95

10/02/95 CREATIVE HOUSE 00520891209 8911/95

P.01

The undersigned hereby gives Jeffery L. Ihnen power of attorney to sign the Declaration and Power of Attorney for US application serial numbers 08/289,221; 08/300,266; and 08/308,104.

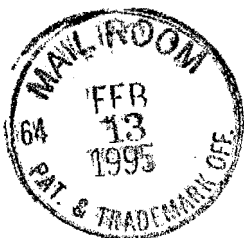
Feb 7 1995
Dated

David E. Goldgar
David E. Goldgar, Ph.D.

0002 001

MYRIAD GENETICS

02 09:83 18:30 FAX 501 582 3405



DECLARATION AND POWER OF ATTORNEY

As 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

<u>Prior Foreign Application(s)</u>			<u>Priority Claimed</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes No</u>

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

<u>08/300,266</u> (Appln. No.)	<u>02 September 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/289,221</u> (Appln. No.)	<u>12 August 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/249,602</u> (Appln. No.)	<u>20 May 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

Full Name of Sole or
First Inventor

Mark H. SKOLNICK

Inventor's Signature

Date of Signature

Residence Address

460 Northmount Way
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Second
Inventor

David E. GOLDSAR

Inventor's Signature

David E. Goldsar by Jeffrey L. Ihnen

Date of Signature

13 February 1975

Residence Address

2707 East Kelly Lane
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Third
Inventor

Yoshio MIKI

Inventor's Signature

Date of Signature

Residence Address

40 South 900 East
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

JAPAN

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature

Date of Signature

Residence Address

745 West Creek Bend, #2807
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

445 976 01 445 976 01

Full Name of Seventh
Inventor

Donna M. SHATTUCK-EIDENS

Inventor's Signature

Date of Signature

Residence Address

1874 Texas Street
Salt Lake City, Utah 84108

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Eighth
Inventor

Sean V. TAVTIGIAN

Inventor's Signature

Date of Signature

Residence Address

557 East 1st Avenue, #3
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Ninth
Inventor

Roger W. WISEMAN

Inventor's Signature

Date of Signature

Residence Address

Post Office Address

SAME

Country of Citizenship

USA

44-38861-10000000

Full Name of Tenth
Inventor

P. Andrew FUTREAL

Inventor's Signature

Date of Signature

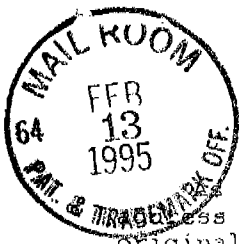
Residence Address

Post Office Address

SAME

Country of Citizenship

USA



DECLARATION AND POWER OF ATTORNEY

I, the 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

<u>Prior Foreign Application(s)</u>			<u>Priority Claimed</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes No</u>

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

<u>08/300,266</u> (Appln. No.)	<u>02 September 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/289,221</u> (Appln. No.)	<u>12 August 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/349,602</u> (Appln. No.)	<u>20 May 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

Full Name of Sole or First Inventor	<u>Mark H. SKOLNICK</u>
Inventor's Signature	_____
Date of Signature	_____
Residence Address	<u>460 Northmount Way</u> <u>Salt Lake City, Utah 84103</u>
Post Office Address	<u>SAME</u>
Country of Citizenship	<u>USA</u>
Full Name of Second Inventor	<u>David E. GOLDGAR</u>
Inventor's Signature	_____
Date of Signature	_____
Residence Address	<u>2707 East Kelly Lane</u> <u>Salt Lake City, Utah 84117</u>
Post Office Address	<u>SAME</u>
Country of Citizenship	<u>USA</u>
Full Name of Third Inventor	<u>Yoshio MIKI</u>
Inventor's Signature	_____
Date of Signature	_____
Residence Address	<u>40 South 900 East</u> <u>Salt Lake City, Utah 84102</u>
Post Office Address	<u>SAME</u>
Country of Citizenship	<u>JAPAN</u>

Full Name of Fourth
Inventor

JEFF SWENSON

Inventor's Signature

Jeff Swenson

Date of Signature

2/9/95

Residence Address

~~745 West Creek Bend, #2807~~
~~Salt Lake City, Utah 84117~~

1210 E Lowaine Dr
SLC, UT 84106

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature

Alexander Kamb

Date of Signature

2/9/95

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

KD Harshman

Date of Signature

2-9-95

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature

Date of Signature

Residence Address

745 West Creek Bend, #2807
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

208 104

Full Name of Seventh
Inventor

Donna M. SHATTUCK-EIDENS

Inventor's Signature

Date of Signature

Residence Address

1874 Texas Street
Salt Lake City, Utah 84108

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Eighth
Inventor

Sean V. TAVTIGIAN

Inventor's Signature

Date of Signature

Residence Address

557 East 1st Avenue, #3
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Ninth
Inventor

Roger W. WISEMAN

Inventor's Signature

Date of Signature

Residence Address

Roger Wiseman
2/9/95
515 Woodwinds Drive
Durham, NC 27713

Post Office Address

SAME

Country of Citizenship

USA

308,104

Full Name of Tenth
Inventor

P. Andrew FUTREAL

Inventor's Signature

P. Andrew Futreal

Date of Signature

2/9/95

Residence Address

220 East Markham Ave
Durham NC 27701

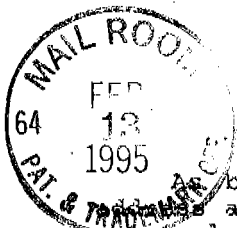
Post Office Address

SAME

Country of Citizenship

USA

16 MAR 1995



DECLARATION AND POWER OF ATTORNEY

As 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

Prior Foreign Application(s)			Priority Claimed
Appln. No.	Country	Date Filed	Yes No
Appln. No.	Country	Date Filed	Yes No
Appln. No.	Country	Date Filed	Yes No

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

08/300,266 (Appln. No.)	02 September 1994 (Filing Date)	Pending (Status: patented, pending, abandoned)
08/289,221 (Appln. No.)	12 August 1994 (Filing Date)	Pending (Status: patented, pending, abandoned)
08/249,602 (Appln. No.)	20 May 1994 (Filing Date)	Pending (Status: patented, pending, abandoned)

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,

-2-

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

Full Name of Sole or
First Inventor

Mark H. SKOLNICK

Inventor's Signature

Date of Signature

Residence Address

460 Northmount Way
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Second
Inventor

David E. GOLDGAR

Inventor's Signature

Date of Signature

Residence Address

2707 East Kelly Lane
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Third
Inventor

Yoshio MIKI

Inventor's Signature

Date of Signature

Residence Address

40 South 900 East
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

JAPAN

308,104

- 3 -

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature

Date of Signature

Residence Address

745 West Creek Bend, #2807
Salt Lake City, Utah 84117

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

1. 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352

308.104

-4-

Full Name of Seventh
Inventor

Donna M. SHATTUCK-EIDENS

Inventor's Signature

Date of Signature

Residence Address

1874 Texas Street
Salt Lake City, Utah 84108

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Eighth
Inventor

Sean V. TAVTIGIAN

Inventor's Signature

Date of Signature

Residence Address

557 East 1st Avenue, #3
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Ninth
Inventor

Roger W. WISEMAN

Inventor's Signature

Date of Signature

Residence Address

515 Woodwinds Drive
Durham, NC 27713

Post Office Address

SAME

Country of Citizenship

USA

-5-

Full Name of Tenth

Inventor:

P. And w. FUTREAL

Inventor's Signature

[Handwritten Signature]

Date of Signature

2/9/95

Residence Address

220 East Markham Ave
Durham NC 27701

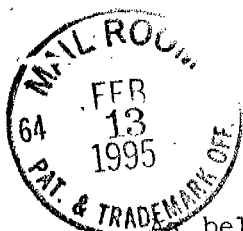
Post Office Address

SAME

Country of Citizenship

USA

202 962 8300



DECLARATION AND POWER OF ATTORNEY

I, 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

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

and which was filed on 16 September 1994 and assigned Serial No. 08/308,104.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) on which priority is claimed:

<u>Prior Foreign Application(s)</u>			<u>Priority Claimed</u>	
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes</u>	<u>No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes</u>	<u>No</u>
<u>Appln. No.</u>	<u>Country</u>	<u>Date Filed</u>	<u>Yes</u>	<u>No</u>

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose all information known to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

<u>08/300,266</u> (Appln. No.)	<u>02 September 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/289,221</u> (Appln. No.)	<u>12 August 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)
<u>08/249,602</u> (Appln. No.)	<u>20 May 1994</u> (Filing Date)	<u>Pending</u> (Status: patented, pending, abandoned)

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.

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; Royal W. Craig, Registration No. 34,145; and E. Brendan Magrab, Registration No. 36,205; all of Venable, Baetjer, Howard & Civiletti, whose address is 1201 New York Avenue, N.W., Suite 1000, Washington, D.C. 20005, telephone number (202) 962-4800, my attorneys and/or agents with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Please direct all correspondence and telephone calls to Venable, Baetjer, Howard & Civiletti at the above address and telephone number.

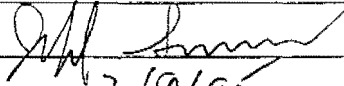
Full Name of Sole or First Inventor	Mark H. SKOLNICK
Inventor's Signature	<i>Mark Skolnick</i>
Date of Signature	Feb 9, 1995
Residence Address	460 Northmount Way Salt Lake City, Utah 84103
Post Office Address	SAME
Country of Citizenship	USA
Full Name of Second Inventor	David E. GOLDGAR
Inventor's Signature	
Date of Signature	
Residence Address	2707 East Kelly Lane Salt Lake City, Utah 84117
Post Office Address	SAME
Country of Citizenship	USA
Full Name of Third Inventor	Yoshio MIKI
Inventor's Signature	<i>Yoshio Miki</i>
Date of Signature	Feb. 9, 1995
Residence Address	40 South 900 East Salt Lake City, Utah 84102
Post Office Address	SAME
Country of Citizenship	JAPAN

filed to 100000

Full Name of Fourth
Inventor

Jeff SWENSON

Inventor's Signature


2/9/95

Date of Signature

Residence Address

~~745 West Creek Bend, #2807~~ 1210 E Lorraine Dr
~~Salt Lake City, Utah 84117~~ SL, UT 84106 

Post Office Address

SAME

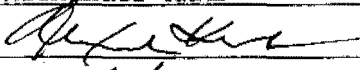
Country of Citizenship

USA

Full Name of Fifth
Inventor

Alexander KAMB

Inventor's Signature


2/9/95

Date of Signature

Residence Address

1103 East 600 South
Salt Lake City, Utah 84102

Post Office Address

SAME

Country of Citizenship

USA

Full Name of Sixth
Inventor

Keith D. HARSHMAN

Inventor's Signature

Date of Signature

Residence Address

159 North F Street
Salt Lake City, Utah 84103

Post Office Address

SAME

Country of Citizenship

USA

100-100000-100000

Residence Address

2/9/95

~~745 West Creek Bend, #2807~~
~~Salt Lake City, Utah 84117~~

1210 E Lowry Ave Dr
SIC, UT 84106

SAME

USA

2/9/95

2/9/95

1103 East 600 South
Salt Lake City, Utah 84102

SAME

USA

2-9-95

2.9.95

159 North F Street
Salt Lake City, Utah 84103

SAME

USA

P. Andrew FUTREAL

SAME

USA

[illegible]



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re: Application of:

MARKUS SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

PRELIMINARY AMENDMENT

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

In response to the Notice to File Missing Parts mailed 6 December 1994, please amend the above-identified application as follows:

Please substitute attached pages 94-145 for pages 94-129 of the application as originally filed.

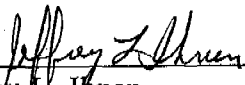
REMARKS

The specification has been amended to substitute pages 94-145 for originally filed pages 94-129. The substitute pages contain the required paper "Sequence Listing" (pages 94-135) for DNA sequences found in the sequence listing of the originally filed specification and for the DNA sequences detailed in Table 9 of the specification as originally filed, and the

-2-

original claims and abstract (pages 136-145) as appeared at original pages 120-129. It is believed that none of these amendments constitute new matter.

Respectfully submitted,



Jeffrey L. Ihnen
Registration No. 28,957

JLI:ch

VENABLE, BAETJER, HOWARD & CIVILETTI
1201 New York Avenue, N.W., #1000
Washington, D.C. 20005
(202) 962-4810

Atty. Docket No. 24884-109347-3
Dated: 13 February 1995



-94-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Skolnick, Mark H.
Goldgar, David E.
Miki, Yoshio
Swenson, Jeff
Kamb, Alexander
Harshman, Keith D.
Shattuck-Eidens, Donna M.
Tavtigian, Sean V.
Wiseman, Roger W.
Futreal, P. Andrew
- (ii) TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
Susceptibility Gene
- (iii) NUMBER OF SEQUENCES: 61
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Venable, Baetjer, Howard & Civiletti
 - (B) STREET: 1201 New York Avenue, N.W., Suite 1000
 - (C) CITY: Washington
 - (D) STATE: DC
 - (F) ZIP: 20005
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: WordPerfect 5.1/5.2
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/308,104
 - (B) FILING DATE: 16-SEP-1994
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/300,266
 - (B) FILING DATE: 02-SEP-1994
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/289,221
 - (B) FILING DATE: 12-AUG-1994
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/246,602
 - (B) FILING DATE: 20-MAY-1994
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Ihnen, Jeffrey L.
 - (B) REGISTRATION NUMBER: 28,957
 - (C) REFERENCE/DOCKET NUMBER: 24884-109347-3
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 202-962-4810
 - (B) TELEFAX: 202-962-8300

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5857 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 64..5769

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

GGCCCCTGCG CTCAGGAGGC CTTCAACCCTC TGCTCTGGGT AAAGTTCATT GGAACAGAAA	60
GAAATGGATT TATCTGCTCT TCGCGTTGAA GAAGTACAAA ATGTCATTAA TGCTATGCAG	120
AAAATCTTAG AGTGTCCCAT CTGTCTGGAG TTGATCAAGG AACCTGTCTC CACAAAGTGT	180
GACCACATAT TTTGCAAATT TTGCATGCTG AAACCTTCTCA ACCAGAAGAA AGGGCCTTCA	240
CAGTGTCTTT TATGTAAGAA TGATATAACC AAAAGGAGCC TACAAGAAAG TACGAGATTT	300
AGTCAACTTG TTGAAGAGCT ATTGAAAATC ATTTGTGCTT TTCAGCTTGA CACAGGTTTG	360
GAGTATGCAA ACAGCTATAA TTTTGCAAAA AAGGAAAATA ACTCTCCTGA ACATCTAAAA	420
GATGAAGTTT CTATCATCCA AAGTATGGGC TACAGAAACC GTGCCAAAAG ACTTCTACAG	480
AGTGAACCCG AAAATCCTTC CTTGCAGGAA ACCAGTCTCA GTGTCCAACT CTCTAACCTT	540
GGAAGTGTGA GAACTCTGAG GACAAAGCAG CGGATACAAC CTCAAAGAC GTCTGTCTAC	600
ATTGAATTGG GATCTGATTC TTCTGAAGAT ACCGTTAATA AGGCAACTTA TTGCAGTGTG	660
GGAGATCAAG AATTGTTACA AATCACCCTT CAAGGAACCA GGGATGAAAT CAGTTTGGAT	720
TCTGCAAAAA AGGCTGCTTG TGAATTTTCT GAGACGGATG TAACAAATAC TGAACATCAT	780
CAACCCAGTA ATAATGATTT GAACACCACT GAGAAGCGTG CAGCTGAGAG GCATCCAGAA	840
AAGTATCAGG GTAGTTCTGT TTCAAACCTG CATGTGGAGC CATGTGGCAC AAATACTCAT	900
GCCAGCTCAT TACAGCATGA GAACAGCAGT TTATTACTCA CTAAAGACAG AATGAATGTA	960
GAAAAGGCTG AATTCTGTAA TAAAAGCAAA CAGCCTGGCT TAGCAAGGAG CCAACATAAC	1020
AGATGGGCTG GAAGTAAGGA AACATGTAAT GATAGGCGGA CTCCCAGCAC AGAAAAAAG	1080
GTAGATCTGA ATGCTGATCC CCTGTGTGAG AGAAAAGAAT GGAATAAGCA GAACTGCCA	1140
TGCTCAGAGA ATCCTAGAGA TACTGAAGAT GTTCCTTGGA TAACACTAAA TAGCAGCATT	1200
CAGAAAGTTA ATGAGTGGTT TTCCAGAAGT GATGAACTGT TAGGTTCTGA TGAATCACAT	1260
GATGGGGAGT CTGAATCAAA TGCCAAAGTA GCTGATGTAT TGGACGTCCT AATTGAGGTA	1320
GATGAATATT CTGGTTCTTC AGAGAAAATA GACTTACTGG CCAGTGATCC TCATGAGGCT	1380
TTAATATGTA AAAGTGACAG AGTTCACTCC AAATCAGTAG AGAGTAATAT TGAAGGCCAA	1440
ATATTTGGGA AAACCTATCG GAAGAAGGCA AGCCTCCCCA ACTTAAGCCA TGTAAGTAA	1500

AATCTAATTA	TAGGNGCATT	TGTTTCTGAG	CCACAGATAA	TACAAGAGCG	TCCCCTCACA	1560
AATAAATTAA	AGCGTAAAAG	GAGACCTACA	TCAGGCCTTC	ATCCTGAGGA	TTTTATCAAG	1620
AAAGCAGATT	TGGCAGTTCA	AAAGACTCCT	GAAATGATAA	ATCAGGGAAC	TAACCAAACG	1680
GAGCAGAATG	GTCAAGTGAT	GAATATTACT	AATAGTGGTC	ATGAGAATAA	AACAAAAGGT	1740
GATTCTATTC	AGAATGAGAA	AAATCCTAAC	CCAATAGAAT	CACTCGAAAA	AGAATCTGCT	1800
TTCAAAACGA	AAGCTGAACC	TATAAGCAGC	AGTATAAGCA	ATATGGAAC	CGAATTAAAT	1860
ATCCACAATT	CAAAAGCACC	TAAAAAGAAT	AGGCTGAGGA	GGAAGTCTTC	TACCAGGCAT	1920
ATTCATGCGC	TTGAACTAGT	AGTCAGTAGA	AATCTAAGCC	CACCTAATTG	TACTGAATTG	1980
CAAATTGATA	GTTGTTCTAG	CAGTGAAGAG	ATAAAGAAAA	AAAAGTACAA	CCAAATGCCA	2040
GTCAGGCACA	GCAGAAACCT	ACAACATCATG	GAAGGTAAAG	AACCTGCAAC	TGGAGCCAAG	2100
AAGAGTAACA	AGCCAAATGA	ACAGACAAGT	AAAAGACATG	ACAGTGATAC	TTTCCCAGAG	2160
CTGAAGTTAA	CAAATGCACC	TGGTTCTTTT	ACTAAGTGTT	CAAATACCAG	TGAACTTAAA	2220
GAATTTGTCA	ATCCTAGCCT	TCCAAGAGAA	GAAAAAGAAG	AGAACTAGA	AACAGTTAAA	2280
GTGTCTAATA	ATGCTGAAGA	CCCCAAAGAT	CTCATGTTAA	GTGGAGAAAG	GGTTTTGCAA	2340
ACTGAAAGAT	CTGTAGAGAG	TAGCAGTATT	TCCTGGTAC	CTGGTACTGA	TTATGGCACT	2400
CAGGAAAGTA	TCTCGTTACT	GGAAGTTAGC	ACTCTAGGGA	AGGCAAAAAC	AGAACCAAAT	2460
AAATGTGTGA	GTCAGTGTGC	AGCATTGTGA	AACCCCAAGG	GAATAATTCA	TGGTTGTTCC	2520
AAAGATAATA	GAAATGACAC	AGAAGGCTTT	AAGTATCCAT	TGGGACATGA	AGTTAACCAC	2580
AGTCGGGAAA	CAAGCATAGA	AATGGAAGAA	AGTGAACCTG	ATGCTCAGTA	TTTGCAGAAT	2640
ACATTCAAGG	TTTCAAAGCG	CCAGTCATTT	GCTCTGTTTT	CAAATCCAGG	AAATGCAGAA	2700
GAGGAATGTG	CAACATTCTC	TGCCCCACTCT	GGGTCCTTAA	AGAAACAAAG	TCCAAAAGTC	2760
ACTTTTGAAT	GTGAACAAAA	GGAAGAAAAT	CAAGGAAAGA	ATGAGTCTAA	TATCAAGCCT	2820
GTACAGACAG	TTAATATCAC	TGCAGGCTTT	CCTGTGGTTG	GTCAGAAAGA	TAAGCCAGTT	2880
GATAATGCCA	AATGTAGTAT	CAAAGGAGGC	TCTAGGTTTT	GTCTATCATC	TCAGTTCAGA	2940
GGCAACGAAA	CTGGACTCAT	TACTCCAAAT	AACATGGAC	TTTTACAAAA	CCCATATCGT	3000
ATACCACCAC	TTTTTCCCAT	CAAGTCATTT	GTTAAAACTA	AATGTAAGAA	AAATCTGCTA	3060
GAGGAAAAC	TTGAGGAACA	TTCAATGTCA	CCTGAAAGAG	AAATGGGAAA	TGAGAACATT	3120
CCAAGTACAG	TGAGCACAAT	TAGCCGTAAT	AACATTAGAG	AAAATGTTTT	TAAAGAAGCC	3180
AGCTCAAGCA	ATATTAATGA	AGTAGGTTCC	AGTACTAATG	AAGTGGGCTC	CAGTATTAAT	3240
GAAATAGGTT	CCAGTGATGA	AAACATTCAA	GCAGAACTAG	GTAAGAACAG	AGGGCCAAAA	3300
TTGAATGCTA	TGCTTAGATT	AGGGGTTTTG	CAACCTGAGG	TCTATAAACA	AAGTCTTCCT	3360
GGAAGTAATT	GTAAGCATCC	TGAAATAAAA	AAGCAAGAAT	ATGAAGAAGT	AGTTCAGACT	3420
GTTAATACAG	ATTTCTCTCC	ATCTCTGATT	TCAGATAACT	TAGAACAGCC	TATGGGAAGT	3480
AGTCATGCAT	CTCAGGTTTG	TTCTGAGACA	CCTGATGACC	TGTTAGATGA	TGGTGAAATA	3540

AAGGAAGATA CTAGTTTTGC TGAAAATGAC ATTAAGGAAA GTTCTGCTGT TTTTAGCAAA	3600
AGCGTCCAGA AAGGAGAGCT TAGCAGGAGT CCTAGCCCTT TCACCCATAC ACATTTGGCT	3660
CAGGGTTACC GAAGAGGGGC CAAGAAATTA GAGTCCTCAG AAGAGAACTT ATCTAGTGAG	3720
GATGAAGAGC TTCCCTGCTT CCAACACTTG TTATTTGGTA AAGTAAACAA TATACCTTCT	3780
CAGTCTACTA GGCATAGCAC CGTTGCTACC GAGTGTCTGT CTAAGAACAC AGAGGAGAAT	3840
TTATTATCAT TGAAGAATAG CTTAAATGAC TGCAGTAACC AGGTAATATT GGCAAAGGCA	3900
TCTCAGGAAC ATCACCTTAG TGAGGAAACA AAATGTTCTG CTAGCTTGTT TTCTTCACAG	3960
TGCAGTGAAT TGGAAGACTT GACTGCAAT ACAAACACCC AGGCTCCTTT CTTGATTGGT	4020
TCTTCCAAAC AAATGAGGCA TCAGTCTGAA AGCCAGGGAG TTGGTCTGAG TGACAAGGAA	4080
TTGGTTTCAG ATGATGAAGA AAGAGGAACG GGCTTGGAAG AAAATAATCA AGAAGAGCAA	4140
AGCATGGATT CAAACTTAGG TGAAGCAGCA TCTGGGTGTG AGAGTGAAAC AAGCGTCTCT	4200
GAAGACTGCT CAGGGCTATC CTCTCAGAGT GACATTTTAA CCACTCAGCA GAGGGATACC	4260
ATGCAACATA ACCTGATAAA GCTCCAGCAG GAAATGGCTG AACTAGAAGC TGTGTTAGAA	4320
CAGCATGGGA GCCAGCCTTC TAACAGCTAC CTTTCCATCA TAAGTGACTC TTCTGCCCTT	4380
GAGGACCTGC GAAATCCAGA ACAAAGCACA TCAGAAAAAG CAGTATTAAC TTCACAGAAA	4440
AGTAGTGAAT ACCCTATAAG CCAGAATCCA GAAGGCCTNT CTGCTGACAA GTTTGAGGTG	4500
TCTGCAGATA GTTCTACCAG TAAAAATAAA GAACCAGGAG TGGAAAGGTC ATCCCCTTCT	4560
AAATGCCCAT CATTAGATGA TAGGTGGTAC ATGCACAGTT GCTCTGGGAG TCTTCAGAAT	4620
AGAAACTACC CACCTCAAGA GGAGCTCATT AAGGTTGTTG ATGTGGAGGA GCAACAGCTG	4680
GAAGAGTCTG GGCCACACGA TTTGACGGAA ACATCTTACT TGCCAAGGCA AGATCTAGAG	4740
GGAACCCCTT ACCTGGAATC TGGAAATCAGC CTCTTCTCTG ATGACCCTGA ATCTGATCCT	4800
TCTGAAGACA GAGCCCCAGA GTCAGCTCGT GTTGGCAACA TACCATCTTC AACCTCTGCA	4860
TTGAAAGTTC CCCAATTGAA AGTTGCAGAA TCTGCCCAGA GTCCAGCTGC TGCTCATACT	4920
ACTGATACTG CTGGGTATAA TGCAATGGAA GAAAGTGTGA GCAGGGAGAA GCCAGAATTG	4980
ACAGCTTCAA CAGAAAGGGT CAACAAAAGA ATGTCCATGG TGGTGTCTGG CCTGACCCCA	5040
GAAGAATTTA TGCTCGTGTA CAAGTTTGCC AGAAAACACC ACATCACTTT AACTAATCTA	5100
ATTACTGAAG AGACTACTCA TGTGTATTATG AAAACAGATG CTGAGTTTGT GTGTGAACGG	5160
ACACTGAAAT ATTTTCTAGG AATTGCGGGA GGAAAATGGG TAGTTAGCTA TTTCTGGGTG	5220
ACCCAGTCTA TTAAAGAAAAG AAAAATGCTG AATGAGCATG ATTTTGAAGT CAGAGGAGAT	5280
GTGGTCAATG GAAGAAACCA CCAAGGTCCA AAGCGAGCAA GAGAATCCCA GGACAGAAAG	5340
ATCTTCAGGG GGCTAGAAAT CTGTTGCTAT GGGCCCTTCA CCAACATGCC CACAGATCAA	5400
CTGGAATGGA TGGTACAGCT GTGTGGTGCT TCTGTGGTGA AGGAGCTTTC ATCATTCACC	5460
CTTGGCACAG GTGTCCACCC AATTGTGGTT GTGCAGCCAG ATGCCTGGAC AGAGGACAAT	5520
GGCTTCCATG CAATTGGGCA GATGTGTGAG GCGCCTGTGG TGACCCGAGA GTGGGTGTTG	5580

GACAGTGTAG CACTCTACCA GTGCCAGGAG CTGGACACCT ACCTGATACC CCAGATCCCC 5640
CACAGCCACT ACTGACTGCA GCCAGCCACA GGTACAGAGC CACAGGACCC CAAGAATGAG 5700
CTTACAAAGT GGCCTTTCCA GGCCCTGGGA GCTCCTCTCA CTCTTCAGTC CTTCTACTGT 5760
CCTGGCTACT AAATATTTTA TGTACATCAG CCTGAAAAGG ACTTCTGGCT ATGCAAGGGT 5820
CCCTTAAAGA TTTTCTGCTT GAAGTCTCCC TTGGAAA 5857

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1863 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(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
 Ile 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 Asp Arg Val His
 435 440 445
 Ser Lys Ser Val Glu Ser Asn Ile Glu Gly Gln 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 Xaa Ala Phe Val Ser 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

-100-

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 Ser 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 Ala 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 Xaa
 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 Pro 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:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: s754 A

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

20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: s754 B

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

20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: s975 A

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

20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: s975 B

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

AGGCAACTTT GCAATGAGTG

20

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: tdj1474 A

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

CAGAGTGAGA CCTTGCTCTCA AA

22

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: tdj1474 B

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

TTCTGCAAAC ACCTTAACT CAG

23

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: tdj1239 A

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

AACCTGGAAG GCAGAGGTTG

20

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE:
 - (B) CLONE: tdj1239 B

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

TCTGTACCTG CTAAGCAGTG G

21

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 111 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

GGKCTTACTC TGTTGTCCCA GCTGGAGTAC AGWGTGCGAT CATGAGGCTT ACTGTTGCTT	60
GACTCCTAGG CTCAAGCGAT CCTATCACCT CAGTCTCCAA GTAGCTGGAC T	111

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

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

Xaa	Leu	Leu	Cys	Cys	Pro	Ser	Trp	Ser	Thr	Xaa	Cys	Asp	His	Glu	Ala
1			5					10						15	
Tyr	Cys	Cys	Leu	Thr	Pro	Arg	Leu	Lys	Arg	Ser	Tyr	His	Leu	Ser	Leu
			20				25					30			
Gln	Val	Ala	Gly	Leu											
			35												

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1540 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

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

GAGGCTAGAG GGCAGGCACT TTATGGCAAA CTCAGGTAGA ATTCTTCCTC TTCCGTCTCT	60
TTCTTTTAC GTCATCCGGG GGCAGACTGG GTGGCCAATC CAGAGCCCCG AGAGACGCTT	120
GGCTCTTCT GTCCCTCCCA TCCTCTGATT GTACCTTGAT TTCGTATTCT GAGAGGCTGC	180
TGCTTAGCGG TAGCCCCCTG GTTCCCGTGG CAACGGAAAA GCGCGGGAAT TACAGATAAA	240
TTAAAACTGC GACTGCGCGG CGTGAGCTCG CTGAGACTTC CTGGACGGGG GACAGGCTGT	300

GGGGTTTCTC AGATAACTGG GCCCCTGCGC TCAGGAGGCC TTCACCCTCT GCTCTGGGTA	360
AAGGTAGTAG AGTCCCGGGA AAGGGACAGG GGGCCCAAGT GATGCTCTGG GGTACTGGCG	420
TGGGAGAGTG GATTTCCGAA GCTGACAGAT GGGTATTCTT TGACGGGGGG TAGGGGCGGA	480
ACCTGAGAGG CGTAAGGCGT TGTGAACCCT GGGGAGGGGG GCAGTTTGTA GGTGCGGAGG	540
GAAGCGCTGA GGATCAGGAA GGGGGCACTG AGTGTCCGTG GGGGAATCCT CGTGATAGGA	600
ACTGGAATAT GCCTTGAGGG GGACACTATG TCTTTAAAAA CGTCGGCTGG TCATGAGGTC	660
AGGAGTTCCA GACCAGCCTG ACCAACGTGG TGAAACTCCG TCTCTACTAA AAATACNAAA	720
ATTAGCCGGG CGTGGTGCCG CTCCAGCTAC TCAGGAGGCT GAGGCAGGAG AATCGCTAGA	780
ACCCGGGAGG CGGAGGTTGC AGTGAGCCGA GATCGCGCCA TTGCACTCCA GCTTGGGCGA	840
CAGAGCGAGA CTGTCTCAAA AAAAAACAAA AAAAAACAAA AAAAAAACA CCGGCTGGTA	900
TGTATGAGAG GATGGGACCT TGTGGAAGAA GAGGTGCCAG GAATATGTCT GGGGAGGGGA	960
GGAGACAGGA TTTTGTGGGA GGGAGAACTT AAGAACTGGA TCCATTTGCG CCATTGAGAA	1020
AGCGCAAGAG GGAAGTAGAG GAGCGTCAGT AGTAACAGAT GCTGCCGGCA GGGATGTGCT	1080
TGAGGAGGAT CCAGAGATGA GAGCAGGTCA CTGGGAAAAGG TTAGGGGCGG GGAGGCCTTG	1140
ATTGGTGTTG GTTGGTTCGT TGTGATTTT GGTTTTATGC AAGAAAAAGA AAACAACCAG	1200
AAACATTGGR GAAAGCTAAG GCTACCACCA CCTACCCGGT CAGTCACTCC TCTGTAGCTT	1260
TCTCTTYCTT GGAGAAAGGA AAAGACCCAA GGGGTTGGCA GCRATATGTG AAAAAATTCA	1320
GAATTTATGT TGTCTAATTA CAAAAAGCAA CTTCTAGAAT CTTTAAAAAT AAAGGACGTT	1380
GTCATTAGTT CTTTGGTTTG TATYATTCTA AAACCTTCCA AATCTTAAAT TTACTIONATT	1440
TTAAATGAT AAAATGAAGT TGTCATTTTA TAAACCTTTT AAAAAGATAT ATATATATGT	1500
TTTCTAATG TGTANAGTT CATTGGAACA GAAAGAAATG	1540

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21715 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

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

CCTTTTACGT CATCGGGGAG ACTGGGTGGC AATCGCAGCC CGAGAGACGC ATGGCTCTTT	60
CTGCCCTCCA TCCTCTGATG TACCTTGATT TCGTATTCTG AGAGGCTGCT GCTTAGCGGT	120
AGCCCCTTGG TTTCCGTGGC AACGGAAAAG CGCGGGAATT ACAGATAAAT TAAAACTGCG	180

ACTGCGCGGC GTGAGCTCGC TGAGACTTCC TGGACCCCGC ACCAGGCTGT GGGGTTTCTC	240
AGATAACTGG GCCCCTGCGC TCAGGAGGCC TTCACCTCT GCTCTGGGTA AAGVVVVVAA	300
AAATAAAGGA CGCTGTCGAY TAGTTCCTCT GGTGTGTATT ATTCTMAAAC CTTCCAAATC	360
TTCAAATTTA CTTYATTTTA AAATGATGAA AATGAAGTTG TCATTTTATA AACCTTTTAA	420
AAAGATATAT ATATATGTTT TTCTAATGTG TTAAAGTTCA TTGGAACAGA AAGAAATGGA	480
TTTATCTGCT CTTGCGGTTG AAGAAGTACA AAATGTCATT AATGCTATGC AGAAAATCTT	540
AGAGTGTCCT ATCTGGTAAG TCAGCACAAG AGTGATTATA TTTGGGATTC CTATGATTAT	600
CTCCTATGCA AATGAACAGA ATTGACCTTA CATACTAGGG AAGAAAAGAC ATGTCTAGTA	660
AGATTAGGCT ATTGTAATTG CTGATTTTCT TAACTGAAGA ACTTTAAAAA TATAGAAAAT	720
GATTCCTTGT TCTCCATCCA CTCTGCCTCT CCCACTCCTC TCCTTTTCAA CACATCCTGT	780
GGTCCGGGAA AGCAGGGCTC TGTCTTGATT GGTCTGTCAC TGGGCAGGAT CTGTTAGATA	840
CTGCATTGTC TTTCTCCAGC TCTAAAVVVV VVVVVVVVAA ATGCTGATGA TAGTATAGAG	900
TATTGAAGGG ATCAATATAA TTCTGTTTTG ATATCTGAAA GCTCACTGAA GGTAAGGATC	960
GTATTCTCTG CTGTATTCTC AGTTCCTGAC ACAGCAGACA TTTAATAAAT ATTGAACGAA	1020
CTTGAGGCCT TATGTTGACT CAGTCATAAC AGCTCAMAGT TGAACCTTATT CACTAAGAAT	1080
AGCTTTATTT YTAAATAAAT TATTGAGCCT CATTTATTTT CTTTTTCTCV VVVVTCTGGA	1140
GTTGATCAAG GAACCTGTCT CCACAAAGTG TGACCACATA TTTTGCAAVV VVVATGTGGC	1200
TCCATTATTA GCTTTTGTTT TTGTCCTTCA TAACCCAGGA AACACCTAAC TTTATAGAAG	1260
CTTTACTTTC TTCAATTAAG TGAGAACGAA AATCCAACCTC CATTTTCATTC TTTCTCAGAG	1320
AGTATATAGT TATCAAAAGT TGGTTGTAAT CATAGTTCCT GGTAAAGTTT TGACATATAT	1380
TATCTTTTTT TTTTTTTGAG ACAAGTCTCG CTCTGTCGCC CAGGCTGGAG TGCAGTGGCA	1440
TGAGGCTTGC TCACTGCACC TCCGCCCCCG AGTTCAGCGA CTCTVVVVVV VVVVVVTGA	1500
GATCTAGACC ACATGGTCAA AGAGATAGAA TGTGAGCAAT AAATGAACCT TAAATTTTTC	1560
AACAGCTACT TTTTTTTTTT TTTTTTGAGA CAGGGKCTTA CTCTGTTGTC CCAGCTGGAG	1620
TACAGWGTGC GATCATGAGG CTTACTGTTG CTTGACTCCT AGGCTCAAGC GATCCTATCA	1680
CCTCAGTCTC CAAGTAGCTG GACTGTAAGT GCACACCACC ATATCCAGCT AAATTTTGTG	1740
TTTTCTGTAG AGACGGGGTT TCGCCATGTT TCCCAGGCTG GTCTTGAAC TTTGGCTTAA	1800
CCCGTCTGCC CACCTAGGCA TCCCAAAGTG CTAGGATTAC AGGTGTGAGT CATCATGCCT	1860
GGCCAGTATT TTAGTTAGCT CTGTCTTTTC AAGTCATATA CAAGTTCATT TTCTTTAAGT	1920
TAGTTAACAA CCTTATATCA TGTATTCTTT TCTAGCAVVV VVVVVVVVVV TGTGATCATA	1980
ACAGTAAGCC ATATGCATGT AAGTTCAGTT TCATAGATCA TTGCTTATGT AGTTTAGGTT	2040
TTTGCTTATG CAGCATCCAA AAACAATTAG GAAACTATTG CTTGTAATTC ACCTGCCATT	2100
ACTTTTTTAA TGGCTCTTAA GGGCAGTTGT GAGATTATCT TTTTCATGGCT ATTTGCCTTT	2160
TGAGTATTCT TTCTACAAA GGAAGTAAAT TAAATTGTTC TTTCTTTCTT TATAATTTAT	2220

AGATVVVVVA	TTTTGCATGC	TGAAACTTCT	CAACCAGAAG	AAAGGGCCTT	CACAGTGTCC	2280
TTTATGTAAG	AATGATATAA	CCAAAAGGAT	ATAANNNGAA	NTGATGCTAG	GTTGGAAGCA	2340
ACCACAGTAG	GAAAAGTAGA	AATTATTTAA	TAACATAGCG	TTCCTATAAA	ACCATTTCATC	2400
AGAAAAATTT	ATAAAAGAGT	TTTTCAGACA	CAGTAAATTA	TTTCCAAAGT	TATTTTCCTG	2460
AAAGTTTTAT	GGGCATCTGC	CTTATACAGG	TNTTGVVVVV	VVVVVVVVGG	TAGGCTTAAA	2520
TGAATGACAA	AAAGTTACTA	AATCACTGCC	ATCACACGGT	TATATACAGA	TGTCAATGAT	2580
GTATTGATTT	AGAGGTTTTT	TACTGTTGCT	GCATCTTATT	TTTATTTGTT	TACATGTCTT	2640
TTCTTATTTT	AGTGTCTTAA	AAAGGTTGAT	AATCACTTGC	TGAGTGTGTT	TCTCAA VVV	2700
VVGAGCCTAC	AAGAAAGTAC	GAGATTTAGT	CAACTTGTTG	AAGAGCTATT	GAAAATCATT	2760
TGTGCTTTTC	AGCTTGACAC	AGGTTTGGAG	TATGCAAACA	GCTATAATTT	TGCAAAAAAG	2820
GAAAATAACT	CTCCTGAACA	TCTAAAAGAT	GAAGTTTCTA	TCATCCAAAG	TATGGGCTAC	2880
AGAAACCGTG	CCAAAAGACT	TCTACAGAGT	GAACCCGAAA	ATCCTTCCTT	GCAGAAACCA	2940
TTTGTTTTCT	TCTTCTCTTT	CTTCTTCTTT	TCTTTTTTTT	TTCTTTTTTT	TTTCWTGAGA	3000
GTGGAGTCYT	TGGCTCGTSS	YGGCGCCGAG	GCTAGGAAGC	AGCTCCTCCT	GCGCTCGTAC	3060
GCCACGCTAG	TAGGTAGCGT	GGGCAKACAC	GGCACGCGCA	CCATGGCCAG	GCTAATTTTT	3120
GTATTGTTVV	VVVVVVVVVV	VATGATGGAG	ATCTTAAAAA	GTAATCATTC	TGGGGCTGGG	3180
CGTAGTAGCT	TGCACCTGTA	ATCCCAGCAC	TTCCGGGAGGC	TGAGGCAGGC	AGATAATTTG	3240
AGGTCAGGAG	TTTGAGACCA	GCCTGGCCAA	CATGGTGAAA	CCCATCTCTA	CTAAAAATAC	3300
AAAAATTAGC	TGGGTGTGGT	GGCACGTACC	TGTAATCCCA	GCTACTCGGG	AGGCGGAGGC	3360
ACAAGAATTG	CTTGAACCTA	GGACGCGGAG	GTTGCAGCGA	GCCAAGATCG	CGCCACTGCA	3420
CTCCAGCCTG	GGCCGTAGAG	TGAGACTCTG	TCTCAAAAAA	GAAAAAAAAG	TAATTGTTCT	3480
AGCTGGGCGC	AGTGGCTCTT	GCCTGTAATC	CCAGCACTTT	GGGAGGCCAA	GGCGGGTGGA	3540
TCTCGAGTCC	TAGAGTTCAA	GACCAGCCTA	GGCAATGTGG	TGAAACCCCA	TCGCTACAAA	3600
AAATACAAAA	ATTAGCCAGG	CATGGTGGCG	TGCGCATGTA	GTCCCAGCTC	CTTGGGAGGC	3660
TGAGGTGGGA	GGATCACTTG	AACCCAGGAG	ACAGAGGTTG	CAGTGAACCG	AGATCACGCC	3720
ACCACGCTCC	AGCCTGGGCA	ACAGAACAAG	ACTCTGTCTA	AAAAAATACA	AATAAAATAA	3780
AAGTAGTTCT	CACAGTACCA	GCATTCATTT	TTCAAAAAGAT	ATAGAGCTAA	AAAGGAAGGA	3840
AAAAAAAAGT	AATGTTGGGC	TTTAAATAC	TCGTTCCCTAT	ACTAAATGTT	CTTAGGAGTG	3900
CTGGGGTTTT	ATTGTCATCA	TTTATCCTTT	TTAAAAATGT	TATTGGCCAG	GCACGGTGGC	3960
TCATGGCTGT	AATCCCAGCA	CTTTGGGAGG	CCGAGGCAGG	CAGATCACCT	GAGGTCAGGA	4020
GTGTGAGACC	AGCCTGGCCA	ACATGGCGAA	ACCTGTCTCT	ACTAAAAATA	CAAAAATTAA	4080
CTAGGCGTGG	TGGTGTACGC	CTGTAGTCCC	AGCTACTCGG	GAGGCTGAGG	CAGGAGAATC	4140
AACTGAACCA	GGGAGGTGGA	GTTTGCAGTG	TGCCGAGATC	ACGCCACTGC	ACTCTAGCCT	4200
GGCAACAGAG	CAAGATTCTG	TCTCAAAAAA	AAAAAACATA	TATACACATA	TATCCCAAAG	4260

TGCTGGGATT ACATATATAT ATATATATAT ATATTATATA TATATATATA TATATATGTG	4320
ATATATATGT GATATATATA TAACATATAT ATATGTAATA TATATGTGAT ATATATATAA	4380
TATATATATG TAATATATAT GTGATATATA TATATACACA CACACACACA TATATATGTA	4440
TGTGTGTGTA CACACACACA CACAAATTAG CCAGGCATAG TTGCACACGC TTGGTAGACC	4500
CAGCTACTCA GGAGGCTGAG GGAGGAGAAT CTCTTGAAC TTAGGAGGCGG AGGTTGCAGT	4560
GAGCTGAGAT TGCGCCACTG CACTCCAGCC TGGGTGACAG AGCAGGACTC TGTACACCCC	4620
CCAAAACAAA AAAAAAAGTT ATCAGATGTG ATTGGAATGT ATATCAAGTA TCAGCTTCAA	4680
AATATGCTAT ATTAATACTT CAAAAATTAC ACAAATAATA CATAATCAGG TTTGAAAAAT	4740
TTAAGACAAC MSAARAAAA AWYCMAATCA CAMATATCCC ACACATTTTA TTATTMCTMC	4800
TMCWATTATT TTGWAGAGMC TGGGTCTCAC YCYKTTGCTW ATGCTGGTCT TTGAACYCCY	4860
KGCCYCAARC ARTCCTSCCT CABCCCTCCA ARGTGCTGGG GATWATAGGC ATGARCTACC	4920
CGCTCCCAGC CCMAGACAAT TTTAGTGTGW GCAAAATTCC TGGGCATTTT TTCMAGGCHD	4980
NMWAYAYVTW DSMTGCWHRT GAHBGGCTGA TGWGTCAATT TATTTTGYCC ATGGTGTCAA	5040
KTYTCTCTTC AGGAGGAAAA GCACAMGAAC TGGCCAAYAA TTGCTTGACT GTTCTTTACC	5100
AWACTGTTWA RCAGGAAACC ARTCTCAGTG TCCCAACTCT CTAACCTTGG AACTTGTGAG	5160
AACTCTGAGG ACAAAGCAGC GGATACAACC TCAAAAGACG TCTGTCTACA TTGAATTGGG	5220
TAAGGGTCTC AGGTTTTTTA AGTATTTAAT AATAATTGCT GGATTCCTTA TCTTATAGTT	5280
TTGCCAMAAA TCTTGGTCAT AATTTGTATT TGTGGTAGGC AGCWTGCGGA AGTGAATTTT	5340
ATGAGCCCTA TGGTGAGTTA TAAAAAATGT AAAAGACGCA GTTCCCACCT TGAAGAATCT	5400
TACTTTAAAA AGGGAGCAAA AGAGGCCAGG CATGGTGGCT CACACCTGTA ATCCCAGCAC	5460
TTTGGGAGGC CAAAGTGGGT GGATCACCTG AGGTCGGGAG TTCGAGACCA GCCTAGCCAA	5520
CATGGAGAAA CTCTGTCTGT ACCAAAAAAT AAAAAATTAG CCAGGTGTGG TGGCACATAA	5580
CTGTAATCCC AGCTACTCGG GAGGCTGAGG CAGGAGAATC ACTTGAACCC GGGAGGTGGA	5640
GGTTGCGGTG AACCGAGATC GCACCATTGC ACTCCAGCCT GGGCAAAAAT AGCGAAACTC	5700
CATCTAAAAA AAAAAAAGAG AGCAAAAGAA ASAMTMTCTG GTTTTAAMTM TGTGTAAATA	5760
TGTTTTTGGA AAGATGGAGA GTAGCAATAA GAAAAACAT GATGGATTGC TACAGTATTT	5820
AGTTCCAAGA TAAATTGTAC TAGATGAGGA AGCCTTTTAA GAAGAGCTGA ATTGCCAGGC	5880
GCAGTGGCTC ACGCCTGTAA TCCCAGCACT TTGGGAGGCC GAGGTGGGCG GATCACCTGA	5940
GGTCGGGAGT TCAAGACCAG CCTGACCAAC ATGGAGAAAC CCCATCTCTA CTAAAAAAA	6000
AAAAAATAAA ATTAGCCGGG GTGGTGGCTT ATGCCTGTAA TCCCAGCTAC TCAGGAGGCT	6060
GAGGCAGGAG AATCGCTTGA ACCCAGGAAG CAGAGGTTGC AGTGAGCCAA GATCGCACCA	6120
TTGCACTCCA GCCTAGGCAA CAAGAGTGAA ACTCCATCTC AAAAAAATA AAAAAAGAGCT	6180
GAATCTTGGC TGGGCAGGAT GGCTCGTGCC TGTAAATCCTA ACGCTTTGGA AGACCGAGGC	6240
AGAAGGATTG GTTGAGTCCA CGAGTTTAA ACCAGCCTGG CCAACATAGG GGAACCCTGT	6300

CTCTATTTTT AAAATAATAA TACATTTTTG GCCGGTGC GG TGGCTCATGC CTGTAATCCC	6360
AATACTTTGG GAGGCTGAGG CAGGTAGATC ACCTGAGGTC AGAGTTCGAG ACCAGCCTGG	6420
ATAACCTGGT GAAACCCCTC TTTACTAAAA ATACAAAAAA AAAAAAAAAAT TAGCTGGGTG	6480
TGGTAGCACA TGCTTGTAAT CCCAGCTACT TGGGAGGCTG AGGCAGGAGA ATCGCTTGAA	6540
CCAGGGAGGC GGAGGTTACA ATGAGCCAAC ACTACACCAC TGCCTCCAG CCTGGGCAAT	6600
AGAGTGAGAC TGCATCTCAA AAAATAATA ATTTTTAAAA ATAATAAATT TTTTAAAGCT	6660
TATAAAAAGA AAAGTTGAGG CCAGCATAGT AGCTCACATC TGTAATCTCA GCAGTGGCAG	6720
AGGATTGCTT GAAGCCAGGA GTTTGAGACC AGCCTGGGCA ACATAGCAAG ACCTCATCTC	6780
TACAAAAAAA TTTCTTTTTT AAATTAGCTG GGTGTGGTGG TGTGCATCTG TAGTCCCAGC	6840
TACTCAGGAG GCAGAGGTGA GTGGATACAT TGAACCCAGG AGTTTGAGGC TGAGTGAGC	6900
TATGATCATG CCACTGCACT CCAACCTGGG TGACAGAGCA AGACCTCCAA AAAAAAAAAA	6960
AAAAGAGCTG CTGAGCTCAG AATTCAAAC TGGCTCTCAA ATTGGATTTT CTTT TAGAAT	7020
ATATTTATAA TTA AAAAGGA TAGCCATCTT TTGAGCTCCC AGGCACCACC ATCTATTTAT	7080
CATAACACTT ACTGTTTTCC CCCCTTATGA TCATAAATTC CTAGACAACA GGCATTGTAA	7140
AAATAGTTAT AGTAGTTGAT ATTTAGGAGC ACTTAACTAT ATTCCAGGCA CTATTGTGCT	7200
TTTCTTGAT AACTCATTAG ATGCTTGTC GACCTCTGAG ATTGTTCTTA TTATACTTAT	7260
TTTACAGATG AGAAAATTAA GGCACAGAGA AGTTATGAAA TTTTCCAAG GTATTAAACC	7320
TAGTAAGTGG CTGAGCCATG ATTCAAACCT AGGAAGTTAG ATGTCAGAGC CTGTGCTTTT	7380
TTTTTGTTTT TGTTTTTGTT TTCAGTAGAA ACGGGGGTCT CACTTTGTTG GCCAGGCTGG	7440
TCTTGAAC TC AACCTCAA ATAATCCACC CATCTCGGCC TCCTCAAGTG CTGGGATTAC	7500
AGGTGAGAGC CACTGTGCCT GGCGAAGCCC ATGCCTTTAA CCACTTCTCT GTATTACATA	7560
CTAGCTTAAC TAGCATTGTA CCTGCCACAG TAGATGCTCA GTAAATATTT CTAGTTGAAT	7620
ATCTGTTTTT CAACAAGTAC ATTTTTTTAA CCCTTTTAAT TAAGAAAAC TTTATTGATT	7680
TATTTTTTGG GGGGAAATTT TTTAGGATCT GATTCTTCTG AAGATACCGT TAATAAGGCA	7740
ACTTATTGCA GGTGAGTCAA AGAGAACCTT TGTCTATGAA GCTGGTATTT TCCTATTTAG	7800
TTAATATTAA GGATTGATGT TTCTCTCTTT TTAAAAATAT TTTAACTTTT ATTTTAGGTT	7860
CAGGGATGTA TGTGCAGTTT GTTATATAGG TAAACACACG ACTTGGGATT TGGTGTATAG	7920
ATTTTTTTCA TCATCCGGGT ACTAAGCATA CCCACAGTT TTTTGTTTGC TTTCTTTCTG	7980
AATTTCTCCC TCTTCCCACC TTCCTCCCTC AAGTAGGCTG GTGTTTCTCC AGACTAGAAT	8040
CATGGTATTG GAAGAAACCT TAGAGATCAT CTAGTTTAGT TCTCTCATTT TATAGTGGAG	8100
GAAATACCCT TTTTGTTTGT TGGATTTAGT TATTAGCACT GTCCAAAGGA ATTTAGGATA	8160
ACAGTAGAAC TCTGCACATG CTTGCTTCTA GCAGATTGTT CTCTAAGTTC CTCATATACA	8220
GTAATATTGA CACAGCAGTA ATTGTGACTG ATGAAAATGT TCAAGGACTT CATTTTCAAC	8280
TCTTTCTTTC CTCTGTTCTT TATTCCACA TATCTCTCAA GCTTTGTCTG TATGTTATAT	8340

AATAAACTAC	AAGCAACCCC	AACTATGTTA	CCTACCTTCC	TTAGGAATTA	TTGCTTGACC	8400
CAGGTTTTTT	TTTTTTTTTT	TTTGAGACG	GGGTCTTGCC	CTGTTGCCAG	GATGGAGTGT	8460
AGTGGCGCCA	TCTCGGCTCA	CTGCAATCTC	CAACTCCCTG	GTTCAAGCGA	TTCTCCTGTC	8520
TCAATCTCAC	GAGTAGCTGG	GACTACAGGT	ATACACCACC	ACGCCCCGTT	AATTGACCAT	8580
TCCATTTCTT	TCTTTCTCTC	TTTTTTTTTT	TTTTTTTTGA	GACAGAGTCT	TGCTCTGTTG	8640
CCCAGGCTGG	AGTACAGAGG	TGTGATCTCA	CCTCTCCGCA	ACGTCTGCCT	CCCAGGTTGA	8700
AGCCATACTC	CTGCCTCAGC	CTCTCTAGTA	GCTGGGACTA	CAGGCGCGCG	CCACCACACC	8760
CGGCTAATTT	TTGTATTTTT	AGTAGAGATG	GGGTTTCACC	ATGTTGGCCA	GGCTGGTCTT	8820
GAACTCATGA	CCTCAAGTGG	TCCACCCGCC	TCAGCCTCCC	AAAGTGCTGG	AATTACAGGC	8880
TTGAGCCACC	GTGCCCAGCA	ACCATTTTCAT	TTCAACTAGA	AGTTTCTAAA	GGAGAGAGCA	8940
GCTTTCACCTA	ACTAAATAAG	ATTGGTCAGC	TTTCTGTAAT	CGAAAGAGCT	AAAATGTTTG	9000
ATCTTGGTCA	TTTGACAGTT	CTGCATACAT	GTAAGTAGTG	TTTCTTATTA	GGACTCTGTC	9060
TTTTCCCTAT	AGTGTGGGAG	ATCAAGAATT	GTTACAAATC	ACCCCTCAAG	GAACCAGGGA	9120
TGAAATCAGT	TTGGATTCTG	CAAAAAAGGG	TAATGGCAAA	GTTTGCCAAC	TTAACAGGCA	9180
CTGAAAAGAG	AGTGGGTAGA	TACAGTACTG	TAATTAGATT	ATTCTGAAGA	CCATTTGGGA	9240
CCTTTACAAC	CCACAAAATC	TCTTGGCAGA	GTTAGAGTAT	CATTCTCTGT	CAAATGTCGT	9300
GGTATGGTCT	GATAGATTTA	AATGGTACTA	GACTAATGTA	CCTATAATAA	GACCTTCTTG	9360
TAAGTGATTG	TTGCCCTTTC	GCTTTTTTTT	TTGTTTGTTT	GTTTGTTTTT	TTTTGAGATG	9420
GGGTCTCACT	CTGTTGCCCA	GGCTGGAGTG	CAGTGATGCA	ATCTTGGCTC	ACTGCAACCT	9480
CCACCTCCAA	AGGCTCAAGC	TATCCTCCCA	CTTCAGCCTC	CTGAGTAGCT	GGGACTACAG	9540
GCGCATGCCA	CCACACCCGG	TTAATTTTTT	GTGGTTTTAT	AGAGATGGGG	TTTCACCATG	9600
TTACCGAGGC	TGGTCTCAAA	CTCCTGGACT	CAAGCAGTCT	GCCCCATTCA	GCCTCCCAAA	9660
GTGCTGCAGT	TACAGGCTTG	AGCCACTGTG	CCTGGCCTGC	CCTTTACTTT	TAATTGGTGT	9720
ATTTGTGTTT	CATCTTTTAC	CTACTGGTTT	TTAAATATAG	GGAGTGGTAA	GTCTGTAGAT	9780
AGAACAGAGT	ATTAAGTAGA	CTTAATGGCC	AGTAATCTTT	AGAGTACATC	AGAACCAGTT	9840
TTCTGATGGC	CAATCTGCTT	TTAATTCACT	CTTAGACGTT	AGAGAAATAG	GTGTGGTTTT	9900
TGCATAGGGA	AAATTCTGAA	ATTAAVVVVV	VVVVVVVVGA	TCCTAAGTGG	AAATAATCTA	9960
GGTAAATAGG	AATTAAATGA	AAGAGTATGA	GCTACATCTT	CAGTATACTT	GGTAGTTTAT	10020
GAGGTTAGTT	TCTCTAATAT	AGCCAGTTGG	TTGATTTCCA	CCTCCAAGGT	GTATGAAGTA	10080
TGTATTTTTT	TAATGACAAT	TCAGTTTKTG	AGTACCTTGT	TATTTTTGTA	TATTTTCAGC	10140
TGCTTGTGAA	TTTTCTGAGA	CGGATGTAAC	AAATACTGAA	CATCATCAAC	CCAGTAATAA	10200
TGATTTGAAC	ACCACTGAGA	AGCGTGCAGC	TGAGAGGCAT	CCAGAAAAGT	ATCAGGGTAG	10260
TTCTGTTTCA	AACTTGTCATG	TGGAGCCATG	TGGCACAAAT	TACTCATGCC	AGCTCATTAC	10320
AGCATGAGAA	CAGCAGTTTA	TTACTCACTA	AAGACAGAAT	GAATGTAGAA	AAGGCTGAAT	10380

TCTGTAATAA	AAGCAAACAG	CCTGGCTTAG	CAAGGAGCCA	ACATAACAGA	TGGGCTGGAA	10440
GTAAGGAAAC	ATGTAATGAT	AGGCGGACTC	CCAGCACAGA	AAAAAAGGTA	GATCTGAATG	10500
CTGATCCCCT	GTGTGAGAGA	AAAGAATGGA	ATAAGCAGAA	ACTGCCATGC	TCAGAGAATC	10560
CTAGAGATAC	TGAAGATGTT	CCTTGGATAA	CACTAAATAG	CAGCATTGAG	AAAGTTAATG	10620
AGTGGTTTTT	CAGAAGTGAT	GAAGTGTAG	GTTCTGATGA	CTCACATGAT	GGGGAGTCTG	10680
AATCAAATGC	CAAAGTAGCT	GATGTATTGG	ACGTCCTAAT	TGAGGTAGAT	GAATATTCTG	10740
GTTCTTCAGA	GAAAATAGAC	TTACTGGCCA	GTGATCCTCA	TGAGGCTTTA	ATATGTAAAA	10800
GTGACAGAGT	TCAGTCCAAA	TCAGTAGAGA	GTAATATTGA	AGGCCAAATA	TTTGGGAAAA	10860
CCTATCGGAA	GAAGGCAAGC	CTCCCCAACT	TAAGCCATGT	AACTGAAAAAT	CTAATTATAG	10920
GNGCATTTGT	TTCTGAGCCA	CAGATAATAC	AAGAGCGTCC	CCTCACAAAT	AAATTAAAGC	10980
GTAAAAGGAG	ACCTACATCA	GGCCTTCATC	CTGAGGATTT	TATCAAGAAA	GCAGATTTGG	11040
CAGTTCAAAA	GACTCCTGAA	ATGATAAATC	AGGGAAGTAA	CCAAACGGAG	CAGAATGGTC	11100
AAGTGATGAA	TATTACTAAT	AGTGGTCATG	AGAATAAAAC	AAAAGGTGAT	TCTATTGAGA	11160
ATGAGAAAAA	TCCTAACCCA	ATAGAATCAC	TCGAAAAAGA	ATCTGCTTTC	AAAACGAAAG	11220
CTGAACCTAT	AAGCAGCAGT	ATAAGCAATA	TGGAAGTCGA	ATTAAATATC	CACAATTCAA	11280
AAGCACCTAA	AAAGAATAGG	CTGAGGAGGA	AGTCTTCTAC	CAGGCATATT	CATGCGCTTG	11340
AACTAGTAGT	CAGTAGAAAT	CTAAGCCAC	CTAATTGTAC	TGAATTGCAA	ATTGATAGTT	11400
GTTCTAGCAG	TGAAGAGATA	AAGAAAAAAA	AGTACAACCA	AATGCCAGTC	AGGCACAGCA	11460
GAAACCTACA	ACTCATGGAA	GGTAAAGAAC	CTGCAACTGG	AGCCAAGAAG	AGTAACAAGC	11520
CAAATGAACA	GACAAGTAAA	AGACATGACA	GTGATACTTT	CCCAGAGCTG	AAGTTAACAA	11580
ATGCACCTGG	TTCTTTTACT	AAGTGTTCAA	ATACCAGTGA	ACTTAAAGAA	TTTGTCAATC	11640
CTAGCCTTCC	AAGAGAAGAA	AAAGAAGAGA	AACTAGAAAC	AGTTAAAGTG	TCTAATAATG	11700
CTGAAGACCC	CAAAGATCTC	ATGTTAAGTG	GAGAAAGGGT	TTTGCAAAC	GAAAGATCTG	11760
TAGAGAGTAG	CAGTATTTCA	CTGGTACCTG	GTACTGATTA	TGGCACTCAG	GAAAGTATCT	11820
CGTTACTGGA	AGTTAGCACT	CTAGGGAAGG	CAAAAACAGA	ACCAAATAAA	TGTGTGAGTC	11880
AGTGTGCAGC	ATTTGAAAAC	CCCAAGGGAC	TAATTCATGG	TTGTTCCAAA	GATAATAGAA	11940
ATGACACAGA	AGGCTTTAAG	TATCCATTGG	GACATGAAGT	TAACCACAGT	CGGGAAACAA	12000
GCATAGAAAT	GGAAGAAAAGT	GAAGTTGATG	CTCAGTATTT	GCAGAATACA	TTCAAGGTTT	12060
CAAAGCGCCA	GTCATTTGCT	CTGTTTTCAA	ATCCAGGAAA	TGCAGAAGAG	GAATGTGCAA	12120
CATTCTCTGC	CCACTCTGGG	TCCTTAAAGA	AACAAAGTCC	AAAAGTCACT	TTTGAATGTG	12180
AACAAAAGGA	AGAAAATCAA	GGAAAGAATG	AGTCTAATAT	CAAGCCTGTA	CAGACAGTTA	12240
ATATCACTGC	AGGCTTTCCT	GTGGTTGGTC	AGAAAGATAA	GCCAGTTGAT	AATGCCAAAT	12300
GTAGTATCAA	AGGAGGCTCT	AGGTTTTGTC	TATCATCTCA	GTTCAGAGGC	AACGAAACTG	12360
GACTCATTAC	TCCAAATAAA	CATGGACTTT	TACAAAACCC	ATATCGTATA	CCACCACTTT	12420

TTCCCATCAA	GTCATTTGTT	AAAACATAAT	GTAAGAAAAA	TCTGCTAGAG	GAAAACTTTG	12480
AGGAACATTC	AATGTCACCT	GAAAGAGAAA	TGGGAAATGA	GAACATTCCA	AGTACAGTGA	12540
GCACAATTAG	CCGTAATAAC	ATTAGAGAAA	ATGTTTTTAA	AGAAGCCAGC	TCAAGCAATA	12600
TTAATGAAGT	AGGTTCCAGT	ACTAATGAAG	TGGGCTCCAG	TATTAATGAA	ATAGGTTCCA	12660
GTGATGAAAA	CATTCAAGCA	GAAGTAGGTA	GAAACAGAGG	GCCAAAATTG	AATGCTATGC	12720
TTAGATTAGG	GGTTTTGCAA	CCTGAGGTCT	ATAAACAAAG	TCTTCCTGGA	AGTAATTGTA	12780
AGCATCCTGA	AATAAAAAAG	CAAGAATATG	AAGAAGTAGT	TCAGACTGTT	AATACAGATT	12840
TCTCTCCATC	TCTGATTTCA	GATAACTTAG	AACAGCCTAT	GGGAAGTAGT	CATGCATCTC	12900
AGGTTTGTTT	TGAGACACCT	GATGACCTGT	TAGATGATGG	TGAAATAAAG	GAAGATACTA	12960
GTTTTGCTGA	NAATGACATT	AAGGAAAGTT	CTGCTGTTTT	TAGCAAAAGC	GTCCAGAGGG	13020
GAGAGCTTAG	CAGGNGTCCT	AGCCCTTTCA	CCCATACACA	TTTGGCTCAG	GGTTACCGAA	13080
GAGGGGCCAA	GAAATTAGAG	TCCTCAGAAG	AGAAGTTATC	TAGTGAGGAT	GAAGAGCTTC	13140
CCTGCTTCCA	ACACTTGTTA	TTTGGTAAAG	TAAACAATAT	ACCTTCTCAG	TCTACTAGGC	13200
ATAGCACCGT	TGCTACCGAG	TGTCTGTCTA	AGAACACAGA	GGAGAATTTA	TTATCATTGA	13260
AGAATAGCTT	AAATGACTGC	AGTAACCAGG	TAATATTGGC	AAAGGCATCT	CAGGAACATC	13320
ACCTTAGTGA	GGAAACAAAA	TGTTCTGCTA	GCTTGTTTTT	TTCACAGTGC	AGTGAATTGG	13380
AAGACTTGAC	TGCAAATACA	AACACCCAGG	CTCCTTTCTT	GATTGGTTCT	TCCAAACAAA	13440
TGAGGCATCA	GTCTGAAAGC	CAGGGAGTTG	GTCTGAGTGA	CAAGGAATTG	GTTTCAGATG	13500
ATGAAGAAAG	AGGAACGGGC	TTGGAAGAAA	ATAATCAAGA	AGAGCAAAGC	ATGGATTCAA	13560
ACTTAGGTAT	TGGAACCAGG	TTTGTGTGTG	TGGCCCAGTC	TATTTATAGA	AGTGAGCTAA	13620
ATGTTTATGC	TTTTGGGGAG	CACATTTTAC	AAATTTCCAA	GTATAGTTAA	AGGAACTGCT	13680
TCTTAAACTT	GANACATGTT	CCTCCTAAGG	TGGTTNNCAT	AGAAAAAAGT	CCTTCACACA	13740
GCTAGGACGT	CATCTTTGGC	TTGAATGAGC	TTTAACATCC	TANTTACTGG	TGGACTTACT	13800
TCTGGTTTCA	TTTTATTAAA	GAAATCCAGG	TGTCCCAAGC	AAGGATTTAT	CACTGGGTGA	13860
AAGAAAGTAA	TCCAGCCCCG	CANTGAGAGG	AAAGCANAGA	AGTGGAGNGN	TTGGCTTGTT	13920
NCACTGACCC	NCTTTGGNNT	TAGGGAGAGA	CTGTCCVVVV	VVVVVVVVVG	GTGTCCCCAA	13980
GCAAGGAATT	TAATCATTTT	GTGTGACATG	AAAGTANATC	CAGTCCTGCC	AATGAGAAGA	14040
AAAAGCACAG	CAAGNNNGCA	GCGTTTATAG	TCTGCTTTTA	CATCTGAACC	TCTGTTTTTG	14100
TTATTTAAGG	TAAGGTGAAG	CAGCATCTGG	GTGTGAGAGT	GAAACAAGCG	TCTCTGAAGA	14160
CTGCTCAGGG	CTATCCTCTC	AGAGTGACAT	TTTAACCACT	CAGCAGAGVV	VVVVVVVVVV	14220
VNGNGAATGT	AATCCTAATA	TTTCNCNCCN	ACTTAAAGA	ATACCACTCC	AANGGCATCN	14280
CAATACATCA	ATCAATTGGG	GAATTGGGAT	TTTCCCTCNC	TAACATCANT	GGAATAATTT	14340
CATGGCATT	ATTGCATGAA	TGTGGTTAGA	TTAAAAGGTG	TTCATGCTAG	AACTTGTAGT	14400
TCCATACTAG	GTATTTCAAT	CNTTGTTTCCT	CAAANAATTT	GTATGATATA	TTNTCATTTA	14460

ATGGAAAGCT TCTCAAAGTA TTTCATTTTC TTGGTACCAT NNATCGTTTT TGAATNAGGG	14520
ATACCATGCA ACATAACCTG ATAAAGCTCC AGCAGGAAAT GGCTGAACTA GAAGCTGTGT	14580
TAGAACAGCA TGGGAGCCAG CCTTCTAACA GCTACCCTTC CATCATAAGT GACTCTTCTG	14640
CCCTTGAGGA CCTGCGAAAT CCAGAACAAA GCACATCAGA AAAAGGTGTG TATTGTTGGC	14700
CAAACACTGA TATCTTAANC AAAATTCTTT CCTTCCCCTT TATCTCCTTC TGAAGAGTAA	14760
GGACCTAGCT CCAACATTTT ATGATCCTTG CTCAGCACAT GGGTAATTAT GGAGCCTTGG	14820
TTCTTGTCCT TGCTCACAAC TAATATACCA GTCAGAGGGA CCCAAGGCAG TCATTCATGT	14880
TGTCATCTGA GAVVVVVVV VVVVVAATT CTGCCTGATA TACTTGTTNC AAAACCAATT	14940
TGTGTATCAT AGATTGATGC TTTTGAANAA AATCAGTATT CTAACCTGAA TTATCACTAT	15000
CAGAACAAAG CAGTAAAGTA GATTTGTTTT CTCATTCCAT TTAAAGCAGT ATTAACCTCA	15060
CAGAAAAGTA GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTNTCTGC TGACAAGTTT	15120
GAGGTGTCTG CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGG AGTAAGAAAC	15180
ATCAATGTAA AGATGCTGTG GTATCTGACA TCTTTATTTA TATTGAACTC TGATTGTTAA	15240
TTTTTTTCAC CATACTTTCT CCAGTTTTTT TGCATACAGG CATTTATACA CTTTATTGTC	15300
TCTAGGATAC TTCTTTTGTT TAATCCTATA TAGGVVVVV VVVVVVGGG TAAGNTCAAG	15360
AGATATTTTG ATAGGTGATG CAGTGATNAA TTGNAAAAAT TTNCTGCCTG CTTTAAATCT	15420
TCCCCCGTTC TTTCTTCCTN CCTCCCTCCC TTCTNCCTC CCGTCCTTNC CTTTCTTTC	15480
CCTCCCTTCC NCCTTCTTTC CNTCTNCTT TCCTTTCTTT CCTGTCTACC TTTCTTTCCT	15540
TCCTCCCTTC CTTTCTTTT CTTTCTTTC TTTCTTTTC TTCTCTTCT TTCTTTCCT	15600
TTCTTTCTTG ACAGAGTCTT GCTCTGTCAC TCAGGCTGGA GTGCAGTGGC GTGATCTCGN	15660
CTCACTGCAA CCTCTGTCTC CCAGGTTCAA GCAATTTTCC TGCCTCAGCC TCCCAGTAG	15720
CTGAGATTAC AGGCGCCAGC CACCACACCC AGCTACTGAC CTGCTTTTVV VAAACAGCTG	15780
GGAGATATGG TGCCTCAGAC CATCCCCATG TTATATGTCA ACCCTGACAT ATTGGCAGGC	15840
AACATGAATC CAGACTTCTA GGCTGTCTAT CGGGCTCTTT CTTTGCCAGT CATTNCTGAT	15900
CTCTCTGACA TGAAGTGTNT CAGGTATGCT TTGGCTGCCC AGCAAGTATG AAATGTCCTN	15960
TCACAAGGGG TGGCGATGGT TTTCTCCTTC CATTTATCTT TCTAGGTCAT CCCCTTCTAA	16020
ATGCCCATCA TTAGATGATA GGTGGTACAT GCACAGTTGC TCTGGGAGTC TTCAGAATAG	16080
AAACTACCCA CCTCAAGAGG AGCTCATTAA GGTGTGTGAT GTGGAGGAGC AACAGCTGGA	16140
AGAGTCTGGG CCACACGATT TGACGGAAAC ATCTTACTTG CCAAGGCAAG ATCTAGGTAA	16200
TATTTTCATCT GCTGTATTGG ARCAMACACT YTGATTTTAC TCTGAATCCT ACATAAAGAT	16260
ATTCTGGTTA RCCAACTTTT AGATGTRCTA GTCTATCATG GACACTTTTG TTATACTTAA	16320
TTAAGCCAC TTTAGAAAA TAGCTCAAGT GTTAATCWAG GTTTACTTGW AAATTATTGA	16380
AACTGTTAAT CCATCTATAT TTTAATTAAT GGTTTAACTA ATGATTTTGA GGATGWGGGA	16440
GTCKTGGTGT ACTCTAMATG TATTATTTCA GGCCAGGCAT AGTGGCTCAC GCCTGGTAAT	16500

CCCAGTAYYC MRGAGCCCGA GGCAGGTGGA GCCAGCTGAG GTCAGGAGTT CAAGACCTGT	16560
CTTGGCCAAC ATGGGNGAAA CCCTGTCTTC TTCTTAAAAA ANACAAAAAA AATTAACTGG	16620
GTTGTGCTTA GGTGNATGCC CCGNATCCTA GTTNTTCTTG NGGGTTGAGG GAGGAGATCA	16680
CNTTGGACCC CGGAGGGGNG GGTGGGGGNG AGCAGGNCAA AACACNGACC CAGCTGGGGT	16740
GGAAGGGAAG CCCACTCNAA AAAANNNTNV VVVVVVVVVV VVAGGGAACC CCTTACCTGG	16800
AATCTGGAAT CAGCCTCTTC TCTGATGACC CTGAATCTGA TCCTTCTGAA GACAGAGCCC	16860
CAGAGTCAGC TCGTGTGGC AACATACCAT CTTCAACCTC TGCATTGAAA GTTCCCCAAT	16920
TGAAAGTTGC AGAATCTGCC CAGAGTCGAG CTGCTGCTCA TACTACTGAT ACTGCTGGGT	16980
ATAATGCAAT GGAAGAAAGT GTGAGCAGGG AGAAGCCAGA ATTGACAGCT TCAACAGAAA	17040
GGGTCAACAA AAGAATGTCC ATGGTGGTGT CTGGCCTGAC CCCAGAAGAA TTTGTGAGTG	17100
TATCCATATG TATCTCCCTA ATGACTAAGA CTTAACAACA TTCTGGAAAG AGTTTTATGT	17160
AGGTATTGTC AATTAATAAC CTAGAGGAAG AAATCTAGAA ANCAATCACA GTTCTGTGTA	17220
ATTTAATTTT GATTACTAAT TTCTGAAAAT TTAGAAYVVV VVVVVVVVVV NCCCNCCCC	17280
CNAATCTGAA ATGGGGGTAA CCCCCCCCCA ACCGANACNT GGGTNGCNTA GAGANTTTAA	17340
TGGCCCN TTC TGAGGNACAN AAGCTTAAGC CAGGNGACGT GGANCNATGN GTTGTNTNT	17400
GTTTGGTTAC CTCCAGCCTG GTGACAGANC AAGCTCTGTC TNGNAAAANT AGGNGAGGGN	17460
NAAGATTAC NNTAAATAGT TCCAGGACAC GTGTAGAACG TGCAGGATTG CTACGTAGGT	17520
AAACATATGC CATGGTGGGA TAACTAGTAT TCTGAGCTGT GTGCTAGAGG TAACTCATGA	17580
TAATGGAATA TTTGATTTAA TTTCAGATGC TCGTGACAA GTTTGCCAGA AAACACCACA	17640
TCACTTTAAC TAATCTAATT ACTGAAGAGA CTACTCATGT TGTTATGAAA ACAGATATAC	17700
CCAAGAACCT TTACAGAATA CCTTKCATCT SCTGCATAAA ACCACATGAG GCGAGGCACA	17760
GTGGCGCATG CCTGTAWTCG CARCACTTTG GGAGGCCGAG GCGGGCAGAT ACCCCGAGRT	17820
TMSGAGTCGA GCCATCCTGG CCAGCATGGT GAAACCCCGT CTCTACTANN AAATGGNAAA	17880
ATTANCTGGG TGTGGTCGCG TGCNCCTGTA GTCCCAGCTA CTCGTGAGGC TGAGGCAGGA	17940
GAATCACTTG AACCAGGGGAA ATGGAGGTTT CAGTGAGCAG AGATCATNCC CCTNCATTCC	18000
AGCCTGGCGA CAGAGCAAGG CTCCGTCNCC NAAAAAATAA AAAAAACGT GAACAAATAA	18060
GAATATTTGT TGAGCATAGC ATGGATGATA GTCTTCTAAT AGTCAATCAA TTACTTTATG	18120
AAAGACAAAT AATAGTTTTG CTGCTTCCTT ACCTCCTTTT GTTTTGGGTT AAGATTTGGA	18180
GTGTGGGCCA GGCACVVVVV VVVVVVVVGA TCTATAGCTA GCCTTGGCGT CTAGAAGATG	18240
GGTGTGAGA AGAGGGAGTG GAAAGATATT TCCTCTGGTC TTAACCTCAT ATCAGCCTCC	18300
CCTAGACTTC CAAATATCCA TACCTGCTGG TTATAATTAG TGGTGTTTTC AGCCTCTGAT	18360
TCTGTCACCA GGGGTTTTAG AATCATAAAT CCAGATTGAT CTTGGGAGTG TAAAAAACTG	18420
AGGCTCTTTA GCTTCTTAGG ACAGCACTTC CTGATTTTGT TTTCAACTTC TAATCCTTTG	18480
AGTGTKTCTC ATTCTGCAGA TGCTGAGTTT GTGTGTGAAC GGACACTGAA ATATTTTCTA	18540

GGAATTGCGG GAGGAAAATG GGTAGTTAGC TATTTCTGGK AAGKATAATA CTATWTCTCC	18600
CCYCCTCCCC TAACAMCTCA GAATTGSATT TTACACCTAA CATTTACAMC TAAGNYYYTK	18660
GCTGATGCTT GAGTCTGAGT CMCMAAAGKC CTTTAATTGT AATACTAACT ACTTTTATCT	18720
TTAATATCAC TTTGTTTACA TAAGCTGGTG ATGCTGGGAA AATGGTCTCT TTTATAACTA	18780
ATAGGACCTA ATCTGCTCCT AGCAATGTTA GCATATGAGC TAGGGATTTA TTTAATAGTC	18840
GGCAGGAATC CATGTGCARC AGNCAAACCT ATAATGTTTA AATTAAACAT CAACTCTGTC	18900
TCCAGAAGGA AACTGCTGCT ACAAGCCTTA TTAAAGGGCT GTGGCTTTAG AGGGAAGGAC	18960
CTCTCCTCTG TCATTCTTCC TGTGCTCTTT TGTGAATCGC TGACCTCTCT ATCTCCGTGA	19020
AAAGAGCACG TTCTTCTGCT GTATGTAACC TGTCTTTTCT ATGATCTCTV VVVVVVVVVV	19080
VVNAAAAACG GGGNNGGGAN TGGGCCTTAA ANCCAAAGGG CNAACTCCCC AACCATTNAA	19140
AAANTGACNG GGGATTATTA AAANCGGCGG GAAACATTTC ACNGCCCAAC TAATATTGTT	19200
AAATTAAAC CACCACCNCT GCNCCAAGGA GGGACCTTCT CCNNCAAANC TTATTAAAGG	19260
GCCTGNGCCT TAAAGAGGGA AGGACCCTCT CCTCTGTCAA TCTTCCTGTG CTCTNNTTGT	19320
GAATCGCTGA CCTCTCTATC TCCGTGAAAA AGAGCACGTT CTTCTGCTGT ATGTAACCTG	19380
TCTTTTCTAT GATCTCTTTA GGGGTGACCC AGTCTATTAA AGAAAGAAAA ATGCTGAATG	19440
AGGTAAGTAC TTGATGTTAC AAATAACCA GAGATATTCA TTCAGTCATA TAGTTAAAAA	19500
TGTATTTGCT TCCTTCCATC AATGCACCAC TTTCTTAAAC AATGCACAAA TTTTCCATGA	19560
TAATGAGGAT CATCAAGAAT TATGCAGGCC TGCCTGTGG CTCATACCTA TAATCCCAGC	19620
GCTTTGGGAG GCTGAGGCGC TTGGATCVVV VVVVVVVVVV TGCACCCCAA GACGGATGGG	19680
NNCTCCCCCN CCNCTATTTT GATTTAGAAA NTGGGTTCCT ATGNGTAAGC NGTGCGACCC	19740
CTGCCTCAAT GACTNCTCCT AATCTCCCAA ATTCTAGGAT AGNGGGTGAN CCTCCNCTG	19800
NCTGATCCNT AAAATGGACG TGTCTGCTNC NCCACTCCAT TGAAGGAGCT TCTCTTCTC	19860
TTATCCTGAT GGGTGTGNTG TTTCTTTCAG CATGATTTTG AAGTCAGAGG AGATGTGGTC	19920
AATGGAAGAA ACCACCAAGG TCCAAAGCGA GCAAGAGAAT CCCAGGACAG AAAGGTAAAG	19980
CTCCCTCCCT CAAGTTGACN NAAATCTCAC CECACCACTC TGTATTCCAC TCCCCTTTGC	20040
AGAGATGGGC CGCTTCATTT TGTAAGACTT ATTACATACA TACACAGTGC TAGATACTTT	20100
CACACAGGTT CTTTTTTCAC TCTTCCATCC CAACCACATA AATAAGTATT GTCTCTACTT	20160
TATGAATGAT AAAACTAAGA GATTTAGAGA GGCTGTGTAA TTTGGATTCC CGTCTCGGGT	20220
TCAGATCVVV VVVVVVVVVV ATCTTCAGGG GGCTAGAAAT CTGTTGCTAT GGGCCCTTCA	20280
CCAACATGCC CACAGGTAAG AGCCTCGGGA GAANCCAGAG TTCCAGCACC AGCCTTTGTC	20340
TTACATAGTG GAGTATTATA AGCAAGGTCC CACGATGGGG GTTCCCTCAGA TTGCTGAAAT	20400
GTTCTAGAGG CTATTCTATT TCTCTACCAC TCTCCAAACA AACAGCACC TAAATGTTAT	20460
CCTATGGCAA AAAAAAATA TACCTTGTC CCCTTCTCAA GAGCATGAAG GTGGTTAATA	20520
GTTAGGATTC AGTATGTTAT GTGTTGAGAT GGCCTTGAGC TGCTGTTAGT GCCVVVVVVV	20580

VVVVVVTTTG AGAGACTATC AAACCTTATA CCAAGTGGCC TTATGGAGAC TGATAACCAG	20640
AGTACATGGC ATATCAGTGG CAAATTGACT TAAAATCCAT ACCCCTACTA TTTTAAGACC	20700
ATTGTCCTTT GGAGCAGAGA GACAGACTCT CCCATTGAGA GGTCTTGCTA TAAGCCTTCA	20760
TCCGGAGAGT GTAGGGTAGA GGGCCTGGGT TAAGTATGCA GATTACTGCA GTGATTTTAG	20820
ATCCTAAATG TCCATTTTAG ATCAACTGGA ATGGATGGTA CAGCTGTGTG GTGCTTCTGT	20880
GGTGAAGGAG CTTTCATCAT TCACCCCTGG CACAGVVVVV VVVVVVVCA GAGCAAGACC	20940
CTGTCTCAAA AACAAACAAA AAAAATGATG AAGTGACAGT TCCAGTAGTC CTACTTTGAC	21000
ACTTTGAATG CTCTTTCCTT CCTGGGGANC CAGGTGTCCA CCCAATTGTG GTTGTGCAGC	21060
CAGATGCCTG GACAGAGGAC AATGGCTTCC ATGTTTATCA CTCATTACCT GGTGCTTGAG	21120
TAGCACAGTT CTTGGCACAT TTTTAAATAT TTGTTGAATG AATGGCTAAA ATGTCTTTTT	21180
GATGTTTTTA TTGTTATTTG TTTTATATTG TAAAAGTAAT ACATGAACTG TTTCCATGGG	21240
GTGGGAGTAA GATATGAATG TTCATCACVV VVVVVVVVV VCAGTAATCC TNAGAACTCA	21300
TACGACCGGG CCCCTGGAGT CGNTGNTTNG AGCCTAGTCC NGGAGAATGA ATTGACACTA	21360
ATCTCTGCTT GTGTTCTCTG TCTCCAGCAA TTGGGCAGAT GTGTGAGGCG CCTGTGGTGA	21420
CCCGAGAGTG GGTGTTGGAC AGTGTAGCAC TCTACCAGTG CCAGGAGCTG GACACCTACC	21480
TGATACCCCA GATCCCCCAC AGCCACTACT GACTGCAGCC AGCCACAGGT ACAGAGCCAC	21540
AGGACCCCAA GAATGAGCTT ACAAAGTGGC CTTTCCAGGC CCTGGGAGCT CCTCTCACTC	21600
TTCAGTCCTT CTAAGTGCCT GGCTACTAAA TATTTTATGT ACATCAGCCT GAAAAGGACT	21660
TCTGGCTATG CAAGGGTCCC TTAAAGATTT TCTGCTTGAA GTCTCCCTTG GAAAT	21715

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

GATAAATTAA AACTGCGACT GCGCGGCGTG	30
----------------------------------	----

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

AAAAATAAAG GACGCTGTCG AYTAGTTCTT

30

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

AAGATATATA TATATGTTTT TCTAATGTGT

30

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

GTAAGTCAGC ACAAGAGTGY ATTAATTTGG

30

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

TTATTGAGCC TCATTTATTT TCTTTTCTC

30

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

ATGTGGCTCC ATTATTAGCT TTTGTTTTTG

30

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

AGCTACTTTT TTTTTTTTTT TTTTGAGCAG

30

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GTAAGTGCAC ACCACATATC CAGCTAAATG

30

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTGTTCTTTC TTTCTTTATA ATTTATAGAT

30

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GATATAANNN GAANTGATGC TAGGTTGGAA

30

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
GATAATCACT TGCTGAGTGT GTTTCTCAA 30

(2) INFORMATION FOR SEQ ID NO:26:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
GTAAGTGTG AATATCCCAA GAATGAGAGT 30

(2) INFORMATION FOR SEQ ID NO:27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
AAACATAATG TTTTCCCTTG TATTTTACAG 30

(2) INFORMATION FOR SEQ ID NO:28:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

AAACCATTG TTTTCTTCTT CTTCTTCTTC

30

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTGACTGTTC TTTACCAWAC TGTTWARCAG

30

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GGTAAGGGTC TCAGGTTTTT TAAGTATTTA

30

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTGATTATTATT TTTTGGGGGG AAATTTTTTTA

30

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GGTGAGTCAA AGAGAACCTT TGTCTATGAA

30

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTCTTATTAG GACTCTGTCT TTTCCCTATA

30

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GGTAATGGCA AAGTTTGCCA ACTTAACAGG

30

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TGAGTACCTT GTTATTTTTG TATATTTTCA

30

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GTATTGGAAC CAGGTTTGTG TGTGTGGCCC

30

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

ACATCTGAAC CTCIGTTTTT GTTATTTAAG

30

- (2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: homo sapiens

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

NGNGAATGTA ATCCTAATAT TTCNCNCNA

30

- (2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

TCTTGGTACC ATNNATCGTT TTTGAATNAG

30

- (2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

30

TGTGTATTGT TGGCCAAACA CTGATATCTT

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

30

AGATTTGTTT TCTCATTCCA TTAAAGCAG

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

30

GTAAGAAACA TCAATGTAAA GATGCTGTGG

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:
CGATGGTTTT CTCCTTCCAT TTATCTTTCT 30
- (2) INFORMATION FOR SEQ ID NO:44:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:
GTAATATTTTC ATCTGCTGTA TTGGARCAMA 30
- (2) INFORMATION FOR SEQ ID NO:45:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:
TGGAAGGGAA GCCCACTCNA AAAAANNTTN 30
- (2) INFORMATION FOR SEQ ID NO:46:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GTGAGTGTAT CCATATGTAT CTCCTAATG

30

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

ATGATAATGG AATATTTGAT TTAATTTTCAG

30

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TATACCCAAG AACCTTTACA GAATACCTTK

30

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:
TAATCCTTTG AGTGTKTCTC ATTCTGCAGA

30

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:
KAAGKATAAT ACTATWTCTC CCCYCCTCCC

30

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:
TAACCTGTCT TTTCTATGAT CTCTTTAGGG

30

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GTAAGTACTT GATGTTACAA ACTAACCAGA

30

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTATCCTGAT GGGTGTGNTG TTTCTTTCAG

30

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

GTAAAGCTCC CTCCTCAAG TTGACNNAAA

30

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

TAATTTGGAT TCCCGTCTCG GGTTCAGATC

30

- (2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

GTAAGAGCCT CGGGAGAANC CAGAGTTCCA

30

- (2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens

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

GTGATTTTATG ATCCTAAATC TCCATTTTATG

30

- (2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

CAGAGCAAGA CCCTGTCTCA AAAACAAACA

30

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTGAATGCTC TTCCTTCCT GGGGANCCAG

30

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

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

TTTATCACTC ATTACCTGGT GCTTGAGTAG

30

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

-135-

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:
CTAATCTCTG CTTGTGTTCT CTGTCTCCAG

30

WHAT IS CLAIMED IS:

1. An isolated DNA consisting essentially of DNA coding for a BRCA1 polypeptide, said DNA selected from the group consisting of cDNA and genomic DNA.
2. The isolated DNA of claim 1, wherein said BRCA1 polypeptide comprises the amino acid sequence set forth in SEQ ID No:2.
3. An isolated DNA consisting essentially of DNA having at least 15 nucleotides of the DNA of claim 1.
4. The isolated DNA of claim 3 which comprises regulatory sequences.
5. A pair of single-stranded DNA primers for determination of a nucleotide sequence of a BRCA1 gene by a polymerase chain reaction, the sequence of said primers being derived from human chromosome 17q, wherein the use of said primers in a polymerase chain reaction results in the synthesis of DNA having all or part of the sequence of the BRCA1 gene.
6. The pair of primers of claim 5 wherein said sequence of the BRCA1 gene is set forth in SEQ ID No:1.
7. An isolated DNA sequence consisting essentially of DNA coding for a mutant BRCA1 polypeptide, said DNA selected from the group consisting of cDNA and genomic DNA.
8. The isolated DNA of claim 7 selected from the group consisting of:
 - (a) SEQ ID NO:1 having T at nucleotide position 4000;
 - (b) SEQ ID NO:1 having extra C at nucleotide position 5329;
 - (c) SEQ ID NO:1 having G at nucleotide position 5387; and,
 - (d) SEQ ID NO:1 having 11 base pairs at nucleotide positions 133-143 deleted.
9. A nucleic acid probe complementary to human wild-type BRCA1 gene sequences.

10. A nucleic acid probe complementary to human altered BRCA1 gene sequences.
11. The nucleic acid probe which hybridizes to part of the altered BRCA1 gene, said part of the BRCA1 gene selected from the group consisting of:
 - (a) SEQ ID NO:1 having T at nucleotide position 4000;
 - 5 (b) SEQ ID NO:1 having extra C at nucleotide position 5329;
 - (c) SEQ ID NO:1 having G at nucleotide position 5387; and
 - (d) SEQ ID NO:1 having 11 base pairs at nucleotide positions 133-143 deleted.
12. A replicative cloning vector which comprises the isolated DNA of claim 1 and a replicon operative in a host cell.
- 10 13. An expression system which comprises the isolated DNA of claim 1 operably linked to suitable control sequences.
14. Recombinant host cells transformed with the expression system of claim 13.
15. A method of producing recombinant BRCA1 polypeptide which comprises culturing the cells of claim 14 under conditions effective for the production of said BRCA1 polypeptide.
- 15 16. A preparation of human BRCA1 polypeptide substantially free of other human proteins, the amino acid sequence of said protein corresponding to that shown in SEQ ID NO:2.
17. A preparation of human BRCA1 polypeptide substantially free of other human proteins, the amino acid sequence of said polypeptide having substantial homology with the amino acid sequence set forth in SEQ ID NO:1, said polypeptide having substantially similar function
20 as wild-type BRCA1 polypeptide.
18. An antibody immunoreactive with a human BRCA1 polypeptide and not substantially immunoreactive with other human polypeptides.
19. The antibody of claim 18 which is a monoclonal antibody.

20. A method for diagnosing a polymorphism associated with predisposition for breast and ovarian cancer in a human comprising: detecting a germline alteration of the wild-type BRCA1 gene or its expression products in a human sample, said alteration indicating a predisposition to said cancer.
- 5 21. The method of claim 20 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.
22. The method of claim 20 wherein said expression product is mRNA of BRCA1.
23. The method of claim 22 wherein the alteration of the wild-type BRCA1 gene mRNA is detected by hybridization of the mRNA from said tissue sample to a BRCA1 gene probe.
- 10 24. The method of claim 20 wherein an alteration is detected in the regulatory regions of the BRCA1 gene.
25. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by observing shifts in electrophoretic mobility of single-stranded DNA on non-denaturing polyacrylamide gels.
- 15 26. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by hybridization of a BRCA1 gene probe to genomic DNA isolated from said tissue.
27. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by hybridization with an allele-specific probe.
- 20 28. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplifying all or part of the BRCA1 gene in said tissue to produce an amplified sequence and sequencing the amplified sequence.
29. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplifying all or part of the BRCA1 gene for a specific BRCA1 allele.

30. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by molecularly cloning all or part of the BRCA1 gene in said tissue to produce a cloned sequence and sequencing the cloned sequence.
- 5 31. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by identifying a mismatch between molecules (1) BRCA1 gene genomic DNA or BRCA1 mRNA isolated from said tissue, and (2) a nucleic acid probe complementary to the human wild-type BRCA1 gene DNA, when molecules (1) and (2) are hybridized to each other to form a duplex.
- 10 32. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplification of BRCA1 gene sequences in said tissue and hybridization of the amplified sequences to nucleic acid probes which comprise wild-type BRCA1 gene sequences.
33. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by amplification of BRCA1 gene sequences in said tissue and hybridization of the amplified sequences to nucleic acid probes which comprise non-wild-type BRCA1 gene sequences.
- 15 34. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by screening for a deletion mutation.
35. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by screening for a point mutation.
- 20 36. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by screening for an insertion.
37. The method of claim 20 wherein the alteration of the wild-type BRCA1 gene is detected by *in situ* hybridization of the BRCA1 gene with nucleic acid probes which comprise the BRCA1 gene.
38. The method of claim 20 wherein the expression products are protein molecules.

39. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by immunoblotting.
40. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by immunocytochemistry.
- 5 41. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by assaying for binding interactions between BRCA1 gene protein isolated from said tissue and its ligand.
42. The method of claim 38 wherein the alteration of wild-type BRCA1 gene protein is detected by assaying for the inhibition of biochemical activity of its binding partner.
- 10 43. A method for confirming the lack of a predisposition for cancer at the BRCA1 locus in a human, which comprises detecting the wild-type BRCA1 gene or its expression products in a human sample, the presence of the wild-type gene or expression products indicating no predisposition to cancer at the BRCA1 locus.
- 15 44. The method of claim 43 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.
45. The method of claim 43 wherein said expression product is mRNA of BRCA1.
46. The method of claim 43 wherein the expression products are protein molecules.
47. A method for diagnosing or prognosing tissue of a lesion in a human, comprising: detecting a somatic alteration of wild-type BRCA1 gene or its expression products in tissue isolated from said human, said alteration indicating mutation of said tissue.
- 20 48. The method of claim 47 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.

49. The method of claim 47 wherein said expression product is mRNA of BRCA1.
50. The method of claim 49 wherein the alteration of the wild-type BRCA1 gene mRNA is detected by hybridization of the mRNA from said tissue sample to a BRCA1 gene probe.
51. The method of claim 47 wherein the alteration is detected in the regulatory region of the
5 BRCA1 gene.
52. The method of claim 47 wherein the expression products are protein molecules.
53. A method for confirming the lack of a mutation at the BRCA1 locus in a human, which comprises detecting the wild-type BRCA1 gene or its expression products in a human sample, the presence of the wild-type gene or expression products indicating no mutation
10 resulting from the BRCA1 locus.
54. The method of claim 53 wherein said BRCA1 gene has the sequence set forth in SEQ ID NO:1.
55. The method of claim 53 wherein said expression product is mRNA of BRCA1.
56. The method of claim 53 wherein the expression products are protein molecules.
- 15 57. A method for supplying a wild-type BRCA1 gene function or a BRCA1 function substantially similar to the wild-type to a cell which has lost said gene function or has altered gene function by virtue of a mutation in the BRCA1 gene, comprising: introducing a wild-type BRCA1 gene nucleic acid or nucleic acid substantially homologous to said wild-type BRCA1 gene nucleic acid and capable of suppression of a transformed state of said cell.
- 20 58. The method of claim 57 wherein said nucleic acid is a portion of wild-type BRCA1 gene, said portion encoding a part of the wild-type BRCA1 gene polypeptide which is required for suppression of a transformed state of said cell.

59. The method of claim 57 wherein said nucleic acid is a portion of wild-type BRCA1 gene, said portion encoding all of the wild-type BRCA1 gene polypeptide which is required for suppression of a transformed state of said cell.
- 5 60. The method of claim 57 wherein said nucleic acid is substantially homologous to a portion of the wild-type BRCA1 gene, said portion encoding a polypeptide which is substantially homologous to part of the wild-type BRCA1 polypeptide and has a substantially similar function to the part of the wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell.
- 10 61. The method of claim 57 wherein said nucleic acid is substantially homologous to a portion of the wild-type BRCA1 gene, said portion encoding a polypeptide which is substantially homologous to all of the wild-type BRCA1 polypeptide and has a substantially similar function to all of the wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell.
- 15 62. The method of claim 57 wherein said nucleic acid contains the BRCA1 gene regulatory sequences.
63. The method of claim 57 wherein said nucleic acid is incorporated into the genome of said cell.
- 20 64. A method for supplying a wild-type BRCA1 gene function or a BRCA1 function substantially similar to the wild-type to a cell which has lost said gene function or has altered gene function by virtue of a mutation in the BRCA1 gene, comprising: introducing into the cell all or a part of a wild-type BRCA1 polypeptide which is required for suppression of a transformed state of said cell.
- 25 65. A method for supplying a wild-type BRCA1 gene function or a BRCA1 function substantially similar to the wild-type to a cell which has lost said gene function or has altered gene function by virtue of a mutation in the BRCA1 gene, comprising: introducing into the cell a molecule which mimics the function of wild-type BRCA1 polypeptide.

66. The method of claim 65 wherein said molecule is a polypeptide substantially homologous to the complete wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell and has a substantially similar function as the wild-type BRCA1 polypeptide.
- 5 67. The method of claim 65 wherein said molecule is a polypeptide substantially homologous to a portion of the wild-type BRCA1 polypeptide required for suppression of a transformed state of said cell and has a substantially similar function as the portion of the wild-type BRCA1 polypeptide.
- 10 68. A method for screening potential cancer therapeutics which comprises: combining a BRCA1 binding partner, BRCA1 polypeptide and a compound suspected of being a cancer therapeutic and determining the amount of binding of the BRCA1 polypeptide to its binding partner.
69. A method for screening potential cancer therapeutics which comprises: combining a BRCA1 binding partner and a compound suspected of being a cancer therapeutic and measuring the biological activity of the binding partner.
- 15 70. A method for screening potential cancer therapeutics which comprises: growing a transformed eukaryotic host cell containing an altered BRCA1 gene in the presence of a compound suspected of being a cancer therapeutic and determining the rate of growth of said host cell.
- 20 71. A transgenic animal which carries an altered BRCA1 allele from a second animal in its genome.
72. The transgenic animal of claim 71 wherein the altered BRCA1 allele is a deletion.
73. The transgenic animal of claim 71 wherein the altered BRCA1 allele is a disrupted allele.

74. A method for identifying a mutant BRCA1 nucleotide sequence in a suspected mutant BRCA1 allele which comprises comparing the nucleotide sequence of the suspected mutant BRCA1 allele with the wild-type BRCA1 nucleotide sequence, wherein a difference between the suspected mutant and the wild-type sequences identifies a mutant BRCA1 nucleotide sequence.

5

TITLE OF THE INVENTION

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

ABSTRACT OF THE DISCLOSURE

5 The present invention relates generally to the field of human genetics. Specifically, the present invention relates to methods and materials used to isolate and detect a human breast and ovarian cancer predisposing gene (BRCA1), some mutant alleles of which cause susceptibility to cancer, in particular breast and ovarian cancer. More specifically, the invention relates to germline mutations in the BRCA1 gene and their use in the diagnosis of predisposition to breast and ovarian cancer. The present invention further relates to somatic mutations in the BRCA1 gene
10 in human breast and ovarian cancer and their use in the diagnosis and prognosis of human breast and ovarian cancer. Additionally, the invention relates to somatic mutations in the BRCA1 gene in other human cancers and their use in the diagnosis and prognosis of human cancers. The invention also relates to the therapy of human cancers which have a mutation in the BRCA1 gene, including gene therapy, protein replacement therapy and protein mimetics. The invention
15 further relates to the screening of drugs for cancer therapy. Finally, the invention relates to the screening of the BRCA1 gene for mutations, which are useful for diagnosing the predisposition to breast and ovarian cancer.

Application No. 09/308100

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 CFR 1.821 - 1.825 for the following reason(s):

- ☐ 1. This application clearly fails to comply with the requirements of 37 CFR 1.821 - 1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c).
- ☒ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e).
- ☐ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached marked-up copy of the "Raw Sequence Listing."
- ☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A substitute computer readable form must be submitted as required by 37 CFR 1.825(d).
- ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 CFR 1.821(e).
- ☐ 7. Other: _____

Applicant must provide:

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing"
- ☐ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 CFR 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d)

For questions regarding compliance with these requirements, please contact:

For Rules Interpretation, call (703) 308-1123
For CRF submission help, call (703) 308-4212
For PatentIn software help, call (703) 308-6856

Please return a copy of this notice with your response.

CL

143 1/2

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/308,104

DATE: 03/17/95
TIME: 09:18:33

INPUT SET: S2800.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Skolnick, Mark H.
Goldgar, David E.
Miki, Yoshio
Swenson, Jeff
Kamb, Alexander
Harshman, Keith D.
Shattuck-Eidens, Donna M.
Tavtigian, Sean V.
Wiseman, Roger W.
Futreal, P. Andrew

(ii) TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
Susceptibility Gene

(iii) NUMBER OF SEQUENCES: 61

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Venable, Baetjer, Howard & Civiletti
(B) STREET: 1201 New York Avenue, N.W., Suite 1000
(C) CITY: Washington
(D) STATE: DC
(F) ZIP: 20005

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WordPerfect 5.1/5.2

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/308,104
(B) FILING DATE: 16-SEP-1994
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/300,266
(B) FILING DATE: 02-SEP-1994

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/289,221
(B) FILING DATE: 12-AUG-1994

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/308,104DATE: 03/17/95
TIME: 09:18:37

INPUT SET: S2800.raw

47 (vii) PRIOR APPLICATION DATA:
48 (A) APPLICATION NUMBER: US 08/246,602
49 (B) FILING DATE: 20-MAY-1994
50
51 (viii) ATTORNEY/AGENT INFORMATION:
52 (A) NAME: Ihnen, Jeffrey L.
53 (B) REGISTRATION NUMBER: 28,957
54 (C) REFERENCE/DOCKET NUMBER: 24884-109347-3
55
56 (ix) TELECOMMUNICATION INFORMATION:
57 (A) TELEPHONE: 202-962-4810
58 (B) TELEFAX: 202-962-8300
59
60
61 (2) INFORMATION FOR SEQ ID NO:1:
62
63 (i) SEQUENCE CHARACTERISTICS:
64 (A) LENGTH: 5857 base pairs
65 (B) TYPE: nucleic acid
66 (C) STRANDEDNESS: single
67 (D) TOPOLOGY: linear
68
69 (ii) MOLECULE TYPE: cDNA
70
71 (iii) HYPOTHETICAL: NO
72
73 (iv) ANTI-SENSE: NO
74
75 (vi) ORIGINAL SOURCE:
76 (A) ORGANISM: Homo sapiens
77
78 (ix) FEATURE:
79 (A) NAME/KEY: CDS
80 (B) LOCATION: 64..5769
81
82
83 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
84
85 GGCCCCCTGCG CTCAGGAGGC CTTACCCCTC TGCTCTGGGT AAAGTTCATT GGAACAGAAA 60
86
87 GAAATGGATT TATCTGCTCT TCGCGTTGAA GAAGTACAAA ATGTCATTAA TGCTATGCAG 120
88
89 AAAATCTTAG AGTGTCCCAT CTGTCTGGAG TTGATCAAGG AACCTGTCTC CACAAAGTGT 180
90
91 GACCACATAT TTTGCAAATT TTGCATGCTG AAACCTTCTCA ACCAGAAGAA AGGGCCTTCA 240
92
93 CAGTGTCCCTT TATGTAAGAA TGATATAACC AAAAGGAGCC TACAAGAAAG TACGAGATTT 300
94
95 AGTCAACTTG TTGAAGAGCT ATTGAAAATC ATTTGTGCTT TTCAGCTTGA CACAGGTTTG 360
96
97 GAGTATGCAA ACAGCTATAA TTTTGCAAAA AAGGAAAATA ACTCTCCTGA ACATCTAAAA 420
98
99 GATGAAGTTT CTATCATCCA AAGTATGGGC TACAGAAACC GTGCCAAAAG ACTTCTACAG 480

DATE: 03/17/95
TIME: 09:18:40

INPUT SET: S2800.raw

100									
101	AGTGAACCCG	AAAATCCTTC	CTTGCGAGGAA	ACCAGTCTCA	GTGTCCAAC	CTCTAACCTT		540	
102									
103	GGAAGTGTGA	GAAGTCTGAG	GACAAAGCAG	CGGATACAAC	CTCAAAGAC	GTCTGTCTAC		600	
104									
105	ATTGAATTGG	GATCTGATTC	TTCTGAAGAT	ACCGTTAATA	AGGCAACTTA	TTGCAGTGTG		660	
106									
107	GGAGATCAAG	AATTGTTACA	AATCACCCC	CAAGGAACCA	GGGATGAAAT	CAGTTTGGAT		720	
108									
109	TCTGCAAAAA	AGGCTGCTTG	TGAATTTTCT	GAGACGGATG	TAACAAATAC	TGAACATCAT		780	
110									
111	CAACCCAGTA	ATAATGATTT	GAACACCACT	GAGAAGCGTG	CAGCTGAGAG	GCATCCAGAA		840	
112									
113	AAGTATCAGG	GTAGTTCTGT	TTCAAACCTG	CATGTGGAGC	CATGTGGCAC	AAATACTCAT		900	
114									
115	GCCAGCTCAT	TACAGCATGA	GAACAGCAGT	TTATTACTCA	CTAAAGACAG	AATGAATGTA		960	
116									
117	GAAAAGGCTG	AATTCTGTAA	TAAAAGCAAA	CAGCCTGGCT	TAGCAAGGAG	CCAACATAAC		1020	
118									
119	AGATGGGCTG	GAAGTAAGGA	AACATGTAAT	GATAGGCGGA	CTCCCAGCAC	AGAAAAAAAAAG		1080	
120									
121	GTAGATCTGA	ATGCTGATCC	CCTGTGTGAG	AGAAAAGAAT	GGAATAAGCA	GAAACTGCCA		1140	
122									
123	TGCTCAGAGA	ATCCTAGAGA	TACTGAAGAT	GTTCCTTGGA	TAACACTAAA	TAGCAGCATT		1200	
124									
125	CAGAAAGTTA	ATGAGTGGTT	TTCCAGAAGT	GATGAAGTGT	TAGGTTCTGA	TGACTCACAT		1260	
126									
127	GATGGGGAGT	CTGAATCAAA	TGCCAAAGTA	GCTGATGTAT	TGGACGTCCT	AATTGAGGTA		1320	
128									
129	GATGAATATT	CTGGTTCTTC	AGAGAAAATA	GACTTACTGG	CCAGTGATCC	TCATGAGGCT		1380	
130									
131	TTAATATGTA	AAAGTGACAG	AGTTCAC'TCC	AAATCAGTAG	AGAGTAATAT	TGAAGGCCAA		1440	
132									
133	ATATTTGGGA	AAACCTATCG	GAAGAAGGCA	AGCCTCCCCA	ACTTAAGCCA	TGTAAGTGAA		1500	
134									
135	AATCTAATTA	TAGNGCATT	TGTTTCTGAG	CCACAGATAA	TACAAGAGCG	TCCCCTCACA		1560	
136									
137	AATAAATTAA	AGCGTAAAAG	GAGACCTACA	TCAGGCCTTC	ATCCTGAGGA	TTTTATCAAG		1620	
138									
139	AAAGCAGATT	TGGCAGTTCA	AAAGACTCCT	GAAATGATAA	ATCAGGGAAC	TAACCAAACG		1680	
140									
141	GAGCAGAATG	GTCAAGTGAT	GAATATTACT	AATAGTGGTC	ATGAGAATAA	AACAAAAGGT		1740	
142									
143	GATTCTATTG	AGAATGAGAA	AAATCCTAAC	CCAATAGAAT	CACTCGAAAA	AGAATCTGCT		1800	
144									
145	TTCAAAACGA	AAGCTGAACC	TATAAGCAGC	AGTATAAGCA	ATATGGAACT	CGAATTAAAT		1860	
146									
147	ATCCACAATT	CAAAAGCACC	TAAAAAGAAT	AGGCTGAGGA	GGAAGTCTTC	TACCAGGCAT		1920	
148									
149	ATTCATGCGC	TTGAACTAGT	AGTCAGTAGA	AATCTAAGCC	CACCTAATTG	TACTGAATTG		1980	
150									
151	CAAATTGATA	GTTGTTCTAG	CAGTGAAGAG	ATAAAGAAAA	AAAAGTACAA	CCAAATGCCA		2040	
152									

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/308,104

DATE: 03/17/95
 TIME: 09:18:44

INPUT SET: S2800.raw

153	GTCAGGCACA	GCAGAAACCT	ACAACTCATG	GAAGGTAAAG	AACCTGCAAC	TGGAGCCAAG	2100
154							
155	AAGAGTAACA	AGCCAAATGA	ACAGACAAGT	AAAAGACATG	ACAGTGATAC	TTTCCCAGAG	2160
156							
157	CTGAAGTTAA	CAAATGCACC	TGGTTCTTTT	ACTAAGTGTT	CAAATACCAG	TGAACTTAAA	2220
158							
159	GAATTTGTCA	ATCCTAGCCT	TCCAAGAGAA	GAAAAAGAAG	AGAAACTAGA	AACAGTTAAA	2280
160							
161	GTGTCTAATA	ATGCTGAAGA	CCCCAAAGAT	CTCATGTTAA	GTGGAGAAAG	GGTTTTGCAA	2340
162							
163	ACTGAAAGAT	CTGTAGAGAG	TAGCAGTATT	TCACTGGTAC	CTGGTACTGA	TTATGGCACT	2400
164							
165	CAGGAAAGTA	TCTCGTTACT	GGAAGTTAGC	ACTCTAGGGA	AGGCAAAAAC	AGAACCAAAT	2460
166							
167	AAATGTGTGA	GTCAGTGTGC	AGCATTTGAA	AACCCCAAGG	GAATAATTCA	TGGTTGTTCC	2520
168							
169	AAAGATAATA	GAAATGACAC	AGAAGGCTTT	AAGTATCCAT	TGGGACATGA	AGTTAACCAC	2580
170							
171	AGTCGGGAAA	CAAGCATAGA	AATGGAAGAA	AGTGAACCTG	ATGCTCAGTA	TTTGCAGAAT	2640
172							
173	ACATTCAAGG	TTTCAAAGCG	CCAGTCATTT	GCTCTGTTTT	CAAATCCAGG	AAATGCAGAA	2700
174							
175	GAGGAATGTG	CAACATTCTC	TGCCCCACTCT	GGGTCCTTAA	AGAAACAAAG	TCCAAAAGTC	2760
176							
177	ACTTTTGAAT	GTGAACAAAA	GGAAGAAAAAT	CAAGGAAAGA	ATGAGTCTAA	TATCAAGCCT	2820
178							
179	GTACAGACAG	TTAATATCAC	TGCAGGCTTT	CCTGTGGTTG	GTCAGAAAGA	TAAGCCAGTT	2880
180							
181	GATAATGCCA	AATGTAGTAT	CAAAGGAGGC	TCTAGGTTTT	GTCTATCATC	TCAGTTCAGA	2940
182							
183	GGCAACGAAA	CTGGACTCAT	TACTCCAAAT	AAACATGGAC	TTTTACAAAA	CCCATATCGT	3000
184							
185	ATACCACCAC	TTTTTCCCAT	CAAGTCATTT	GTTAAAACTA	AATGTAAGAA	AAATCTGCTA	3060
186							
187	GAGGAAAACT	TTGAGGAACA	TTCAATGTCA	CCTGAAAGAG	AAATGGGAAA	TGAGAACATT	3120
188							
189	CCAAGTACAG	TGAGCACAAT	TAGCCGTAAT	AACATTAGAG	AAAATGTTTT	TAAAGAAGCC	3180
190							
191	AGCTCAAGCA	ATATTAATGA	AGTAGGTTCC	AGTACTAATG	AAGTGGGCTC	CAGTATTAAT	3240
192							
193	GAAATAGGTT	CCAGTGATGA	AAACATTCAA	GCAGAACTAG	GTAGAAACAG	AGGGCCAAAA	3300
194							
195	TTGAATGCTA	TGCTTAGATT	AGGGGTTTTG	CAACCTGAGG	TCTATAAACA	AAGTCTTCCT	3360
196							
197	GGAAGTAATT	GTAAGCATCC	TGAAATAAAA	AAGCAAGAAT	ATGAAGAAGT	AGTTCAGACT	3420
198							
199	GTAAATACAG	ATTTCTCTCC	ATCTCTGATT	TCAGATAACT	TAGAACAGCC	TATGGGAAGT	3480
200							
201	AGTCATGCAT	CTCAGGTTTG	TTCTGAGACA	CCTGATGACC	TGTTAGATGA	TGGTGAAATA	3540
202							
203	AAGGAAGATA	CTAGTTTTGC	TGAAAATGAC	ATTAAGGAAA	GTTCTGCTGT	TTTTAGCAAA	3600
204							
205	AGCGTCCAGA	AAGGAGAGCT	TAGCAGGAGT	CCTAGCCCTT	TCACCCATAC	ACATTTGGCT	3660

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/308,104

INPUT SET: S2800.raw

GeneDX 1021. pg. 250



UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
ASSISTANT SECRETARY AND COMMISSIONER
OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

Paper No. 4

Venable, Baetjer, Howard & Civiletti
1201 New York Avenue, N.W., Suite 1000
Washington, D.C. 20005

COPY MAILED
AUG 20 1995
SPECIAL PROGRAMS OFFICE
OAC FOR PATENTS

In re Application of :
Mark H. Skolnick, et al : REQUIREMENT FOR INFORMATION
Application No. 08/308,104 : RE PETITION UNDER 37 CFR
Filed: September 16, 1994 : 1.48(a)
Attorney Docket No.24884-109347-3 :

This is in response to a petition under 37 CFR 1.48(a) filed February 13, 1995, requesting that the name of the inventors in the application be corrected. The petition is being decided by the undersigned rather than in the examining group so that application processing, including the issuance of a filing receipt, may proceed with the actual inventors being named in the application.

The petition under § 1.48(a) to correct the inventorship cannot be granted at this time.

Applicants are given TWO MONTHS from the date of mailing of this Requirement for Information to correct the deficiencies noted below (a § 1.63 declaration and a verified statement of facts signed by Goldgar and clarification regarding the existence of any assignee). Failure to respond completely to the Requirement will result in abandonment of the application. The time for response, however, may be extended by compliance with the provisions of 37 CFR 1.136(a).

The instant application was filed on September 16, 1994 (and is related to prior applications), pursuant to 37 CFR 1.53(b) without the filing fees and without an executed oath or declaration under 37 CFR 1.63 and named as the inventors: Skolnick, Goldgar, Miki, Swenson, Kamb, Harshman, Shattuck-Eidens and Tavigian.

Application Processing Division mailed a Notice to File Missing Parts of Application under 37 CFR 1.53(b) on December 6, 1994, requiring an oath or declaration in compliance with 37 CFR 1.63, the filing fees and a surcharge for their late submission.

In response to the Notice to File Missing Parts of Application, applicants timely filed on February 13, 1995 (with a petition and fee for a two month extension of time under 37 CFR 1.136(a)), inter alia, the instant \$ 1.48(a) petition and fee, the filing fees, the surcharge, four combined Declarations under 37 CFR 1.63 and Powers of Attorney (two being duplicates) by the actual inventors except for Goldgar and two verified statements of facts by the original named inventors except for Goldgar (the heading refers to a filing date of September 2, 1994 rather than September 16, 1994).

The petition and verified statements of facts state that an error occurred in failing to name Wiseman and Futreal as inventors, that the instant application was filed to avoid a loss of rights, that it was determined at the time of filing that the inventorship of the instant application was the same as that of prior applications and that after the instant application was filed the issue was raised as to the inventorship of Wiseman and Futreal.

37 CFR 1.48(a) requires a verified statement of facts from each original named inventor. A verified statement of facts from original named inventor Goldgar has not been submitted as is required. The statement of facts signed by counsel on behalf of Goldgar is ineffective notwithstanding a power of attorney by Goldgar on behalf of counsel to sign a Verified Statement of Facts. Goldgar is the party with first hand knowledge of what Goldgar knows in regard to the inventorship error and he is required to sign a verified statement of facts if only to the effect that he lacks any first hand knowledge thereof. Similarly, the \$ 1.63 declaration signed by counsel on behalf of Goldgar is ineffective as to the averments contained therein. A declaration by Goldgar is required.

In view of the references to the University of Utah, Myriad Genetics, Inc. and the National Institute of Environmental Health Sciences, clarification is requested regarding the existence of any assignees that would require their written consents supported by a \$ 3.73(b) certification.

Receipt is acknowledged of a February 13, 1995 Preliminary Amendment.

Further correspondence with respect to this matter should be addressed as follows:

By mail: Commissioner of Patents and Trademarks
Box DAC
Washington, D.C. 20231

By FAX: (703) 308-6916
Attn: Special Program Law Office

Application No. 08/308,104

Page 3

By hand: One Crystal Park, Suite 520
2011 Crystal Drive
Arlington, VA

Telephone inquiries specific to this matter should be directed to the undersigned at (703) 305-8713.



Hiram H. Bernstein
Senior Legal Advisor
Special Program Law Office
Office of the Deputy Assistant Commissioner
for Patent Policy and Projects

Application No. 08/308104

**NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING
NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES**

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 CFR 1.821 - 1.825 for the following reason(s):

- ☐ 1. This application clearly fails to comply with the requirements of 37 CFR 1.821 - 1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c).
- ☒ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e).
- ☐ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached marked-up copy of the "Raw Sequence Listing."
- ☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A substitute computer readable form must be submitted as required by 37 CFR 1.825(d).
- ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 CFR 1.821(e).
- ☐ 7. Other: _____

Applicant must provide:

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing"
- ☐ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 CFR 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d)

For questions regarding compliance with these requirements, please contact:

For Rules Interpretation, call (703) 308-1123
For CRF submission help, call (703) 308-4212
For PatentIn software help, call (703) 308-6856

Please return a copy of this notice with your response.

#5

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. patent application of:

Mark H. SKOLNICK et al.

Serial No.: 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

RECEIVED

Attn: Office of Deputy Assistant

AUG 15 1995 Commissioner

OFFICE OF PETITIONS
AND APPEALS

COMMUNICATION

Hon. Commissioner of Patents and Trademarks
Washington, D.C. 20231

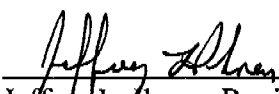
Sir:

Further to the Petition to Correct Inventorship filed on 13 February for the above case, Applicants note the existence of the following assignees: Myriad Genetics, Inc., the University of Utah Research Foundation and the United States of America, as represented by the Secretary of the Department of Health and Human Services. Transmitted herewith are Consents of Assignee for each assignee and Certificates Under 37 CFR §3.73(b) for each assignee.

Consideration of the Petition filed 13 February 1995 is requested in view of this information.

If any Petition for Extension of Time is necessary in view of this paper, an appropriate extension is hereby requested. The Commissioner is authorized to charge such extension fees, if any, to Deposit Account No. 22-0261. This Communication is submitted in triplicate.

Respectfully submitted


Jeffrey L. Ihnen, Registration No. 28,957

VENABLE, BAETJER, HOWARD & CIVILETTI
1201 New York Avenue, N.W., #1000
Washington, D.C. 20005
(202) 962-4810

Atty. Docket No. 24884-109347-3
Dated: 13 August 1995
6565.01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. Application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

CONSENT OF ASSIGNEE

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

An Assignee, Myriad Genetics, Inc., of the above-identified application, assignment of which is being filed concurrently herewith, hereby consents to the amendment of inventorship to add Roger W. Wiseman and P. Andrew Futreal to the inventors named on the application filed on 16 September 1994 under 37 C.F.R. §1.53(d).

MYRIAD GENETICS, INC.

4/26/95
Date

By: Peter D. Meldrum
Peter D. Meldrum
President

CERTIFICATE UNDER 37 CFR §3.73(b)

Applicant: Mark H. SKOLNICK et al.

Application No.: 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

MYRIAD GENETICS, INC., a corporation, certifies that it is an assignee of the entire right, title and interest in the patent application identified above by virtue of:

A chain of title from the inventors of the patent application identified above, to the current assignee as shown below:

1. From: Alexander Kamb, Keith D. Harshman, Donna M. Shattuck-Eidens
and Sean V. Tavtigian
To: Myriad Genetics, Inc.
for which a copy is attached.

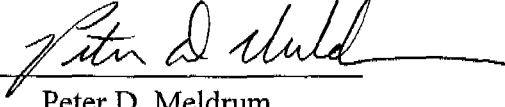
Copies of assignments or other documents in the chain of title are attached.

The undersigned has reviewed all the documents in the chain of title of the patent application identified above and, to the best of his 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, any patent issuing thereon.

For: MYRIAD GENETICS, INC.

Date: 4/26/95 Signature: 
Name: Peter D. Meldrum
Title: President

ASSIGNMENT

WHEREAS, WE, Alexander KAMB, a citizen of the United States of America residing at 1103 East 600 South, Salt Lake City, Utah 84102; Keith D. HARSHMAN, a citizen of the United States of America residing at 159 North F Street, Salt Lake City, Utah 84103; Donna M. SHATTUCK-EIDENS, a citizen of the United States of America residing at 1874 Texas Street, Salt Lake City, Utah 84108; and Sean V. TAVTIGIAN, a citizen of the United States of America residing at 557 East 1st Ave, #3, Salt Lake City, Utah 84103 have jointly invented certain new and useful improvements in

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

for which an application for Letters Patent of the United States was filed on 16 September 1994 and assigned Serial No. 08/308,104, and

WHEREAS, MYRIAD GENETICS, INC., a corporation duly organized under the laws of the State of Utah, having a place of business at 417 Wakara Way, Suite 300, Salt Lake City, Utah 84108 (hereinafter referred to as "Assignee") is desirous of acquiring the entire right, title and interest in and to the aforesaid application in which the invention is disclosed, described and claimed, in and to the above-identified invention for the United States of America and all countries foreign thereto, and in and to any and all Letters Patents to be issued for said invention in the United States of America and all countries foreign thereto, including any and all continuations, divisions, reissues, or extensions, of such application or Letters Patent;

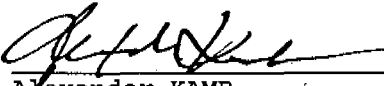
NOW, THEREFORE, TO ALL WHOM IT MAY CONCERN, be it known that for good and valuable consideration, we have bargained, sold and assigned, transferred and set over and by these presents do hereby bargain, sell, assign, transfer and set over unto Assignee, its successors, assigns or other legal representatives, the entire right, title and interest in and to the aforesaid application in which the invention is disclosed, described and claimed, in and to the above-identified invention for the United States of America and all countries foreign thereto, and in and to any and all Letters Patents to be issued for said invention in the United States of America and all countries foreign thereto, and any and all continuations, divisions, reissues, or extensions of any such Letters Patents; the said to be held and enjoyed by Assignee for its own use and on behalf of its successors, assigns, or other legal representatives, to the full end of the term or terms for which said Letters Patents may be granted, reissued, or extended as fully and entirely as the same would have been held and enjoyed by us, jointly and severally, if this assignment and sale had not been made.

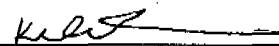
We further covenant and agree that we will at any time, upon request, execute and deliver any and all papers that may be necessary or desirable to perfect the title to said invention, or any Letters Patents that may be granted therefor, in Assignee, its successors, assigns, or other legal representatives, and that if Assignee, its successors, assigns or other legal representatives desire to secure a reissue or extension or division of any Letters Patents, we will upon request sign all papers, make all rightful oaths, and do all legal acts necessary and requisite for application of such reissue, extension or division and the procuring thereof, but at the expense of Assignee, its successors, assigns or other legal representatives.

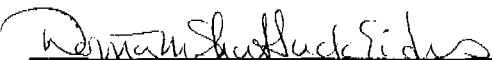
We hereby further covenant and agree that we will at any time, upon request, communicate to Assignee, its successors, assigns or other legal representatives, any facts relating to the said invention and Letters Patents, or to the history thereof known to us, and will testify to the same in any interference or litigation, when requested to do so.

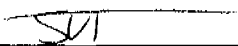
We hereby authorize and request the Commissioner of Patents to issue any and all Letters Patents covering said invention, or resulting from the aforesaid application to Assignee as the assignee of the entire right, title and interest in and to the said application and invention.

IN WITNESS WHEREOF, we have hereunto set our hand and seal on the date(s) given below.

12/21/94
Date

Alexander KAMB

12.14.94
Date

Keith D. HARSHMAN

12/13/94
Date

Donna M. SHATTUCK-EIDENS

12/13/94
Date

Sean V. TAVTIGIAN

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. Application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

CONSENT OF ASSIGNEE

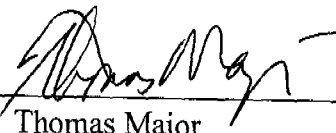
Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

An Assignee, the University of Utah Research Foundation, of the above-identified application, assignments of which are being filed concurrently herewith, hereby consents to the amendment of inventorship to add Roger W. Wiseman and P. Andrew Futreal to the inventors named on the application filed on 16 September 1994 under 37 C.F.R. §1.53(d).

UNIVERSITY OF UTAH RESEARCH FOUNDATION

4/28/95
Date

By: 
Thomas Major
Director, Technology Transfer Office

CERTIFICATE UNDER 37 CFR §3.73(b)

Applicant: Mark H. SKOLNICK et al.

Application No.: 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

UNIVERSITY OF UTAH RESEARCH FOUNDATION, a university foundation, certifies that it is an assignee of the entire right, title and interest in the patent application identified above by virtue of:

A chain of title from the inventors of the patent application identified above, to the current assignee as shown below:

1. From: Mark H. Skolnick, David E. Goldgar, Yoshio Miki and Jeff Swenson
To: University of Utah
for which a copy is attached.
2. From: University of Utah
To: University of Utah Research Foundation
for which a copy is attached.

Copies of assignments or other documents in the chain of title are attached.

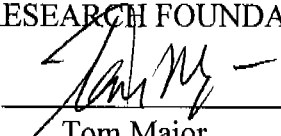
The undersigned has reviewed all the documents in the chain of title of the patent application identified above and, to the best of his knowledge and belief, title is in the assignee identified above.

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, any patent issuing thereon.

For: UNIVERSITY OF UTAH RESEARCH FOUNDATION

Date: 4/28/95

Signature: 

Name:

Tom Major

Title:

Director, Technology Transfer Office

-2-

IN WITNESS WHEREOF, Assignor has hereunto set its hand and seal on the date(s) given below.

12/13/77
Date

THE UNIVERSITY OF UTAH

By:

Thomas Major
Director
Technology Transfer Office

ASSIGNMENT

WHEREAS, WE, Mark H. SKOLNICK, a citizen of the United States of America residing at 460 Northmount Way, Salt Lake City, Utah 84103; David E. GOLDFAR, a citizen of the United States of America residing at 2707 East Kelly Lane, Salt Lake City, Utah 84117; Yoshio MIKI, a citizen of Japan residing at 40 South 900 East, Salt Lake City, Utah 84102; and Jeff SWENSON, a citizen of the United States of America residing at 745 West Creek Bend, #275, Salt Lake City, Utah 84117 have jointly invented certain new and useful improvements in

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE

for which an application for Letters Patent of the United States was filed on 16 September 1994 and assigned Serial No. 08/308,104, and

WHEREAS, THE UNIVERSITY OF UTAH, a non-profit institution duly organized under the laws of the State of Utah, having a place of business at 421 Wakara Way, Suite 170, Salt Lake City, Utah 84108 (hereinafter referred to as "Assignee") is desirous of acquiring the entire right, title and interest in and to the aforesaid application in which the invention is disclosed, described and claimed, in and to the above-identified invention for the United States of America and all countries foreign thereto, and in and to any and all Letters Patents to be issued for said invention in the United States of America and all countries foreign thereto, including any and all continuations, divisions, reissues, or extensions, of such application or Letters Patent;

NOW, THEREFORE, TO ALL WHOM IT MAY CONCERN, be it known that for good and valuable consideration, we have bargained, sold and assigned, transferred and set over and by these presents do hereby bargain, sell, assign, transfer and set over unto Assignee, its successors, assigns or other legal representatives, the entire right, title and interest in and to the aforesaid application in which the invention is disclosed, described and claimed, in and to the above-identified invention for the United States of America and all countries foreign thereto, and in and to any and all Letters Patents to be issued for said invention in the United States of America and all countries foreign thereto, and any and all continuations, divisions, reissues, or extensions of any such Letters Patents; the said to be held and enjoyed by Assignee for its own use and on behalf of its successors, assigns, or other legal representatives, to the full end of the term or terms for which said Letters Patents may be granted, reissued, or extended as fully and entirely as the same would have been held and enjoyed by us, jointly and severally, if this assignment and sale had not been made.

We further covenant and agree that we will at any time, upon request, execute and deliver any and all papers that may be necessary or desirable to perfect the title to said invention, or any Letters Patents that may be granted therefor, in Assignee, its successors, assigns, or other legal representatives, and that if Assignee, its successors, assigns or other legal representatives desire to secure a reissue or extension or division of any Letters Patents, we will upon request sign all papers, make all rightful oaths, and do all legal acts necessary and requisite for application of such reissue, extension or division and the procuring thereof, but at the expense of Assignee, its successors, assigns or other legal representatives.

We hereby further covenant and agree that we will at any time, upon request, communicate to Assignee, its successors, assigns or other legal representatives, any facts relating to the said invention and Letters Patents, or to the history thereof known to us, and will testify to the same in any interference or litigation, when requested to do so.

We hereby authorize and request the Commissioner of Patents to issue any and all Letters Patents covering said invention, or resulting from the aforesaid application to Assignee as the assignee of the entire right, title and interest in and to the said application and invention.

IN WITNESS WHEREOF, we have hereunto set our hand and seal on the date(s) given below.

Dec 13, 1994
Date

Mark H. Skolnick
Mark H. SKOLNICK

Dec 14 1994
Date

David E. Goldgar
David E. GOLDGAR

Dec 14 1994
Date

Yoshio Miki
Yoshio MIKI

Dec 14 1994
Date

Jeff Swenson
Jeff SWENSON

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. Application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

CONSENT OF ASSIGNEE

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

An Assignee, the United States of America, as represented by the Secretary of the Department of Health and Human Services, of the above-identified application, assignment of which is being filed concurrently herewith, hereby consents to the amendment of inventorship to add Roger W. Wiseman and P. Andrew Futreal to the inventors named on the application filed on 16 September 1994 under 37 C.F.R. §1.53(d).

UNITED STATES OF AMERICA, as represented by the
Secretary of the Department of Health and Human Services

8 August 1995
Date

By: _____



Susan Rucker, Patent Advisor
for the Director, Dr. Marie Freire
Office of Technology Transfer

CERTIFICATE UNDER 37 CFR §3.73(b)

Applicant: Mark H. SKOLNICK et al.

Application No.: 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

The UNITED STATES OF AMERICA as represented by the Secretary of the Department of Health and Human Services, certifies that it is an assignee of the entire right, title and interest in the patent application identified above by virtue of:

A chain of title from the inventors of the patent application identified above, to the current assignee as shown below:

1. From: Roger W. Wiseman and P. Andrew Futreal
To: the United States of America as represented by the Secretary of the Department of Health and Human Services
for which a copy is attached.

Copies of assignments or other documents in the chain of title are attached.

The undersigned has reviewed all the documents in the chain of title of the patent application identified above and, to the best of her 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, any patent issuing thereon.

For: The UNITED STATES OF AMERICA as represented by
the Secretary of the Department of Health and Human Services

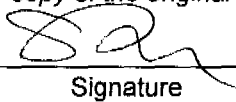
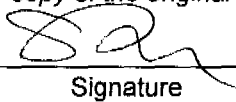
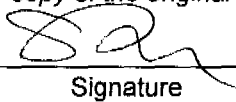
Date: 8/9/95

Signature: 

Name: Susan Rucker

Title: Patent Advisor
for the Director, Dr. Marie Freire
Office of Technology Transfer

RECORDATION FORM COVER SHEET
PATENTS ONLY

1. Name of Conveying party(ies): P. Andrew Futreal Roger W. Wiseman		2. Name and address of receiving party(ies): The Government of the United States of America, as represented by the Secretary, Department of Health and Human Services Internal Address: Office of Technology Transfer National Institutes of Health 6011 Executive Blvd. Rockville, MD 20852-3804 Street Address: same as above				
3. Nature of conveyance: [X] Assignment Execution Date: August 1, 1995 and July 31, 1995						
4. Application number(s) or patent numbers: A. Patent Application No.(s): Serial Number 08/308,104 filed September 16, 1994 entitled: 17q-Linked Breast and Ovarian Cancer Susceptibility Gene						
5. Name and address of party to whom correspondence concerning document should be mailed: Name: <u>Susan S. Rucker</u> Internal Address: Office of Technology Transfer National Institutes of Health 6011 Executive Blvd. Rockville, MD 20852-3804 Street Address: same as above		6. Total Number of patents and applications involved: <u>1</u>				
		7. Total Fee (37 C.F.R. §3.41) \$80.00 [X] Please charge any deficiency or credit any overpayment to the below-numbered deposit account.				
		8. Deposit Account Number: 14-1052				
9. Statement and signature. <i>To the best of my knowledge and belief, the foregoing information is true and correct and any attached copy is a true copy of the original document.</i> <table><tr><td><u>Susan S. Rucker (35,762)</u> Name of Person Signing</td><td> Signature</td><td><u>August 3, 1995</u> Date</td></tr></table>				<u>Susan S. Rucker (35,762)</u> Name of Person Signing	 Signature	<u>August 3, 1995</u> Date
<u>Susan S. Rucker (35,762)</u> Name of Person Signing	 Signature	<u>August 3, 1995</u> Date				
Total Number of pages comprising cover sheet 1						

ASSIGNMENT

WHEREAS, we, Roger W. Wiseman and P. Andrew Futreal, employees of the Department of Health and Human Services and Mark H. Skolnick, David E. Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith D. Harshman, Donna M. Shattuck-Eidens, and Sean V. Tavtigian, have invented "17q-Linked Breast and Ovarian Cancer Susceptibility Gene" for which we have made application for Letters Patent of the United States Serial No. 08/308,104 filed September 16, 1994; and

WHEREAS, we are the applicants named in the above-identified application for Letters Patent; and

WHEREAS, the conditions under which said invention was made are such as to entitle THE UNITED STATES OF AMERICA, as represented by the Secretary, Department of Health and Human Services [hereinafter THE GOVERNMENT] under Paragraph 1(a) of Executive Order 10096 to the entire right, title, and interest herein, both domestic and foreign; and

WHEREAS, THE GOVERNMENT is desirous of acquiring all domestic and foreign right, title, and interest in the above-mentioned invention described in the application for Letters Patent; and

NOW, THEREFORE, for good and valuable consideration, the receipt of which is hereby acknowledged, we hereby assign and transfer to THE GOVERNMENT, represented by the Secretary, Department of Health and Human Services, the full and exclusive rights in and to said invention in the United States and within each and every foreign country in which THE GOVERNMENT elects to file and the entire right, title, and interest in and to such applications, and any continuations, continuations-in-part, divisions, reissues or extensions thereof, and including priority rights as may be filed in the United States and foreign countries, and such Letters Patent as may be granted to be held by THE GOVERNMENT to the end of the term for which the same would have been held by the inventors had this assignment not been made.

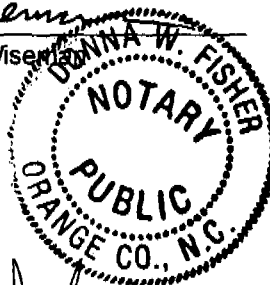
We further agree to make, execute, and deliver to the Secretary, Department of Health and Human Services, upon request, any and all papers, documents, affidavits, or other instruments that may be necessary in the prosecution of any application or applications for improvements or reissues of Letters Patent, and to assist THE GOVERNMENT in every way as may be requested in protecting said invention, provided that any expense of extending such assistance shall be paid by THE GOVERNMENT.

IN WITNESS WHEREOF we subscribe our hands and seals on the dates shown below:

7/31/95
Date

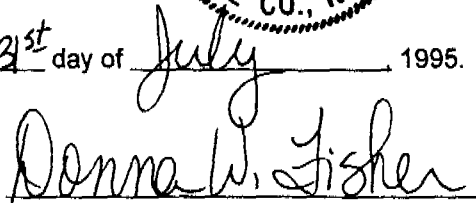

Roger W. Wiseman

County of Durham
Ss:
State of North Carolina



Subscribed and sworn to before me this 31st day of July, 1995.

[Seal]


Notary Public

ASSIGNMENT

WHEREAS, we, Roger W. Wiseman and P. Andrew Futreal, employees of the Department of Health and Human Services and Mark H. Skolnick, David E. Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith D. Harshman, Donna M. Shattuck-Eidens, and Sean V. Tavtigian, have invented "17q-Linked Breast and Ovarian Cancer Susceptibility Gene" for which we have made application for Letters Patent of the United States Serial No. 08/308,104 filed September 16, 1994; and

WHEREAS, we are the applicants named in the above-identified application for Letters Patent; and

WHEREAS, the conditions under which said invention was made are such as to entitle THE UNITED STATES OF AMERICA, as represented by the Secretary, Department of Health and Human Services [hereinafter THE GOVERNMENT] under Paragraph 1(a) of Executive Order 10096 to the entire right, title, and interest herein, both domestic and foreign; and

WHEREAS, THE GOVERNMENT is desirous of acquiring all domestic and foreign right, title, and interest in the above-mentioned invention described in the application for Letters Patent; and

NOW, THEREFORE, for good and valuable consideration, the receipt of which is hereby acknowledged, we hereby assign and transfer to THE GOVERNMENT, represented by the Secretary, Department of Health and Human Services, the full and exclusive rights in and to said invention in the United States and within each and every foreign country in which THE GOVERNMENT elects to file and the entire right, title, and interest in and to such applications, and any continuations, continuations-in-part, divisions, reissues or extensions thereof, and including priority rights as may be filed in the United States and foreign countries, and such Letters Patent as may be granted to be held by THE GOVERNMENT to the end of the term for which the same would have been held by the inventors had this assignment not been made.

We further agree to make, execute, and deliver to the Secretary, Department of Health and Human Services, upon request, any and all papers, documents, affidavits, or other instruments that may be necessary in the prosecution of any application or applications for improvements or reissues of Letters Patent, and to assist THE GOVERNMENT in every way as may be requested in protecting said invention, provided that any expense of extending such assistance shall be paid by THE GOVERNMENT.

IN WITNESS WHEREOF we subscribe our hands and seals on the dates shown below:

Date Roger W. Wiseman

County of _____
Ss:

State of _____

Subscribed and sworn to before me this ____ day of _____, 1995.

[Seal] _____
Notary Public

ASSIGNMENT
Attorney Docket No.: NIH E-237-94/4
UU U-2090

Date P. Andrew Futreal

County of _____
Ss:

State of _____

Subscribed and sworn to before me this ____ day of _____, 1995.

[Seal]

Notary Public



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
National Institutes of Health

Office of Technology Transfer
National Institutes of Health
6011 Executive Boulevard
Rockville, MD 20852

January 3, 1995

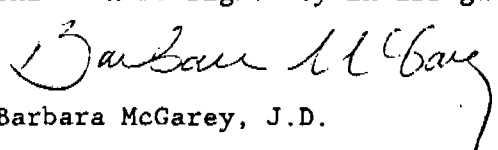
TO: Commissioner of Patents and Trademarks
U.S. Patent and Trademark Office

FROM: Acting Director, Office of Technology Transfer (OTT)
National Institutes of Health

SUBJECT: Delegation of Authority for Certain Patent
Administration and Prosecution Activities

Effective this date and acting under the authority set forth in DHHS Secretary Louis W. Sullivan's May 21, 1991 memorandum appearing in the Friday, June 7, 1991 Federal Register Notices at Volume 56, Number 110, pages 26418-26419 (copy attached), I authorize James C. Haight (Reg.#25,588), Gloria Richmond (Reg.#30,416), Robert Benson (Reg.# 33,612), Jack Spiegel (Reg.# 34,477), Denise Bernstein (Reg.# 35,787), Susan Rucker (Reg.# 35,762), David Sadowski (Reg.# 32,808), Laurence Hyman (Reg.#35,551), Arthur Cohn (Reg.# 37,800) and Steve Ferguson (Reg.# 38,448) to exercise full authority in United States or International patents or patent applications in which the Government of the United States of America as represented by the Secretary, Department of Health and Human Services is an Applicant or has an ownership interest, with respect to the following activities:

- Signing original, associate, or substitute powers of attorney before the US or International Authorities;
- Signing Request Forms, Demand Forms, and other documents filed in applications before the International Authorities;
- Signing assignee's consent to changes in inventorship;
- Signing papers required from an assignee in connection with an application for Reissue and/or Reexamination of a patent;
- Signing terminal disclaimers including but not limited to terminal disclaimers to obviate obviousness-type double patenting rejections; and
- Signing all documents before the United States Patent and Trademark Office, either as the US National Office or as a designated Receiving Office under the Patent Cooperation Treaty that can be signed by an assignee in patent matters.


Barbara McGarey, J.D.

cc: OTT Professional Staff

F.N.B. Corporation, et al: Acquisitions of Companies Engaged in Permissible Nonbanking Activities

The organizations listed in this notice have applied under § 225.23(a)(2) or (f) of the Board's Regulation Y (12 CFR 225.23(a)(2) or (f)) for the Board's approval under section 4(c)(8) of the Bank Holding Company Act (12 U.S.C. 1843(c)(8)) and § 225.21(a) of Regulation Y (12 CFR 225.21(a)) to acquire or control voting securities or assets of a company engaged in a nonbanking activity that is listed in § 225.23 of Regulation Y as closely related to banking and permissible for bank holding companies. Unless otherwise noted, such activities will be conducted throughout the United States.

Each application is available for immediate inspection at the Federal Reserve Bank indicated. Once the application has been accepted for processing, it will also be available for inspection at the offices of the Board of Governors. Interested persons may express their views in writing on the question whether consummation of the proposal can "reasonably be expected to produce benefits to the public, such as greater convenience, increased competition, or gain in efficiency, that outweigh possible adverse effects, such as undue concentration of resources, decreased or unfair competition, conflicts of interests, or unsound banking practices." Any request for a hearing on this question must be accompanied by a statement of the reasons a written presentation would not suffice in lieu of a hearing, identifying specifically any questions of fact that are in dispute, summarizing the evidence that would be presented at a hearing, and indicating how the party commenting would be aggrieved by approval of the proposal.

Unless otherwise noted, comments regarding each of these applications must be received at the Reserve Bank indicated for the application or the offices of the Board of Governors not later than June 22, 1985.

A. Federal Reserve Bank of Cleveland (John J. Wixted, Jr., Vice President) 1435 East Sixth Street, Cleveland, Ohio 44101:

1. **F.N.B. Corporation, Hermitage, Pennsylvania**: to acquire Regency Consumer Discount Company, Inc., Scranton, Pennsylvania, and thereby engage in making and acquiring consumer finance loans pursuant to § 225.23(b)(1); and in the sale of credit life, accident, and health insurance pursuant to § 225.23(b)(8) of the Board's Regulation Y.

B. Federal Reserve Bank of Atlanta (Robert E. Heck, Vice President) 104 Marietta Street, N.W., Atlanta, Georgia 30303:

1. **Evergreen Bancshares, Inc., Tallahassee, Florida**: to establish Evergreen Federal Interim Savings Bank, Tallahassee, Florida (Interim Bank), to acquire certain assets and assume certain liabilities of the Tallahassee, Florida branch office of Anchor Savings Bank, FSL, Hewlett, New York, pursuant to section 4(c)(8) of the Bank Holding Company Act and the Oaker Amendment of FIFSEA, and to facilitate the merger of Interim Bank with and into Evergreen's subsidiary bank, Guaranty National Bank of Tallahassee, Tallahassee, Florida.

2. **First State Corporation, Albany, Georgia**: to acquire Randolph Federal Savings & Loan Association, Culbert, Georgia (Randolph Federal), pursuant to section 4(c)(8) of the Bank Holding Company Act. Applicant also proposes to merge Randolph Federal with and into its bank subsidiary, First State Bank & Trust Company, Albany, Georgia, pursuant to the Oaker Amendment of FIFSEA.

Board of Governors of the Federal Reserve System, June 2, 1985.

Jennifer J. Johnson

Associate Secretary of the Board
[FR Doc. 85-12477 Filed 6-5-85; 8:11 am]
GPO: 1985 500-0000

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Office of the Secretary

Office of the General Counsel
Statement of Organization, Functions and Delegations of Authority

Part A, chapter AG (Office of the General Counsel, Office of the Secretary) of the Statement of Organization, Functions, and Delegations of Authority for the Department of Health and Human Services (38 FR 17003, June 22, 1973, as amended most recently in pertinent part at 55 FR 17500, April 22, 1990), and part H, chapter HN (National Institutes of Health, Public Health Service) (40 FR 22656, May 27, 1975, as amended most recently in pertinent part at 55 FR 35366, August 29, 1990), are amended to:

(1) Delete the Department Patent Officer (AG-35) and transfer patent administration and prosecution responsibilities from this Officer to the Office of Technology Transfer (HNA-432), Office of Intramural Affairs (HNA-43), Office of Intramural Research (HNA-4), Office of the Director (HNA).

Public Health Service (HHS) Patent

administration and prosecution activities will be consolidated with patent licensing functions within a single HHS organizational entity in order to improve program and resource management and

(2) Retain patent legal services of the Department Patent Officer (AG-35) in the Public Health Division (AG-226), Office of the General Counsel, Office of the Secretary, and update the functional statement for the Public Health Division to reflect its responsibilities more accurately.

Office of the Secretary

Under Chapter AG, Office of the General Counsel, Sections AG-35 (Department Patent Officer) and AG-226 (Public Health Division), delete the titles and statements in their entirety and substitute the following:

Public Health Division (AG-226). The Public Health Division shall provide legal services, including patent legal services, for programs administered by the Public Health Service (except the Food and Drug Administration), e.g., the Office of the Assistant Secretary for Health and the agencies and offices of the Public Health Service (other than FDA), including the Office of Technology Transfer and the Patent Policy Board.

Public Health Service

Under Chapter HN, National Institutes of Health, Section HNA, Office of the Director, Office of Intramural Affairs (HNA-43), add the following title and statement:

Office of Technology Transfer (HNA-432). (1) Develops policy and procedures for NIH, ADAMHA, and CDC to follow for the implementation of Cooperative Research and Development Agreements (CRADAs), patent licenses, and other technology transfers; (2) implements Patent Policy Board decisions and policies; (3) drafts, negotiates, and periodically revises model forms and agreements; (4) provides advice to ICDCs on problem licenses and agreements; (5) develops policy statements on various technology transfer issues such as conflicts of interest; (6) tracks the OTT budget and prepares an annual status report to the NIH Office of the Director; (7) provides coordination and management of goals, functions, and operations of the Technology Management Branch, Technology Licensing Branch, and the Technology Transfer Coordination Branch; (8) coordinates and provides planning and liaison support for

international CRADA and technology transfers; (9) creates and implements special programs relating to technology transfer by State and local governments and universities; (10) drafts and presents Congressional testimony, and drafts technology transfer-related responses to other Congressional inquiries; (11) provides operational management activities; (12) assists the Office of the General Counsel (OGC) in evaluating patent-related litigation matters; (13) in consultation with OGC and the involved agency component as appropriate, negotiates settlements on contested matters with licensees or other parties involved with NIH, ADAMHA, and CDC in technology transfer or utilization matters; (14) represents the NIH, ADAMHA, and CDC in technology transfer or utilization matters; (15) represents the above agencies at a variety of professional conferences and other public fora; (16) investigates special issues; (17) evaluates the need for and develops new programs in technology management and technology transfer for the above agencies; (18) develops licensing strategies for NIH/ADAMHA/CDC intramural and CRADA inventions; (19) negotiates licenses and other technology transfers; (20) works with scientist inventors, contract attorneys and others in preparing patent applications and presenting these applications at the Patent Office level; (21) handles infringements in consultation with the OGC at the Patent Office level; and (22) makes recommendations to the OGC for referral of matters to the Department of Justice.

Dated: May 21, 1991.
Louis W. Sullivan,
Secretary.
[FR Doc. 91-12483 Filed 6-6-91; 2:45 am]
BILLING CODE 1120-01

Agency for Health Care Policy and Research Establishment

Pursuant to the Federal Advisory Committee Act, Public Law 92-463 (5 U.S.C. appendix 2), the Administrator, Agency for Health Care Policy and Research (AHCPR), announces the establishment of the following review committee.

Designation: Employer-Based Health Insurance Advisory Committee.

Purpose: The purpose of the Committee is to advise and make recommendations to the Secretary, HHS, and the Administrator, AHCPR, with regard to the awarding of a proposed contract designed to provide AHCPR with a comprehensive policy-oriented

report that describes the current status and problems of employer-based health insurance and evaluates options for improving, reforming or replacing this system.

Function: The Committee shall review and make recommendations to the Administrator on the scientific and technical merit of proposals received in response to the Request for Proposal entitled Employer-Based Health Insurance.

Structure: The Committee shall consist of up to three members, including the Chair, who will serve for the duration of the Committee. No member may be an officer or employee of the Federal Government. Members and Chair shall be selected by the Administrator, AHCPR, from individuals with appropriate expertise and experience in health services research, including but not limited to the areas of health economics, utilization and costs of health insurance, research evaluation and dissemination, and assessment of the impact of the dissemination of research.

Notwithstanding section 14(e) of the Federal Advisory Committee Act, the Committee shall continue to exist until otherwise provided by law or upon a determination by the Administrator, AHCPR, or his delegate, that the purpose of the Committee has been accomplished.

Dated: May 21, 1991.
J. James Clinton,
Administrator, Agency for Health Care Policy and Research.
[FR Doc. 91-12564 Filed 6-6-91; 2:45 am]
BILLING CODE 1120-01

Administration for Children and Families

Forms Submitted to the Office of Management and Budget for Clearance

The Administration for Children and Families will publish on Friday's information collection packages submitted to the Office of Management and Budget (OMB) for clearance, in compliance with the Paperwork Reduction Act (44 U.S.C. chapter 35). This collection package is being submitted for expedited review in compliance with 5 CFR 1320.12.

(For a copy of a package, call the FSA Report Clearance Officer 202-401-5604.)

Plans for the Child Care and Development Block Grant Form ACF-116-NW. The information contained in the Block grant plan is to determine whether the plan can be approved for

Block Grant funding, as required in section 435(d) of the Budget Reconciliation Act; and to determine if the lead agency is operating in accordance with its plan where issues of compliance arise. **Respondents:** States or local governments; **Number of Respondents:** 233; **Frequency of Response:** Biennially (after initial submission); **Average Burden per Response:** 50 hours; **Estimated Annual Burden:** 12,750 hours.

OMB Desk Clearance Officer: Laura Oliver.

Consideration will be given to comments and suggestions received within 30 days of publication. Written comments and recommendations for the proposed information collection should be sent directly to the appropriate OMB Desk Officer designated above at the following address:

OMB Reports Management Branch, New Executive Office Building, room 3201, 725 15th Street, NW, Washington, DC 20503.

Dated: May 21, 1991.

Nancy L. Mares.

Associate Administrator, Office of Management and Information Systems.

Draft Plan For The Child Care & Development Block Grant

(Grantee) _____
for the period _____
through _____

Table of Contents

- Section 1. Assurance
- Section 2. Lead Agency Responsibilities
 - 2.1 Administration
 - 2.2 Consultation
 - 2.3 Coordination of Services
 - 2.4 Public Hearing Process
 - 2.5 Public-Private Partnerships
- Section 3. Child Care Services (75% of Block Grant Funds)
 - 3.1 Description of Services and Process to Receive Such Services
 - 3.2 Description of Activities to Improve the Availability and Quality of Child Care
 - 3.3 Localities
 - 3.4 Basis for Allocating Funds for Such Services and Activities
 - 3.5 Eligibility Criteria and Priority Rules for Children
 - 3.6 Criteria for Awarding Grants and Contracts
 - 3.7 Eligibility Terminology
 - 3.8 Tribal Eligibility Criteria
- Section 4. Activities to Improve the Quality of Child Care and to Increase the availability of early childhood Development Programs and Before- and After-School Care Services (25% of Block Grant Funds)
 - 4.1 Description of Activities
 - 4.2 Localities



UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
ASSISTANT SECRETARY AND COMMISSIONER
OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

COPY MAILED Paper No. 6

SEP 29 1995

**SPECIAL PROGRAMS OFFICE
DAC FOR PATENTS**

Venable, Baetjer, Howard & Civiletti
1201 New York Avenue, N.W., Suite 1000
Washington, D.C. 20005

In re Application of :
Mark H. Skolnick, et al : SUPPLEMENTAL REQUIREMENT
Application No. 08/308,104 : FOR INFORMATION RE PETITION
Filed: September 16, 1994 : UNDER 37 CFR 1.48(a)
Attorney Docket No.24884-109347-3 :

This is a supplement to an August 28, 1995 Requirement for Information that responds to an August 15, 1995 paper recently matched with the file record.

The Requirement for Information required in-part clarification regarding the existence of assignees. The August 15, 1995 paper has submitted three written consents of assignees supported by § 3.73(b) certifications.

Accordingly, the requirement for clarification regarding the existence of assignees is withdrawn. The requirements for a verified statement of facts and a § 1.63 declaration from original named and actual inventor Goldgar remains.

The period for response set forth in the Requirement for Information continues to run with extensions of time available under 37 CFR 1.136(a).

The application will be retained by this Office to await a complete response to the outstanding Requirement for Information as modified herein.

Telephone inquiries specific to this matter should be directed to the undersigned at (703) 305-8713.

Hiram H. Bernstein
Senior Legal Advisor
Special Program Law Office
Office of the Deputy Assistant Commissioner
for Patent Policy and Projects



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. patent application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 16 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

RECEIVED

NOV 01 1995

OFFICE OF PETITIONS

Attn: Office of Deputy Assistant
Commissioner

RESPONSE TO REQUIREMENT FOR INFORMATION

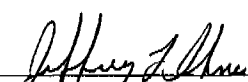
Hon. Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

In response to the Requirement for Information mailed 28 August 1995 and the Supplemental Requirement for Information mailed 29 September 1995, Applicants submit herewith a Verified Statement of Facts and a Declaration and Power of Attorney which have been executed by David E. Goldgar.

Consideration of the Petition filed 13 February 1995 is requested in view of this information.

Respectfully submitted,



Jeffrey L. Innen,
Registration No. 28,957

VENABLE, BAETJER, HOWARD & CIVILETTI, LLP
1201 New York Avenue, N.W., #1000
Washington, D.C. 20005
(202) 962-4810
Atty. Docket No. 24884-109347-3
Dated: 26 October 1995
647.01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. Application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Filed: 2 September 1994

For: 17q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE

VERIFIED STATEMENT OF FACTS

Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

Mark H. Skolnick, David E. Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith D. Harshman, Donna M. Shattuck-Eidens and Sean V. Tavtigian declare that:

(1) The invention described and claimed in the above-identified application was developed in part by us and in part by Roger W. Wiseman and P. Andrew Futreal.

(2) Mark Skolnick, David Goldgar, Yoshio Miki and Jeff Swenson are employees of the University of Utah (University) and have been conducting research to identify and characterize a breast cancer gene on human chromosome 17 since at least January 1991. This research has led to three earlier-filed related applications of which Skolnick, Goldgar, Miki and Swenson are coinventors.

(3) Mark Skolnick and his fellow University researchers began collaborating with Myriad Genetics, Inc. (Myriad) and its researchers on May 1, 1992 to intensify the search for the breast cancer gene. As a result of this collaborative effort, Mark Skolnick, David Goldgar, Yoshio Miki, Jeff Swenson, Alexander Kamb, Keith Harshman and Donna Shattuck-Eidens identified a gene which was disclosed and claimed in a fourth, earlier-filed

related application as a breast cancer gene. This fourth related application also disclosed other candidate DNA sequences.

(4) Mark Skolnick and his fellow University researchers began collaborating with Roger Wiseman and his fellow researchers at the National Institute of Environmental Health Sciences (NIEHS) in April 1992 to intensify the search for the breast cancer gene. The researchers at Myriad were included in this collaboration, as well as researchers at other universities and companies.

(5) Mark Skolnick, Alexander Kamb or Donna Shattuck-Eidens periodically communicated research results with Jeffrey Ihnen, patent counsel to the University and Myriad, to update him on the results of their research. In early August 1994, Jeffrey Ihnen was advised that the gene identified in the fourth related application as being the breast cancer gene was not thought to be the correct gene and that about 70% of the DNA coding sequence for the correct gene had been determined. During a meeting on 10 August 1994 with Jeffrey Ihnen, it was decided to proceed with the filing of an application to claim the gene sequence which was then known as well as the predisposing mutations which had been discovered. It was also decided that the application would be filed on 12 August 1994. This application was filed on 12 August 1994 under Serial No. 08/289,221 ('221 application).

(6) Inventorship of the breast cancer gene was briefly discussed during the 10 August 1994 meeting. Based on the available information, it was determined that the inventorship for the breast cancer gene was as set forth in the transmittal letter for the '221 application. That is, it was determined that Roger Wiseman and Andrew Futreal were not joint inventors of the breast cancer gene, since they did not develop the entire gene sequence (then known) and had not discovered any mutations in the gene.

(7) On 29 August 1994, Mark Skolnick advised Jeffrey Ihnen that the remainder of the DNA coding sequence for the breast cancer gene had been determined and that a manuscript would be submitted to Science on 2 September 1994. It was decided that a continuation-in-part (CIP) application would be filed by that date. Since the remainder of the DNA sequence had been determined at Myriad, it was determined that the inventorship for the CIP

application was the same as for the '221 application. This application was filed on 2 September 1994 under Serial No. 08/300,266 ('266 application).

(8) On 15 September 1994, after the story of the discovery of the breast cancer gene broke in the press, Jeffrey Ihnen learned that Science was releasing copies of the manuscript to anyone who requested a copy. The manuscript which was released was the manuscript submitted on 2 September 1994. Mark Skolnick advised Jeffrey Ihnen that a revised manuscript had been submitted which included additional 5' DNA sequence information and additional mutations in the DNA coding sequence. It was then decided to file a continuation-in-part application on 16 September 1994 to avoid loss of rights if Science released the revised manuscript. Since the additional 5' DNA sequence information and disclosed mutations had been determined at Myriad, it was determined that the inventorship for this CIP application was the same as the '221 and '266 applications.

(9) The present application was filed on 16 September 1994 without a Declaration, pursuant to the provisions of 37 CFR §1.53(d), naming us as joint inventors.

(10) On 5 October 1994, following the filing of the present application, NIEHS claimed that Wiseman and Futreal were joint inventors of the application, since claim 3 in the application was directed to fragments of the breast cancer gene. On 28 November 1994 NIEHS provided information from its research records describing the identification of fragments of coding sequences by Wiseman and Futreal from genetic materials obtained from Myriad. After a thorough analysis was made of all of the pertinent research records of the University and Myriad and of the information provided by NIEHS, it was determined that Wiseman and Futreal had been involved in the identification of two DNA fragments which are part of the breast cancer gene. After the facts were completely analyzed, it was recognized that Wiseman and Futreal are joint inventors as to a fragment claim and had been inadvertently omitted, without any deceptive intent, from the application as filed under 37 CFR §1.53(d) and should be properly joined as coinventors of the claimed subject matter.

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.

Date

10/23/95

Date

Mark H. Skolnick

David E. Goldgar
David E. Goldgar

Date

Yoshio Miki

Date

Jeff Swenson

Date

Alexander Kamb

Date

Keith D. Harshman

Date

Donna M. Shattuck-Eidens

Date

Sean V. Tavtigian



UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
ASSISTANT SECRETARY AND COMMISSIONER
OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

Paper No. 8

COPY MAILED

MAR 1 1996

**OFFICE OF PETITIONS
AND PATENTS**

Venable, Baetjer, Howard & Civiletti
1201 New York Avenue, N.W., Suite 1000
Washington, D.C. 20005

In re Application of :
Mark H. Skolnick, et al :
Application No. 08/308,104 : DECISION ON PETITION UNDER
Filed: September 16, 1994 : 37 CFR 1.48(a)
Attorney Docket No.24884-109347-3 :

This is a decision on a February 13, 1995 petition under 37 CFR 1.48(a), requesting correction of the inventorship and on the paper of October 26, 1995 filed in support thereof, which timely responds to an August 28, 1995 Requirement for Information as supplemented on September 29, 1995.

The §1.48(a) petition to correct inventorship is Granted.

The Requirement for Information as supplemented required a verified statement of facts and § 1.63 declaration from original named inventor Goldgar, which were submitted on October 26, 1995.

The October 26, 1995 verified statement of facts is in accord with the statements previously submitted by the other original named inventors.

Based on the file record as a whole and the facts as set forth in the Requirement for Information and in the statements now of record, the error in inventorship occurred without deceptive intent and was diligently corrected.

The application is being returned to the Initial Patent Examination Division of the Office of Initial Patent Examination for further processing, including issuance of a filing receipt,

Application No. 08-08,104

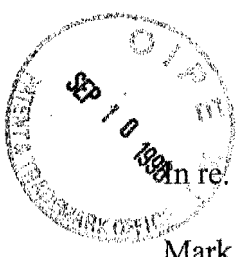
Page 2

with the names of the inventors as shown on the executed declarations under 37 CFR 1.63 filed on February 13, 1995.

A handwritten signature in dark ink, appearing to read 'H. Bernstein', with a stylized, flowing script.

Hiram H. Bernstein
Senior Legal Advisor
Special Program Law Office
Office of the Deputy Assistant Commissioner
for Patent Policy and Projects

(703) 305-9285



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. patent application of:

Mark H. SKOLNICK et al.

Serial No. 08/308,104

Examiner:

Filed: 16 September 1994

Group Art Unit: 1804

For: 17Q-LINKED BREAST AND OVARIAN
CANCER SUSCEPTIBILITY GENE



EXPRESS ABANDONMENT

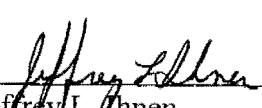
Hon. Commissioner of Patents
and Trademarks
Washington, D.C. 20231

Sir:

Applicant hereby expressly abandon the above-identified application, as of the date
given hereinbelow.

RECEIVED
NOV 14 1996
OFFICE OF THE COMMISSIONER OF PATENTS
AND TRADEMARKS

Respectfully submitted


Jeffrey L. Ahnen
Registration No. 28,957

VENABLE, BAETJER, HOWARD & CIVILETTI
1201 New York Avenue, N.W., #1000
Washington, D.C. 20005
(202) 962-4810

Atty. Docket No. 24884-109347-02
Dated: 09 September 1996

RECEIVED
SEP 12 1994
GROUP 1804



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
--------------------	---------------------	-----------------------	---------------------------

08/308,104

09/16/94

0212/0512

FIRST NAMED APPLICANT

ATTORNEY DOCKET NO./TITLE

VENABLE, BAETJER, HOWARD & CIVILETTI
SUITE 1000
1201 NEW YORK AVENUE, N.W.
WASHINGTON DC 20005

NOT ASSIGNED

0000

05/12/98

DATE MAILED:

NOTICE OF ABANDONMENT UNDER 37 CFR 1.53(f) or (g)
(Filing Date Granted)

The above-identified application is abandoned for failure to timely or properly reply to the Notice to File Missing Parts (Notice) mailed on _____.

- ☐ No reply was received.
- ☐ The reply received on _____ was untimely.
- ☐ The reply received on _____ was improper. The reply did not include:
- ☐ 1. The surcharge under 37 CFR 1.16(e) required for filing the basic filing fee or oath or declaration on a date later than the filing date of a nonprovisional application.
 - ☐ 2. The surcharge under 37 CFR 1.16(i) required for filing the basic filing fee or cover sheet on a date later than the filing date of a provisional application.
 - ☐ 3. The full amount of the basic filing fee under 37 CFR 1.16 (a), (f), (g), (k).
(Note: A nonprovisional application may not be relied on for benefits under 35 U.S.C. 120 and 37 CFR 1.78 unless the processing and retention fee set forth in 37 CFR 1.21(l) is paid within the one year period set forth in 37 CFR 1.53(f). A provisional application may not be relied on for benefits under 35 U.S.C. 119(e) and 37 CFR 1.78 unless the basic filing fee is paid).
 - ☐ 4. The oath or declaration of all the inventors required under 37 CFR 1.63 for this nonprovisional application.
- ☒ The letter of Express Abandonment filed on Sept. 10, 1996 is acknowledged; however, the application is abandoned for failure to timely or properly reply to the Notice as indicated above.

A petition to the Commissioner under 37 CFR 1.137 may be filed requesting that the application be revived.

Under 37 CFR 1.137(a), a petition requesting the application be revived on the grounds of **UNAVOIDABLE DELAY** must be filed promptly after the applicant becomes aware of the abandonment and such petition must be accompanied by: (1) a statement that the entire delay was unintentional; (2) the required reply to the above-identified Notice; (3) the petition fee set forth in 37 CFR 1.17(m); and (4) a terminal disclaimer if required by 37 CFR 1.137(c).

Under 37 CFR 1.137(b), a petition requesting the application be revived on the grounds of **UNINTENTIONAL DELAY** must be filed promptly after applicant becomes aware of the abandonment and such petition must be accompanied by: (1) a statement that the entire delay was unintentional; (2) the required reply to the above-identified Notice; (3) the petition fee set forth in 37 CFR 1.17(m); and (4) a terminal disclaimer if required by 37 CFR 1.137(c).

Any questions concerning petitions to revive should be directed to "Office of Petitions" at (703) 305-9282.

C. Barnes
Customer Service Center
Initial Patent Examination Division (703) 308-1202

File History Report

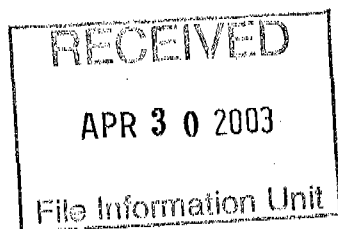


Paper number 10 is missing from the United States Patent and Trademark Office's original copy of the file history. No additional information is available.



The following page(s) _____ of paper number _____ is/are missing from the United States Patent and Trademark Office's original copy of the file history. No additional information is available

Additional comments: _____

REQUEST FOR ACCESS TO AN APPLICATION UNDER 37 CFR 1.14(e)

In re Application of

Application Number

09/308104

Filed

9-9-99

Art Unit

Examiner

Paper No. #11

Assistant Commissioner for Patents
Washington, DC 20231

1. ☒ Thereby request access under 37 CFR 1.14(e)(2) to the application file record of the above-identified ABANDONED Application, which is not within the file jacket of a pending Continued Prosecution Application (CPA) (37 CFR 1.53(d)) and is: (CHECK ONE)

☒ (A) referred to in:

United States Patent Application Publication No. _____, page _____, line _____.

United States Patent Number 5693 423, column _____, line _____, or

an International Application which was filed on or after November 29, 2000 and which

designates the United States, WIPO Pub. No. _____, page _____, line _____.

☐ (B) referred to in an application that is open to public inspection as set forth in 37 CFR 1.11(b) or

1.14(e)(2)(i), i.e., Application No. _____, paper No. _____, page _____, line _____.

2. ☐ I hereby request access under 37 CFR 1.14(e)(1) to an application in which the applicant has filed an authorization to lay open the complete application to the public.

Henry Nung
Signature

4-30-03

Date

HENRY NUNG
Typed or printed name

FOR PTO USE ONLY

Approved by: gm

(initials)

Unit: PTO

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

PATENT APPLICATION FEE DETERMINATION RECORD

Effective October 1, 1994

Application or Docket Number

08/308104

CLAIMS AS FILED - PART I

(Column 1)

(Column 2)

SMALL ENTITY

OR

OTHER THAN
SMALL ENTITY

FOR	NUMBER FILED	NUMBER EXTRA
BASIC FEE		
TOTAL CLAIMS	74 minus 20 = *	54
INDEPENDENT CLAIMS	21 minus 3 = *	18
MULTIPLE DEPENDENT CLAIM PRESENT		

RATE	FEE
	365.00
x\$11=	
x38=	
+120=	
TOTAL	

OR

OR

OR

OR

OR

RATE	FEE
	730.00
x\$22=	1188
x76=	1368
+240=	
TOTAL	3286

* If the difference in column 1 is less than zero, enter "0" in column 2

CLAIMS AS AMENDED - PART II

(Column 1)

(Column 2)

(Column 3)

SMALL ENTITY

OR

OTHER THAN
SMALL ENTITY

AMENDMENT A	CLAIMS REMAINING AFTER AMENDMENT	HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA
Total	*	Minus **	=
Independent	*	Minus ***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM			

RATE	ADDITIONAL FEE
x\$11=	
x38=	
+120=	
TOTAL ADDIT. FEE	

OR

OR

OR

OR

RATE	ADDITIONAL FEE
x\$22=	
x76=	
+240=	
TOTAL ADDIT. FEE	

AMENDMENT B	CLAIMS REMAINING AFTER AMENDMENT	HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA
Total	*	Minus **	=
Independent	*	Minus ***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM			

RATE	ADDITIONAL FEE
x\$11=	
x38=	
+120=	
TOTAL ADDIT. FEE	

OR

OR

OR

OR

RATE	ADDITIONAL FEE
x\$22=	
x76=	
+240=	
TOTAL ADDIT. FEE	

AMENDMENT C	CLAIMS REMAINING AFTER AMENDMENT	HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA
Total	*	Minus **	=
Independent	*	Minus ***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM			

RATE	ADDITIONAL FEE
x\$11=	
x38=	
+120=	
TOTAL ADDIT. FEE	

OR

OR

OR

OR

RATE	ADDITIONAL FEE
x\$22=	
x76=	
+240=	
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.

PACE DATA ENTRY CODING SHEET

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

1ST EXAMINER

DATE

2ND EXAMINER

DATE

APPLICATION NUMBER

TYPE
APPL

FILING DATE

MONTH DAY YEAR

SPECIAL
HANDLING

GROUP
ART UNIT

CLASS

SHEETS OF
DRAWING

08/308104

1

09/16/94

0

1804

435

10

TOTAL
CLAIMS

INDEPENDENT
CLAIMS

SMALL
ENTITY?

FILING FEE

FOREIGN
LICENSE

ATTORNEY DOCKET NUMBER

74

21

Y

248841093473

CONTINUITY DATA

CONT STATUS
CODE CODE

PARENT APPLICATION
SERIAL NUMBER

PCT APPLICATION SERIAL NUMBER

PARENT PATENT
NUMBER

PARENT FILING
DATE

MONTH DAY YEAR

03
13
13

08300266
08289221
08246602

PCT I
PCT I
PCT I
PCT I
PCT I

090294
081294
052094

PCT/FOREIGN APPLICATION DATA

FOREIGN
PRIORITY
CLAIMED

COUNTRY
CODE

PCT/FOREIGN APPLICATION SERIAL NUMBER

FOREIGN
FILING DATE

MONTH DAY YEAR