	Class Subclass Subclass Seve classification			PATENT	ADARIDOI	, VED	
SERIAL NUMBE		and the second sec	SUBCLASS	6	GROUP ART UNIT	EXAMINER Rees	
CONTINU VERIFIEI	ふ (H# Kito WF WF	TOKYO, JAPAM ********** PLN IS A CIP TCH IS A CIP	4: YÚGUKE ****** * OF 08/ * OF 08/ * OF 08/ * OF 08/	NAKAMURA (348,824 (308,104 (300,266	UES SIMARD, , YOKOHAMA, 11/29/94 09/16/94 09/02/94 ABN 08/12/94	QUEBEC. JAPAN.	in states constants of second states and second states and second states and
VERIFIEI NONE J OREIGN I	NG LICENS	E GRANTED 0:		TOTAL	EP. FILING FEE	ATTORNEY'S	and the second se
	Wedged Examiner's Inf L IHNEN BAETJER HOWA V YORK AVENUE	RD & CTVILE			MRS RECEIVED	DOCKET NO.	34
PARTS OF AL FILED SEPAR NOTICE OF A IS Amount Due	PPLICATION NATELY NULOWANCE MAILEE SUEFEE	Assistant Examine	r Prima RED FOR ISSUE	U.S.	DEPT. OF COMM./ PAT.	ons Examiner ALLOWED Print Claim VING IVIG Print Fig Print Fig Osure may be prohibite	

語のなどの語言を

i

O.

	meted in group 1800	
18 409305	PATENT APPLICATION	TIALS APR 1 3 9571
Date Entered	08409305	Date Received
Or Counted	CONTENTS	or Mailed
	ABANDONED	
1. Applicatio	on papers.	alabar
2. <u>-</u> 3. Rau	- Seg Listing (ok)	<u>010775</u>) 9-21-95
4. RAL	Dec. + Jees	9-27-95
5. <u>A</u>	ly free	1-23-96
	ection 3 mas,	7-3-96 1
8.	DOD apand.	9-11-9600
9. <u>[[it</u>	tre of alm	9-19-96
	t La men	12-897
1212		1-16-98
	uest pr access	1-21-88
14. <u>11. []</u> 15	uasi () or unes-	
16		
17		
18 19		
20		
21		
22 23	· · · · · · · · · · · · · · · · · · ·	
24		
25 26		
20 27		
28		
29 		
30 31		
32	(FRONT)	GeneDX 1010, pg. 2

FILING DATE 03/24/95 , SALT LAKE CITY, YO, JAPAN; YUSUKE	CLASS 435	GROUP ART UNIT
, SALT LAKE CITY,	435	1804
IS A CIP OF 08/ IS A CIP OF 08/ IS A CIP OF 08/ IS A CIP OF 08/ NS*****	NAKAMURA, YOKO 348,824 11/29/ 308,104 09/16/ 300,266 09/02/ 289,221 08/12/	HAMA, JAPAN. 94 94 94 ABN 94
INDEPENDENT CLAIMS	FILING FEE RECEIVED	ATTORNEY DOCKET NO.
2	\$880.00	24884-109347
& CIVILETTI SUITE 1000 VARIAN CANCER SUSC	CEPTIBILITY GEN	E
o is a true copy from th application which is ide	ne records of the L entified above.	Inited States
	IS A CIP OF 08/ IS A CIP OF 08/ IS A CIP OF 08/ NS************************************	IS A CIP OF 08/300,266 09/02/ IS A CIP OF 08/289,221 08/12/ IS A CIP OF 08/246,602 05/20/ NS************************************



DB 409305

----- ----- *-*---

!.

PATENT APPLICATION SERIAL NO.

41

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

PTO-1556 (5/87)

08 409305

Sir 24 1995 Transitted her 1th for filing is the patent application of: Inventor 5 Donna M <u>SHATTUCK-EIDENS</u>, Jacques SIMARD, Mitsuru EMI, and Yusuke NAKAMURA

Patents and Trademar

20231

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

Enclosed are:

Commi

Wash

[X] <u>nineteen (19)</u> sheet(s) of drawings (10 figures), in triplicate.

[] Assignment of the invention (with cover sheet) to _____

[X] 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).

OTHER THAN

	(COL. 1)	(COL. 2)	SMALI	, ENTITY		SMAL	L ENTITY
FOR:	NO. FILED	EXTRA	RATE	FEE		RATE	FEE
BASIC FEE:	 			\$365.00			\$730.00
TOTAL CLAIMS:	18-20		X \$ 11	\$	OR	X \$ 22	\$
INDEP CLAIMS:	2-3	1	X \$ 38	\$	OR	X\$76	\$
MULTIPLE DEPE	NDENT CLAIM	PRESENTED:	+ \$120	\$	 OR	+ \$240	\$
			TOTAL	\$		TOTAL	\$730.00

- [] A check in the amount of \$_____ to cover the filing fee and assignment recordation fee is attached.
- [] The Commissioner is authorized to charge the following fees associated with this communication or credit any overpayment to Deposit Account No. 22-0261. Duplicate copies are attached.

[] Additional filing fees required under 37 CFR 1.16.[] Patent application processing fee under 37 CFR 1.17.

Respectfully submitted,

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

VENABLE,	BAETJER,	HOWARD &	CIVILETTI	
			Suite 1000	
Washingto	on, D.C.	20005		
202-962-	4810			

Atty. Docket No. 24884-109347-5 Dated: 24 March 1995 1284.01

AN HELP



TITLE OF THE INVENTION

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE MUTATIONS

UH 109305

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of application Serial No. 08/348,824 filed on 29 November 1994, which is a continuation-in-part of application Serial No. 08/308,104 filed on 16 September 1994, which 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, all incorporated herein by reference.

FIELD OF THE INVENTION

15

5

0

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 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 breast and ovarian 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 List of References.

20

30

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. Fewer than a dozen tumor suppressor genes have been identified, but the number is expected to increase beyond fifty (Knudson, 1993).

-2-

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).

. 10

15

5

20

25

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.

-3-

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.

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).

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.

10

15

5

20

25

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. (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 17q2l; 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).

25

3.0

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.

5

10

15

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 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 base pairs) bounded by Mfdl5 (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.

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

25

30

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

10

15

20

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.

-6-

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 cysteines and one histidine that form the zinc ion binding pocket.

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 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

15

10

5



25

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 variable. Molecular weight standards are in kilobases.

-7-

Figure 8 is a diagram of the 5' untranslated region plus the beginning of the translated region of BRCA1 showing the locations of introns and the variants of BRCA1 mRNA produced by alternative splicing. Intron locations are shown by broken dashed lines. Six alternate splice forms are shown.

Figure 9A shows a nonsense mutation in Kindred 2082. P indicates the person originally screened, b and c are haplotype carriers, a, d, e, 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 9B 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 9C 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.

Figures 10A-10H show genomic sequence of BRCA1. 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 vvvvvvvvvvvvvvv. The intron between exons 6 and 7 is not included. Hyphens (-) are meaningless and should be ignored.

10

15

5

20

25

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 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 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.

-8-

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 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 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.

5

10

15

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.

-9-

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 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 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.

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*.

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.

10

15

5

20

30

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 base pairs, a region which contains a genetic locus, BRCA1, which causes susceptibility to cancer, including breast and ovarian cancer, has been identified.

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.

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 17qlinked 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.

10

5

15

20

30

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 ageadjusted 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.

15

20

25

30

10

5

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.

-12-

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.*, 1994).

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 Mfdl5. Therefore, BRCA1 was shown to be in a region of 6 to 10 million bases bounded by Mfdl5 and 42D6. Marker Mfd191 was subsequently discovered to be distal to Mfd15 and proximal to BRCA1. Thus, Mfdl5 was replaced with Mfd191 as the closest proximal genetic marker. Similarly, it was discovered that genetic marker Mfdl88 could replace genetic marker 42D6, narrowing the region containing the BRCA1 locus to approximately 1.5 million bases. Then the marker Mfdl91 was replaced with tdj1474 as the proximal marker and Mfdl88 was replaced with USR 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.

10

5

20

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.

-13-

<u>Yeast Artificial Chromosomes (YACs)</u>. 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 C_0 t-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

15

10

.

20

25

3.0

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.

-14-

Coverage of the Gap with Pl 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.

15 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.

25

20

(a) <u>Zoo blots</u>. 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 (Clonetech, Cat. 7753-1).



(b) <u>Identifying HTF islands</u>. 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

10

HTF (HpaI tiny fragment) or CpG islands, as restriction enzymes which contain CpG dimers cut frequently in these regions (Lindsay *et al.*, 1987).

-15-

(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.*, 1991) 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) <u>Hybridizing cDNA to Cosmids, P1s, BACs or YACs</u>. 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) <u>Identification of cDNAs</u>. 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 adapters 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 adapters. These adapters serve as amplification sites for PCR. The target and probe sequences are denatured and mixed with human C_0t-1 DNA to block repetitive sequences.

10

15

5

20

30

Solution hybridization is carried out to high $C_0t-1/2$ 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.

-16-

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 then 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.

25

30

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

10

15

20

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 BRCA1 gene product, or to a decrease in mRNA stability or translation efficiency.

-17-

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.

There are several methods that can be used to detect DNA sequence variation. Direct DNA sequencing, either manual sequencing or automated fluorescent sequencing can detect sequence variation. For a gene as large as BRCA1, manual sequencing is very labor-intensive, but under optimal conditions, mutations in the coding sequence of a gene are rarely missed. Another approach is the single-stranded conformation polymorphism assay (SSCA) (Orita *et al.*, 1989). This method does not detect all sequence changes, especially if the DNA fragment size is greater than 200 bp, but can be

10

15

5

20

30

optimized to detect most DNA sequence variation. The reduced detection sensitivity is a disadvantage, but the increased throughput possible with SSCA make it an attractive, viable alternative to direct sequencing for mutation detection on a research basis. The fragments which have shifted mobility on SSCA gels are then sequenced to determine the exact nature of the DNA sequence variation. Other approaches based on the detection of mismatches between the two complementary DNA strands include clamped denaturing gel electrophoresis (CDGE) (Sheffield *et al.*, 1991), heteroduplex analysis (HA) (White *et al.*, 1992) and chemical mismatch cleavage (CMC) (Grompe *et al.*, 1989). None of the methods described above will detect large deletions, duplications or insertions, nor will they detect a regulatory mutation which affects transcription or translation of the protein. Other methods which might detect these classes of mutations such as a protein truncation assay or the asymmetric assay, detect only specific types of mutations and would not detect missense mutations. A review of currently available methods of detecting DNA sequence variation can be found in a recent review by Grompe (1993). Once a mutation is known, an allele specific detection approach such as allele specific oligonucleotide (ASO) hybridization can be utilized to rapidly screen large numbers of other samples for that same mutation.

-18-

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

5

10

20

15

30

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.

-19-

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); 5 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

15

10

25

30

. .

-20-

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.

10

15

20

25

30

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.

-21-

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 wildtype 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

.

15

20

25

30

10

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.

-22-

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.

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.

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 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.

10

15

20

5

25

Definitions

The present invention employs the following definitions:

"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.*, 1989a (for LCR). Reagents and hardware for conducting PCR are commercially available. Primers useful to 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 sequenced 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.

-23-

"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.

"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 several months into rabbits. Rabbit sera is tested for immunoreactivity to the BRCA1 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.

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

10

15

5



30

or rabbit, is selected. Substantially purified antigen is presented to the immune 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.

-24-

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

10

15

5



30

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.

-25-

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

10

15

5



30

.

-26-

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 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.

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.), pendent 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.

25

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.

10

15

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.

-27-

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 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.

30

25

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.

10

15

20

-28-

"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.

30

5

10

15

20

25

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,

5

15

20

25

30

10

specificity.

may be used to isolate other polynucleotide sequences, having sequence similarity by standard methods. For techniques for preparing and labeling probes see, e.g., Sambrook et al., 1989 or Ausubel 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 labeled 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 Such modifications include, for example, acetylation, carboxylation, unusual amino acids. phosphorylation, glycosylation, ubiquitination, labeling, 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 ³²P, ligands, which bind to labeled antiligands (e.g.,

-29-

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

The probes will include an isolated polynucleotide attached to a label or reporter molecule and

5

Ausubel et al., 1992.

10

15

20

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.

30

25

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.

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 labeling polypeptides are well known in the art. See, e.g., Sambrook et al., 1989 or

"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.

-31-

The terms "isolated", "substantially pure", and "substantially homogeneous" 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

10

15

5

20

25
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.

-32-

"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.

10

15

5

20

30

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.

-33-

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 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.

10

15

5

.

20

30

-34-

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, or 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 and Lalouel, 1988.

15 <u>Preparation of recombinant or chemically synthesized</u> nucleic acids: vectors, transformation, host cells

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 Ausubel *et al.*, 1992.

25

20

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 strands together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

10

5

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 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 Ausubel *et al.* 1992.

-35-

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 Ausubel 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 Moloney 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).

30

5

10

15

20

-36-

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, Jakoby and Pastan, 1979. 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

10

5

15



30

skilled practitioner that other cell lines may be appropriate, e.g., to provide higher expression, desirable glycosylation patterns, or other features.

-37-

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.

10

15

5

20

30

Initially, the screening method involves amplification of the relevant BRCA1 sequences. In another preferred embodiment of the invention, the screening method involves a non-PCR based strategy. Such screening methods include two-step label amplification methodologies that are well known in the art. Both PCR and non-PCR based screening strategies can detect target sequences with a high level of sensitivity.

-38-

The most popular method used today is target amplification. Here, the target nucleic acid sequence is amplified with polymerases. One particularly preferred method using polymerase-driven amplification is the polymerase chain reaction (PCR). A preferred PCR based strategy contemplated within the scope of this invention is provided in Example 8. The polymerase chain reaction and other polymerase-driven amplification assays can achieve over a million-fold increase in copy number through the use of polymerase-driven amplification cycles. Once amplified, the resulting nucleic acid can be sequenced or used as a substrate for DNA probes.

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.*,

10

5

20

15

30

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.

-39-

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.

As noted above, non-PCR based screening assays are also contemplated in this invention. An exemplary non-PCR based procedure is provided in Example 11. This procedure hybridizes a nucleic acid probe (or an analog such as a methyl phosphonate backbone replacing the normal phosphodiester), to the low level DNA target. This probe may have an enzyme covalently linked to the probe, such that the covalent linkage does not interfere with the specificity of the hybridization. This enzyme-probeconjugate-target nucleic acid complex can then be isolated away from the free probe enzyme conjugate and a substrate is added for enzyme detection. Enzymatic activity is observed as a change in color development or luminescent output resulting in a 10^3 - 10^6 increase in sensitivity. For an example relating to the preparation of oligodeoxynucleotide-alkaline phosphatase conjugates and their use as hybridization probes see Jablonski *et al.*, 1986.

Two-step label amplification methodologies are known in the art. These assays work on the principle that a small ligand (such as digoxigenin, biotin, or the like) is attached to a nucleic acid probe capable of specifically binding BRCA1. Exemplary probes are provided in Table 9 of this patent application and additionally include the nucleic acid probe corresponding to nucleotide positions 3631 to 3930 of SEQ ID NO:1. Allele specific probes are also contemplated within the scope of this example and exemplary allele specific probes include probes encompassing the predisposing mutations summarized in Tables 11 and 12 of this patent application.

5

10

20

15

30

In one example, the small ligand attached to the nucleic acid probe is specifically recognized by an antibody-enzyme conjugate. In one embodiment of this example, digoxigenin is attached to the nucleic acid probe. Hybridization is detected by an antibody-alkaline phosphatase conjugate which turns over a chemiluminescent substrate. For methods for labeling nucleic acid probes according to this embodiment see Martin *et al.*, 1990. In a second example, the small ligand is recognized by a second ligand-enzyme conjugate that is capable of specifically complexing to the first ligand. A well known embodiment of this example is the biotin-avidin type of interactions. For methods for labeling nucleic acid probes and their use in biotin-avidin based assays see Rigby, *et al.*, 1977 and Nguyen, *et al.*, 1992. Example 11 details a method using a biotin-avidin type assay employing BRCA1.

It is also contemplated within the scope of this invention that the nucleic acid probe assays of this invention will employ a cocktail of nucleic acid probes capable of detecting BRCA1. Thus, in one example to detect the presence of BRCA1 in a cell sample, more than one probe complementary to BRCA1 is employed and in particular the number of different probes is alternatively 2, 3, or 5 different nucleic acid probe sequences. In another example, to detect the presence of mutations in the BRCA1 gene sequence in a patient, more than one probe complementary to BRCA1 is employed where the cocktail includes probes capable of binding to the allele-specific mutations identified in populations of patients with alterations in BRCA1. In this embodiment, any number of probes can be used, and will preferably include probes corresponding to the major gene mutations identified as predisposing an individual to breast cancer. Some candidate probes contemplated within the scope of the invention include probes that include the allele-specific mutations identified in Tables 11 and 12 and those that have the BRCA1 regions corresponding to SEQ ID NO:1 both 5' and 3' to the mutation site.

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. The antibodies may be prepared as discussed above under the heading "Antibodies" and as further shown in Examples 12 and 13. Other techniques for raising and purifying antibodies are well known in the art and any such techniques may be chosen to achieve the preparations claimed in this invention. In a preferred embodiment of the invention, antibodies will immunoprecipitate BRCA1 proteins from solution as well as react with BRCA1 protein

-40-

10

5

20

25

30

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.

-41-

Preferred embodiments relating to methods for detecting BRCA1 or its mutations include enzyme linked immunosorbent assays (ELISA), radioimmunoassays (RIA), immunoradiometric assays (IRMA) and immunoenzymatic assays (IEMA), including sandwich assays using monoclonal and/or polyclonal antibodies. Exemplary sandwich assays are described by David *et al.*, in U.S. Patent Nos. 4,376,110 and 4,486,530, hereby incorporated by reference, and exemplified in Example 14.

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 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 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 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, PCT published application WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or

10

5



15

30

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.

-42-

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 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 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

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 modeling 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.

10

15

20

5

30

5

10

20

25

3.0

-43-

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-idio-typic 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.

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.

15 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.

GeneDX 1010, pg. 48

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.

-44-

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.*, 1995; 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;

.

10

15

5

20

30

-45-

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.*, 1989b; Wolff *et al.*, 1991; Wagner *et al.*, 1990; Wagner *et al.*, 1991; Cotten *et al.*, 1990; Curiel *et al.*, 1991a; Curiel *et al.*, 1991b). 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.

30

5

10

15

20

25

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.

-46-

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.

25 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,

10

15



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.

-47-

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.

30

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

10

15

20

25

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.

-48-

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 cancer. Four women diagnosed with both ovarian and breast cancer are counted in both categories.

.

15

10

TABLE 1

Description of the 19 Kindreds

	· · · · · · · · · · · · · · · · · · ·	- ·	•	Breas	st	·		Ovaria	in		
		of <u>riduals</u>		A	ge at	Dx			Age at	: Dx	-
KINDRED	<u>Total</u>	Sample	<u># Aff.</u>	<u>Min.</u>	Med.	Max.	<u># Aff,</u>	Min.	Med.	Max.	
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 -			-	<u> </u>	
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	-		JZ -	-	
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	-		-	/1	
1917	16	6	4	43	58	61	-	-	-	-	
1920	22	14	3	62	63	68	<u> </u>	_	_		
2079	136	18	14	38	66	84	4	52	- 59	65	
2039	87	40	14	44	68	88	4	41	51	65 75	

+ Includes one case of male breast cancer.

-49-

EXAMPLE 2

-50-

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., 1991), 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).

25

20

5

10

15

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 and Lalouel, 1988). Table 2 shows the results of the pairwise

linkage analysis of each kindred with the three markers 42D6, Mfdl88, and Mfdl5.

-51-

GeneDX 1010, pg. 56

		Pairw:	ise Linkage	Analysis of K	indreds			
		Mfd15 (1 Recomb	D17S250) ination	42D6 (D1 <u>Recombi</u>		Mfd188 (<u>Recombi</u>		
KINDRED		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	
				•				- 52 -
1929		-0.49	-0.25	NT	NT	-0.49	-0.25	N . 1
1901		1.50	1.17	0.78	0.57	0.65	0.37	4
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	

TABLE 2

NT - Kindred not typed for Mfd188.

٠.

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.

-53-

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.

10

15

5

20

25

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, Mfd19l (provided by James Weber), THRA1 (Futreal

15

10

5

25

et al., 1992a), and three polymorphisms kindly provided to us by Dr. Donald Black, NM23 (Hall et al. 1992), SCG40 (D17S181), and 6C1 (D17S293).

-55-

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 the 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 Mfd19l. 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.

10

TABLE 3

<u>CEPH Recombinants</u>

Family	ID	<u>M/P</u>	Mfd15	THRA1	<u>Mfd191</u>	<u>Mfd188</u>	SCG40	_6C1_	<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	Μ	0	0	1	1	. - 1.	-	_	
1334	3	M	1	1	1	1	1	0	0	
1333	4	M	1	1	1	0	-	_	0	
1333	6	Μ	0	0	1	1	-	-	1	
1333	8	Ρ	1	0	0	0	-	- -	0	
1377	8	М	0	-	0	0	0	0	1	
					•					

-56-

Family	ID	M/P	Mfd15	<u>Mfd191</u>	<u>Mfd188</u>	SCG40	<u>6C1</u>	<u>42D6</u>	
75		М	0	1	1	1	· _ ·	. 	
63		M	0	0	1	1	- -	1	
125		M	1	1	1	0	-	0	
4.0		Μ	1	1	- 0	· 0	-	0	

TABLE 4

Kindred 2082 Recombinants

GeneDX 1010 pg. 62

-57

-58-

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 Mfd19l-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 $(GT)_{10}T$. 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

10

15

5

20

30

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.

-59-

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.

10

TABLE 5

Polymorphic	Short	Tandem	Repeat	Markers	Used
for Fine Str	ucture	Mappin	<u>g of th</u>	e BRCA1	Locus

				<u></u>	Al	<u>lele' I</u>	reque	ncy (%)
<u>Clone</u>	Gene	<u>Na**</u>	Hetero- zygosity	<u>1</u>	2	<u>3</u>	_4	_5	_6
Mfd15 THRA1	D17S250 THRA1	10 5	0.82	26	22	15	7	7	23
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.

-60-

5

The four STR polymorphisms which mapped physically to the region (4-7, ED2, AA1, YM29) were analyzed in the meiotic, breakpoint panel shown initially in Tables 3 and 4. Tables 6 and 7 contain the relevant CEPH data and Kindred 2082 data for localization of these four markers. In the tables, 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 6
Key Recombinants Used for Genetic
Ordering of New STR Loci Developed in

 Our Laborator	<u>cy Within</u>	the	BRCA1	Region	of 17a
				and the second secon	

	CEPH													
•	Family	ID	<u>M/P</u>	Mfd15	THRA1	<u>Mfd191</u>	<u>ED2</u>	<u>AA1</u>	<u>Z109</u>	<u>4-7</u>	<u>YM29</u>	<u>Mfd188</u>	<u>SCG40</u>	<u>42D6</u>
	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	М	0	0	1	-	1	-	-	1	1	-	1
	1333	3	М	0	0	1	-	-	-	1	1	1	_	1

-62-

TABLE 7

Kindred 2082 Recombinants

ID	<u>M/P</u>	<u>Mfd15</u>	<u>Mfd191</u>	ED2	<u>AA1</u>	<u>4-7</u>	<u>YM29</u>	<u>Mfd188</u>	SCG40	<u>42D6</u>	
.63	М	0	0	1		1	1	1	1	נ	
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	

-63-- --

From CEPH 1333-04, we see that AAI 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.

-64-

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.

5

15

20

10

.

TABLE 8

Breast Cancer Linked Haplotypes Found in the Three Kindreds

Kin.	НАР	Mfd 15	THRA1	Mfd 191	tdj 1475	ED2	AA1	Z109	CA375	4-7	YM29	Mfd 188	SCG40	6C1	42D
				· · · ·							· · · · ·			······································	
1901	H1	1	5	5 5	3	1	4	NI	NI	1	1	3	NI	NI	4
	R2	9	2	5	3 6	1 1	4	NI	NI	1	ĩ	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	ÑI	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	ं ।	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	ī	2	NI	í
	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		8	2	1	NI	5	1	1	4	3	1	6	8	2	4
	H2	8	2	1.	NI	5	1	1	2	1	ĩ	2	8 3 3	ī	4
	R2	8	2	1	NI	5	1	1	2	1	1	2	ĩ	6	1

-65-

-66-

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

10

5

20

15

-67-

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.

25

30

20

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.

. .

5

.

15
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.

-68-

10 Kindred 2082

5

15

20

25

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 centromere - 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

5

10

15

20

25

30

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 lease 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 Simard *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 Simard 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., USR, 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

5

10

20

15

30

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.

-71-

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 Figures 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 cosmids 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 Pl 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 μ g 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

10

5

20

15

NaPO₄ pH 6.8, 100 μ g/ml Salmon Sperm DNA, lx Denhardt's, 50% formamide, 0.1% SDS, and 2 μ g/ml C_ot-1 DNA.

-72-

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., 1982). The eluted Not I digested DNA (~15 kb to 25 kb) was then digested with EcoRI restriction enzyme to give smaller fragments (~O.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-Rl 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 SaII, 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.

10

5

15

20

.

30

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 EcoRl 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.

-73-

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 Clonetech, Inc. The cDNA library generated from breast tissue of an 8 month pregnant woman is available from Clonetech (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 Clonetech 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 brain (Stratagene, Cat. 936206), human testis (Clonetech Cat. HL1027h), human brain (Clonetech Cat. HL11810), human placenta (Clonetech Cat 1075b), and human skeletal muscle (Clonetech 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.* (1982). 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.

30

5

10

15

20

25

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.

-74-

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

15

10

5

25

20

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 Pl, 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.

-75-

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

10

5

20

15

30

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

-76-

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 cDNA (up through the stop codon) 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 (Smith and 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 cysteine 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

10

5

20

15

30

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 six potential alternate splice donors in this 5' untranslated region, with the longest deletion being 1,155 bp. The predominant 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.

-77-

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. There are six alternative splice forms of this 5' sequence as shown in Figure 8. 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. A fourth splice donor site is the G at position 1238. The T at position 1519 represents the splice acceptor site for all of the above splice donors. A fifth alternate splice form has a first splice donor site at position 357 with a first acceptor site at position 599 and a second splice donor site at position 897 and a second acceptor site at position 1519. A sixth alternate form is unspliced in this 5' region. The A at position 1538 is the canonical start site, which appears at position 120 of SEQ ID NO:1. Partial genomic DNA sequences determined for BRCA1 are set forth in Figures 10A-10H and 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 vvvvvvvvvvvv. The intron/exon junctions are shown in Table 9. The intron between exons 6 and 7 has not been included in this figure or in SEQ ID NO:14. Exons 8 and 14 sometimes have an additional CAG found at their start, i.e., this additional codon is found in some cDNAs but not in others. The cDNA shown in SEQ ID NO:1 shows the most prevalent form.

10

15

5

TABLE 9

Base Exon Position				Intron Borders		
	No.	5'	3'	Length	51	3'
	el	: 1	100	99	GATAAATTAAAACTGCGACTGCGCGGCGTG ^{15*}	AAAAATAAAGGACGCTGTCGAYTAGTTCTT ¹⁶
	e2	101	204	103	AAGATATATATATATGTTTTTTTTAATGTGT ¹⁷	GTAAGTCAGCACAAGAGTGTATTAATTTGG ¹⁸
	e3	205	258	53	TTATTGAGCCTCATTTATTTTCTTTTTCTC	¹⁹ ATGTGGCTCCATTATTAGCTTTTGTTTTG ²⁰
	e4	**	**	110	AGCTACTTTTTTTTTTTTTTTGAGACAG ²¹	GTAAGTGCACACCACCATATCCAGCTAAAT ²²
	e5	259	336	77	TTGTTCTTTCTTTÇTTTATAATTTATAGAT	23 GATATAANNNGAANTGATGCTAGGTTGGAA ²⁴
	e6	337	425	88	GATAATCACTTGCTGAGTGTGTTTCTCAAA	25 GTAAGTGTTGAATATCCCAAGAATGAGAGT ²⁶
	e7	426	568	142	AAACATAATGTTTTCCCTTGTATTTTACAG ²⁷	AAACCATTTGTTTTCTTCTTCTTCTTCTTC ²⁰
	e8	569	670	101	TTGACTGTTCTTTACCAWACTGTTWARCAG ²⁹	GGTAAGGGTCTCAGGTTTTTTAAGTATTTA ³⁰
	e9	671	716	45	TTGATTTATTTTTTGGGGGGAAATTTTTTA ³¹	GGTGAGTCAAAGAGAACCTTTGTCTATGAA ³²
	e10	717	793	76	TTCTTATTAGGACTCTGTCTTTTCCCTATA ³³	GGTAATGGCAAAGTTTGCCAACTTAACAGG ³⁴
	e11	794	4220	3426	TGAGTACCTTGTTATTTTTGTATATTTTCA ³⁵	GTATTGGAACCAGGTTTGTGTGTGTGGCCC ³⁶
	e12	4221	4313	92	ACATCTGAACCTCTGTTTTTGTTATTTAAG ³⁷	NGNGAATGTAATCCTAATATTTCNCNCCNA ³⁸
	e 13	4314	4480	166	TCTTGGTACCATNNATCGTTTTTGAATNAG ³⁹	TGTGTATTGTTGGCCAAACACTGATATCTT ⁴⁰
	e14	4481	4605	124	AGATTTGTTTTCTCATTCCATTTAAAGCAG ⁴¹	GTAAGAAACATCAATGTAAAGATGCTGTGG ⁴²
	e15	4606	4798	192	CGATGGTTTTCTCCTTCCATTTATCTTTCT ⁴³	GTAATATTTCATCTGCTGTATTGGARCAMA44
	e16	4799	5109	310	TGGAAGGGAAGCCCACTCNAAAAAANNTTN	45 GTGAGTGTATCCATATGTATCTCCCTAATG ⁴⁶

Numbers in superscript refer to SEQ ID NOS. e4 from SEQ ID NO:11. *

**

78-

TABLE 9 (cont'd)

Exon No.	Ba: Posi 5'		Length	51	Intron Borders
					.
e17	5110	5198	88	ATGATAATGGAATATTTGATTTAATTTCAG ⁴ "	TATACCCAAGAACCTTTACAGAATACCTTK ⁴⁸
e18	5199	5277	78	TAATCCTTTGAGTGTKTCTCATTCTGCAGA ¹⁹	KAAGKATAATACTATWTCTCCCCYCCTCCC ⁵⁰
e19	5278	5316	38	TAACCTGTCTTTTCTATGATCTCTTTAGGG ⁵ 1	GTAAGTACTTGATGTTACAAACTAACCAGA ⁵²
e20	5317	5400	83	TTATCCTGATGGGTGTGNTGTTTCTTTCAG ⁵³	GTAAAGCTCCCTCCAAGTTGACNNAAA54
e21	5401	5455	54	TAATTTGGATTCCCGTCTCGGGTTCAGATC55	GTÄAGAGCCTCGGGAGAANCCAGAGTTCCA ⁵⁶
e22	5456	5530	74	GTGATTTTAGATCCTAAATGTCCATTTTAG ⁵⁷	CAGAGCAAGACCCTGTCTCAAAAACAAACA ⁵⁸
e23	5531	5590	59	TTGAATGCTCTTTCCTTCCTGGGGANCCAG ⁵⁹	TTTATCACTCATTACCTGGTGCTTGAGTAG*
e24	5591	5918	327	CTAATCTCTGCTTGTGTGTTCTCTGTCTCCAG ⁶¹	•

* Numbers in superscript refer to SEQ ID NOS.

-79-

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.

-80-

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

					Spo	radic	
20	Kindred	<u>Cases (n)</u> Br Br<50 Ov			Cases ¹ (n)	LOD Score	Marker(s)
	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
25	1901	10	7	1*	0	1.50	D178855
	1925	4	3	Ō	0	0.55	D178579
	1910	5	4	0	0	0.36	D17S579/D17S250 ²
•	1927	5	4	Ŏ	1	-0.44	D17S250
	1911	8	5	Õ	2	-0.20	D17S250

3,0

35

40

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.

5

10

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

1	Kindred Number	Mutation	Coding Effect	Location*
	2082	C→T	Gln→Stop	4056
25	1910	extra C	frameshift	5385
	2099	T→G	Met→Arg	5443
	2035	?	loss of transcript	
	1901	11 bp deletion	frameshift	189

30 * 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. 9A), Kindred 1910 contains a single nucleotide insertion in coding exon 19 (Fig. 9B), and Kindred 2099 contains a missense mutation in coding exon 20, resulting in a Met \rightarrow Arg substitution (Fig. 9C). The frameshift and

-81-

20

15

5

10

5

10

15

20

25

30

would contain an altered amino acid sequence beginning with the 24th residue from the wildtype Nterminus. The mutant allele in Kindred 2082 would encode a protein missing 548 residues from the Cterminus. 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

The individual studied in Kindred 2035 evidently contains a regulatory mutation in BRCA1. In her cDNA, a polymorphic site (A \rightarrow G at base 3667) appeared homozygous, whereas her genomic DNA revealed heterozygosity at this position (Fig. 9C). 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. 9B). 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. 9A). 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

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

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 extinguished allele in the cDNA sample was shown to lie on the chromosome that carries the BRCA1 predisposing allele (Fig. 9C).

-83-

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 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) 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 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 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 cancer cases diagnosed in their late 20's or

10

15

5

.

25

20

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.

.

5

10

15

20

1.1.1

30

The BRCA1 protein contains a C_3HC_4 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 turnors 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.*, 1992b; 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.

-85-

15

5

10

20

25

A total of six mutations was found, one in an ovarian tumor, four in breast tumors and one in a male unaffected haplotype carrier (Table 12). One mutation, Glu1541Ter, introduced a stop codon that would create a truncated protein missing 273 amino acids at the carboxy terminus. In addition, two missense mutations were identified. These are Ala1708Glu and Met1775Arg and involve substitutions of small, hydrophobic residues by charged residues. Patients 17764 and 19964 are from the same family. In patient OV24 nucleotide 2575 is deleted and in patients 17764 and 19964 nucleotides 2993-2996 are deleted.

TABLE 12

Predisposing Mutations

	Patient	Codon	Nucleotide Change	Amino Acid <u>Change</u>	Age of <u>Onset</u>	Family <u>History</u>
15	BT098	1541	$\underline{G}AG \rightarrow \underline{I}AG$	$Glu \rightarrow Stop$	39	•
	OV24	819	1 bp deletion	frameshift	44	11 - Ali
	BT106	1708	$G\underline{C}G \rightarrow G\underline{A}G$	Ala \rightarrow Glu	24	+
	MC44	1775	$ATG \rightarrow AGG$	Met \rightarrow Arg	42	· · · · +
	17764	958	4 bp deletion	frameshift	31	· · · · · · · · · · · · · · · · · · ·
20	19964	958	4 bp deletion	frameshift		+*

Unaffected haplotype carrier, male

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

all mutations are present in the germline; (i)

all are absent in appropriate control populations, suggesting they are not common (ii) polymorphisms;

30

each mutant allele is retained in the tumor, as is the case in tumors from patients (iii) 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);

the age of onset in the four breast cancer cases with mutations varied between 24 and 42 (iv) years of age, consistent with the early age of onset of breast cancer in individuals with BRCA1 35

5

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,

-87-

(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 (Met1775Arg) 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 Met1775Arg 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

10

5

20

15

30

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.

5

10

15

20

25

A separate study to analyze tumors for BRCA1 mutations was performed in Japan. A panel of 103 patients representing early-onset cases (<35 years of age) (46 patients), members of multiplyaffected families (12 patients), and/or had developed bilateral breast cancers (59 patients) were screened for mutations in BRCA1. Primary breast tumors from these patients were screened for mutations in coding exons of BRCA1 using single-strand conformation polymorphism (SSCP) analysis. For exon 11, which is 3425 bp long, PCR primers were designed to amplify eleven overlapping segments of this exon separately. Each of the other 22 exons was amplified individually in a single PCR. Thus 33 PCR-SSCP analyses were carried out for each case. Mutations were detected in tumors from four patients, all of whom had developed breast cancers bilaterally (Table 12A). One mutation resulted in a frame shift due to a 2 bp deletion (deletion of AA) at codon 797. This gives rise to a truncated protein missing 1065 amino acids at the COOH terminus. A second mutation was a nonsense mutation at codon 1214 due to a $G \rightarrow T$ transversion of the first nucleotide of the codon. This results in a premature stop codon in place of glutamic acid at this site and results in a truncated protein missing 649 amino acids at the COOH terminus. There were also two missense mutations. One was a $G \rightarrow A$ transition at the first nucleotide of codon 271 resulting in a Val \rightarrow Met substitution. The second was at codon 1150 (a C \rightarrow T transition in the first nucleotide of the codon) causing a Pro \rightarrow Ser substitution, a replacement of a hydrophobic nonpolar amino acid with a polar uncharged amino acid. These mutations were all found to be germline mutations. The mean age of onset in these four patients was 49. These studies also found a common neutral polymorphism of either C or T at the first nucleotide of codon 771.

-88-

-89-

TABLE 12A

Predisposing Mutations

			Nucleotide	Amino Acid	Age of
5	Patient	Codon	Change	Change	Onset
	23	1150	CCT→TCT	Pro→Ser	49 & 64
	44	1214	GAG→TAG	Glu→Stop	51 & 51
	98	271	GTG→ATG	Val→Met	45 & 45
	100	797	2bp deletion	frameshift	50 & 71

EXAMPLE 10

Analysis of the BRCA1 Gene

The structure and function of BRCA 1 gene are determined according to the following methods. <u>Biological Studies</u>. 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.

<u>Molecular Genetics Studies</u>. In vitro mutagenesis is performed to construct deletion mutants and missense mutants (by single base-pair substitutions in individual codons and cluster charged \rightarrow 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.

<u>Structural Studies</u>. 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.

15

20

25

10

.

GeneDX 1010, pg. 94

.30

EXAMPLE 11

-90-

Two Step Assay to Detect the Presence of BRCA1 in a Sample

Patient sample is processed according to the method disclosed by Antonarakis, *et al.* (1985), separated through a 1% agarose gel and transferred to nylon membrane for Southern blot analysis. Membranes are UV cross linked at 150 mJ using a GS Gene Linker (Bio-Rad). BRCA1 probe corresponding to nucleotide positions 3631-3930 of SEQ ID NO:1 is subcloned into pTZ18U. The phagemids are transformed into *E. coli* MV1190 infected with M13KO7 helper phage (Bio-Rad, Richmond, CA). Single stranded DNA is isolated according to standard procedures (see Sambrook, *et al.*, 1989).

Blots are prehybridized for 15-30 min at 65°C in 7% sodium dodecyl sulfate (SDS) in 0.5 M NaPO₄. The methods follow those described by Nguyen, *et al.*, 1992. The blots are hybridized overnight at 65°C in 7% SDS, 0.5 M NaPO₄ with 25-50 ng/ml single stranded probe DNA. Post-hybridization washes consist of two 30 min washes in 5% SDS, 40 mM NaPO₄ at 65°C, followed by two 30-min washes in 1% SDS, 40 mM NaPO₄ at 65°C.

Next the blots are rinsed with phosphate buffered saline (pH 6.8) for 5 min at room temperature and incubated with 0.2% casein in PBS for 30-60 min at room temperature and rinsed in PBS for 5 min. The blots are then preincubated for 5-10 minutes in a shaking water bath at 45°C with hybridization buffer consisting of 6 M urea, 0.3 M NaCl, and 5X Denhardt's solution (see Sambrook, *et al.*, 1989). The buffer is removed and replaced with 50-75 μ l/cm² fresh hybridization buffer plus 2.5 nM of the covalently cross-linked oligonucleotide-alkaline phosphatase conjugate with the nucleotide sequence complementary to the universal primer site (UP-AP, Bio-Rad). The blots are hybridized for 20-30 min at 45°C and post hybridization washes are incubated at 45°C as two 10 min washes in 6 M urea, 1x standard saline citrate (SSC), 0.1% SDS and one 10 min wash in 1x SSC, 0.1% Triton[®]X-100. The blots are rinsed for 10 min at room temperature with 1x SSC.

Blots are incubated for 10 min at room temperature with shaking in the substrate buffer consisting of 0.1 M diethanolamine, 1 mM $MgCl_2$, 0.02% sodium azide, pH 10.0. Individual blots are placed in heat sealable bags with substrate buffer and 0.2 mM AMPPD (3-(2'-spiroadamantane)-4-methoxy-4-(3'-phosphoryloxy)phenyl-1,2-dioxetane, disodium salt, Bio-Rad). After a 20 min

10

15

5

20

30

incubation at room temperature with shaking, the excess AMPPD solution is removed. The blot is exposed to X-ray film overnight. Positive bands indicate the presence of BRCA1.

-91-

EXAMPLE 12

Generation of Polyclonal Antibody against BRCA1

Segments of BRCA1 coding sequence were expressed as fusion protein in *E. coli*. The overexpressed protein was purified by gel elution and used to immunize rabbits and mice using a procedure similar to the one described by Harlow and Lane, 1988. This procedure has been shown to generate Abs against various other proteins (for example, see Kraemer, *et al.*, 1993).

Briefly, a stretch of BRCA1 coding sequence was cloned as a fusion protein in plasmid PET5A (Novagen, Inc., Madison, WI). The BRCA1 incorporated sequence includes the amino acids corresponding to #1361-1554 of SEQ ID NO:2. After induction with IPTG, the overexpression of a fusion protein with the expected molecular weight was verified by SDS/PAGE. Fusion protein was purified from the gel by electroelution. The identification of the protein as the BRCA1 fusion product was verified by protein sequencing at the N-terminus. Next, the purified protein was used as immunogen in rabbits. Rabbits were immunized with 100 μ g of the protein in complete Freund's adjuvant and boosted twice in 3 week intervals, first with 100 μ g of immunogen in incomplete Freund's adjuvant followed by 100 μ g of immunogen in PBS. Antibody containing serum is collected two weeks thereafter.

This procedure is repeated to generate antibodies against the mutant forms of the BRCA1 gene. These antibodies, in conjunction with antibodies to wild type BRCA1, are used to detect the presence and the relative level of the mutant forms in various tissues and biological fluids.

EXAMPLE 13

Generation of Monoclonal Antibodies Specific for BRCA1

Monoclonal antibodies are generated according to the following protocol. Mice are immunized with immunogen comprising intact BRCA1 or BRCA1 peptides (wild type or mutant) conjugated to keyhole limpet hemocyanin using glutaraldehyde or EDC as is well known.

10

15

20

25

30

The immunogen is mixed with an adjuvant. Each mouse receives four injections of 10 to 100 μ g of immunogen and after the fourth injection blood samples are taken from the mice to determine if the serum contains antibody to the immunogen. Serum titer is determined by ELISA or RIA. Mice with sera indicating the presence of antibody to the immunogen are selected for hybridoma production.

-92-

Spleens are removed from immune mice and a single cell suspension is prepared (see Harlow and Lane, 1988). Cell fusions are performed essentially as described by Kohler and Milstein, 1975. Briefly, P3.65.3 myeloma cells (American Type Culture Collection, Rockville, MD) are fused with immune spleen cells using polyethylene glycol as described by Harlow and Lane, 1988. Cells are plated at a density of $2x10^5$ cells/well in 96 well tissue culture plates. Individual wells are examined for growth and the supernatants of wells with growth are tested for the presence of BRCA1 specific antibodies by ELISA or RIA using wild type or mutant BRCA1 target protein. Cells in positive wells are expanded and subcloned to establish and confirm monoclonality.

Clones with the desired specificities are expanded and grown as ascites in mice or in a hollow fiber system to produce sufficient quantities of antibody for characterization and assay development.

EXAMPLE 14

Sandwich Assay for BRCA1

Monoclonal antibody is attached to a solid surface such as a plate, tube, bead, or particle. Preferably, the antibody is attached to the well surface of a 96-well ELISA plate. 100 μ l sample (e.g., serum, urine, tissue cytosol) containing the BRCA1 peptide/protein (wild-type or mutant) is added to the solid phase antibody. The sample is incubated for 2 hrs at room temperature. Next the sample fluid is decanted, and the solid phase is washed with buffer to remove unbound material. 100 μ l of a second monoclonal antibody (to a different determinant on the BRCA1 peptide/protein) is added to the solid phase. This antibody is labeled with a detector molecule (e.g., 125-I, enzyme, fluorophore, or a chromophore) and the solid phase with the second antibody is incubated for two hrs at room temperature. The second antibody is decanted and the solid phase is washed with buffer to remove unbound material.

The amount of bound label, which is proportional to the amount of BRCA1 peptide/protein present in the sample, is quantitated. Separate assays are performed using monoclonal antibodies which

10

15

20

25

30

are specific for the wild-type BRCA1 as well as monoclonal antibodies specific for each of the mutations identified in BRCA1.

-93-

EXAMPLE 15

Analysis of BRCA1 Mutations

The DNA samples which were screened for BRCA1 mutations were extracted from blood or tumor samples from patients with breast or ovarian cancer (or known carriers by haplotype analysis) who were participating in research studies on the genetics of breast cancer. All subjects signed appropriate informed consent. Table 13 details the number of samples, ascertainment criteria, and screening method for each set of samples screened.

10

	Sets of DNA Samples Screened for Mutations in BRCA1									
Source <u>of Samples</u>	Description of Samples ¹	Screening Method ²	No. Samples Screened	No. Mutations Found to Date						
UTAH-2	Br/Ov Families	SEQ	10	2						
MONTREAL	Br/Ov Families	SEQ	30	13						
MSKCC-1	Br and Br/Ov Families	SEQ	14	2						
MSK/UT-1	Early Onset Br Cases	SEQ	24	• • • • • • • • • • • • • • • • • • •						
STRANG	Br and Br/Ov Families	SEQ	12	4						
STOCKHOLM	Br and Br/Ov Families	SEQ	15	4						
USC-1	Bilat Br Proband, High-Ri	.sk SEQ	7	3						
TUMOR-3	Early Onset Br Tumors	SEQ	14	1						
USC-2	Bilat Br <50+1° rel Br	ASO	59	5						
MSK/UT-2	Early Onset Br Cases	ASO	109	3						
YN	Bilateral; Early Onset	SSCA	103	4						
Texas	Br/Ov Families	SEQ	15	2						
Utah	Br/Ov Families	SEQ	10	1						
Pisa	Br/Ov Families	SEQ	21	4						

TABLE 13

-94-

. .

TABLE 13 (cont'd)

Tumor1mod		SEQ		1	
MSKCC-2	Early Onset Br Cases	SEQ	21	3 .	

- 1 Most sample groups contained a heterogeneous mixture of samples. The most representative description of each set is given.
- 2 SEQ Direct sequencing of PCR products; SSCA Single Strand Conformation Assay; ASO Allele-Specific Oligo

-95- -

Although the original mutations described in Miki *et al.*, 1994 were detected through screening of cDNA, 25 pairs of intronic PCR primers were used to amplify the complete coding sequence and splice junctions from genomic DNA for the majority of the remaining samples. Updated primer information is publicly available via anonymous ftp from morgan.med.utah.edu in the directory pub/BRCA1. Where possible, DNA sequence variations were tested for cosegregation with breast or ovarian cancer in the family. Further evidence of a causal role of a sequence variant in cancer was provided by proving the absence of the putative mutation in a set of control individuals. Screening for specific, previously-identified mutations in large sets of selected samples was performed using ASO hybridization.

-96-

Table 14 describes several of the 113 mutations found screening the entire BRCA1 coding sequence, the 21 intron/exon boundaries and by finding polymorphic sites in genomic DNA reduce to monomorphic sites in cDNA. Two common mutations were found and their frequencies in other samples were examined by ASO analysis (Table 15). Tables 16 and 17 describe the distribution of mutations by type and by location within the BRCA1 coding sequence, respectively. By far, the majority of mutations identified were frameshifts which accounted for 65% of the 64 different mutations found to date in BRCA1 and in 72% of all samples shown to have BRCA1 mutations through complete screening. Globally, no statistically significant departure from a random distribution across the coding sequence of BRCA1 was found ($\chi^2 = 2.00, 2$ df, p = 0.37) among the 64 distinct mutations found in the coding sequence of BRCA1 to date.

5

15

TABLE 14

Sample Set	Family		ases				escription
and a start and a start of the		BR	QV	<u>Type</u> ¹	Exon	Codon	Mutation ²
TEXAS	132-000			FS	2	23	185 ins A -> ter 40
MONTREAL	180	2	2	FS	2	23	185 del AG -> ter 39
MONTREAL	235	4	2	FS	2	23	185 del AG -> ter 39
MONTREAL	253	1	3	FS	2	23	185 del AG -> ter 39
MONTREAL	255	0	7	FS	2	23	185 del AG -> ter 39
MSKCC	210311	3	0	FS	2	23	185 del AG -> ter 39
USC-1	008	2	1	FS	2	23	185 del AG -> ter 39
UTAH				SP	I-5	I-5	T->G ins 59 -> ter 75
TUMOR-34	_	1	0.0	FS	11	270	926 ins 11 -> ter 301
MSK/UT-1	-	1	0	FS	11	270	926 ins 11 -> ter 301
YN98	· · · ·	1	0	MS	11	271	Val 271 Met
MONTREAL	270	4	3	FS	11	339	1128 ins A -> ter 345
STRANG	2903	. 1	2	FS	11	339	1128 ins A -> ter 345
MONTREAL	185	1	3	FS	11	392	1294 del 40 -> ter 396
PISA	6			FS	11	461	1499 ins A -> ter 479
PISA	17			FS	11	461	1499 ins A -> ter 479
PISA	31			FS	11	461	1499 ins A -> ter 479
USC-1	052	5	1	FS	11	655	2080 ins A -> ter 672
USC-1	068	2	1	FS	11	655	2080 ins A -> ter 672
STRANG	2802	2	2	FS	11	725	2293 del G -> ter 735

Mutations Identified by Complete Screening of the BRCA1 Gene

-97-

TABLE 14 (Cont'd)

Mutations Identified by Complete Screening of the BRCA1 Gene

Sample Set Family	# Cases	Mutatic	n Description
	BR OV	Type ¹ Exon Cod	on Mutation ²
YN100 -	1 0	FS 11 79	7 2509 del AA -> ter 799
TUMOR1mod OV24	0 1	FS 11 81	9 2575 del C -> ter 845
MONTREAL 179	2 3	MS 11 82	6 Thr 826 Lys
STOCKHOLM AL48	3 1	FS 11 82	6 2596 del C -> ter 845
STOCKHOLM BR33	51	FS 11 82	6 2596 del C -> ter 845
UTAH-2 2305	2 7	FS 11 95	6 2982 del 5 -> ter 968
MONTREAL 218	5 1	FS 11 100	2 3121 del A -> ter 1023
MSK17572		MS 11 100	
STOCKHOLM BR24	2 1	FS 11 101	6 3166 ins 5 -> ter 1025
TEXAS BC110-001		FS 11 111	1 3450 del 4 -> ter 1115
YN23 -	1 0	MS 11 115	
STOCKHOLM PAL33	1 0	FS 11 120	9 3745 del T -> ter 1209
YN44 -	1 0	NS 11 121	
MSK12871		MS 11 121	9 Glu 1219 Asp
UTAH-2 2039	3 2	MS 11 134	7 Arg 1347 Gly
MONTREAL 183	4 1	FS 11 135	
STRANG 1900 ³	3 1	NS 13 144	
PISA #8		FS 16 158	
MSK9646		MS 16 162	
STRANG 8622 ³	4 1	FS 16 165	6 5085 del 19 -> ter 1670

- 98-

TABLE 14 (Cont'd)

Mutations Identified by Complete Screening of the BRCA1 Gene

Sample Set	Family	# Ca	ses		Mut	ation De	scription
		BR	QV	Type ¹	Exon	Codon	Mutation ²
MONTREAL	101	2	2	FS	20	1756	5382 ins C -> ter 1829
MONTREAL	162	3	1	FS	20	1756	5382 ins C -> ter 1829
MONTREAL	166	5	2	FS	20	1756	5382 ins C -> ter 1829
MONTREAL	279	4	0	FS	20	1756	5382 ins C -> ter 1829
MSKCC	193549	0	3	FS	20	1756	5382 ins C -> ter 1829

¹ FS-Frameshift; NS-Nonsense; MS-Missense; SP-Splice Site.

2

For Missense and Nonsense mutations, the mutation description contains: wild type amino acid, affected codon, altered amino acid (or ter). For frameshift mutations, the format is: nucleotide, insertion or deletion, specific nucleotides changed (if <3) or number inserted or deleted (if >2) and the amino acid (accounting for the insertion or deletion) in which the frameshift results in a termination signal. Nucleotides refer to the BRCA1 cDNA sequence in GENBANK under Accession No. U-14680.

The mutation in this family was independently identified in both the Myriad and University of Pennsylvania Labs.

⁴ The mutation identified in this tumor was also found in the germline of the individual.

-99-

-100-. ---

TABLE 15

Frequency of Two Common BRCA1 Mutations

		Number	Number of Mutations Found			
5	Set	Studied	<u>185 del AG</u>	<u>5382 ins C</u>		
	USC-1	59	4	1		
•	MSK/UT-2	109	3	0		
	GLASGOW-2	100	Not tested	3		
10	GLASGOW-3	100	Not tested	2		
	CRC-OV	250	Not tested	1		

TABLE 16

Observed Frequency of Different Types of Mutations

	Number (Percent)		
Mutation Type	Distinct Mutations ¹	All Mutations ²	
Frameshift	42 (65)	81 (72)	
Nonsense	10 (16)	13 (12)	
Missense	9 (14)	14 (12)	
Other	3 (5)	5 (4)	

25

30

20

15

¹ Identical mutations are counted only once in this column.
² Each sample in which a mutation has been identified is counted in this column.

TABLE 17

Distribution of Identified Mutations in BRCA1 Coding Sequence

		Amino Acids		
35 '	Mutations	1-621	<u>622-1242</u>	<u>1243-1863</u>
	Distinct	18	23	21
	All	44	28	39

-101-

Mutations have been found in many different regions of the gene - phenotypically severe mutations have been found both in the extreme 5' end of BRCA1 as well as in the extreme 3' portion of the gene. One such mutation found in a family with seven early-onset breast cancer cases produces a protein that is only missing the terminal 10 amino acids, indicating that this region of BRCA1 plays a role in normal gene function. It is noteworthy the overwhelming majority of alterations in BRCA1 have been either frameshift or nonsense mutations. Of the total of 68 independent individuals with mutations identified by complete screening of the coding sequence, 54 (79%) are frameshift or nonsense mutations, resulting in an unstable or truncated protein product.

In BRCA1, to date, two mutations appear to be relatively common. The 5382 ins C BRCA1 mutation in codon 1756 and the 185 del AG mutation in codon 23 were identified by direct sequencing in seven (10%) and eight (12%) of the 68 probands in which mutations have been identified, respectively. In addition to these common mutations, two mutations were identified three times and there are eleven additional mutations which have been found twice by a complete screen of the cDNA. The total fraction of mutations which are recurrent is 38% and these 15 recurrent mutations comprise 63% of the 68 mutations detected in probands. Many of the probands screened to date for BRCA1 mutations were selected for having a high prior probability of having such mutations. Thus the mutations found in this set may not be representative of those which would be identified in other sets of patients. However, the two most frequent BRCA1 mutations (5382 ins C and 185 del AG) have been found multiple times in targeted screening in sets of probands who were either unselected for family history.

Industrial Utility

25

30

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

10

15

20

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.

15

10

5

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.

-102-

-103-

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. Natl. Acad. Sci. USA 77:5399-5403.

Anderson, D.E. (1972). J. Natl. Cancer Inst. 48:1029-1034.

Anderson, J.A., et al. (1992). J. Otolaryngology 21:321.

Antonarakis, S.E., et al. (1985). New Eng. J. Med. 313:842-848.

Ausubel, 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 (1992). Curr. Top. Microbiol. Immunol. 158:39-61.

Berkner, et al. (1988). BioTechniques 6:616-629.

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: <u>Banbury Report 4: Cancer Incidence in Defined Populations</u> (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.

Bowcock, A.M., et al. (1993). Am. J. Hum. Genet. 52:718.

Brandyopadhyay 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. Natl. Acad. Sci. USA 88:4005-4009.
Cannon-Albright, L., et al. (1994). Cancer Research 54:2378-2385.

-104-

Capecchi, 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. Natl. Acad. Sci. USA 80:278-282.

Constantini and Lacy (1981). Nature 294:92-94.

Cotten, et al. (1990). Proc. Natl. Acad. Sci. USA 87:4033-4037.

Cotton, et al. (1988). Proc. Natl. 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. (1991a). Proc. Natl. Acad. Sci. USA 88:8850-8854.

Curiel, et al. (1991b). 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.

Enhancers and Eurkaryotic Gene Expression, Cold Spring Harbor Press, Cold Spring Harbor, New York (1983).

Erickson, J. et al., (1990). Science 249:527-533.

Fain, P.R. (1992). Cytogen. Cell Genet. 60:178.

Felgner, et al. (1987). Proc. Natl. 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. (1992a). Hum. Molec. Genet. 1:66.

Futreal, P.A., et al. (1992b). 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. Natl. 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.

Grompe, M., (1993). Nature Genetics 5:111-117.

Grompe, M., et al., (1989). Proc. Natl. Acad. Sci. USA 86:5855-5892.

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. (1990). Science 250:1684-1689.

Hall, J.M., et al. (1992). Am. J. Hum. Genet. 50:1235-1241.

Harlow & Lane (1988). <u>Antibodies: A Laboratory Manual</u> (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.).

Jablonski, E., et al. (1986). Nuc. Acids Res. 14:6115-6128.

Jacobs, I.J., et al. (1993). Cancer Res. 53:1218-1221.

Jakoby, W.B. and Pastan, I.H. (eds.) (1979). Cell Culture. Methods in Enzymology, volume 58 (Academic Press, Inc., Harcourt Brace Jovanovich (New York)).

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.

Kohler, G. and Milstein, C. (1975). Nature 256:495-497.

Kozak, M. (1987). Nucleic Acids Res. 15:8125-8148.

Kraemer, F.B. et al. (1993). J. Lipid Res. 34:663-672.

-107-

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.

Little, M.H., et al. (1992). Proc. Natl. Acad. Sci. USA 89:4791.

Little, M.H., et al. (1993). Hum. Mol. Genet. 2:259.

Lovett, et al. (1991). Proc. Natl. 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. (1982). 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.

Martin, R., et al. (1990). BioTechniques 9:762-768.

Matteucci, M.D. and Caruthers, M.H. (1981). J. Am. Chem. Soc. 103:3185.

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.

Miki, Y., et al., 1994. Science 266:66-71.



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 (1992). Hum. Gene Ther. 3:399-410.

Nabel, et al. (1990). Science 249:1285-1288.

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. Natl. Acad. Sci. USA 85:3044-3048.

Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C., and Markham, A.F. (1989). Nucl. Acids Res. 17:2503-2516.

Nguyen, Q., et al. (1992). BioTechniques 13:116-123.

Novack, et al. (1986). Proc. Natl. 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. Natl. Acad. Sci. USA 86:2776-2770.

-109-

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. Natl. 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.

Rigby, P.W.J., et al. (1977). J. Mol. Biol. 113:237-251.

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. Natl. Acad. Sci. USA 86:232-236.

Sheffield, V.C., et al. (1991). Am. J. Hum. Genet. 49:699-706.

Shenk, et al. (1975). Proc. Natl. Acad. Sci. USA 72:989.

Shimada, et al. (1991). J. Clin. Invest. 88:1043-1047.

Shinkai, Y., et al. (1992). Cell 68:855.

Shizuya, H., et al. (1992). Proc. Natl. Acad. Sci. USA 89:8794-8797.

Simard, J., et al. (1993). Human Mol. Genet. 2:1193-1199.

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. (1980). <u>Banbury Report No 4: Cancer Incidence in Defined Populations</u> (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., et al. (1990). Science 250:1715-1720.

Smith, S.A., et al. (1992). Nature Genetics 2:128-131.

Smith, T.F. and Waterman, M.S. (1981). J. Mol. Biol. 147:195-197.

Snouwaert, J.N., et al. (1992). Science 257:1083.

Sorge, et al. (1984). Mol. Cell. Biol. 4:1730-1737.

Srivastava, S., et al. (1993). Cancer Res. 53:4452-5.

Sternberg (1990). Proc. Natl. 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.

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.

Thomas, A. and Skolnick, M.H. (1994). 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. (1990). Proc. Natl. Acad. Sci. USA 87:3410-3414.

Wagner, et al. (1991). Proc. Natl. Acad. Sci. USA 88:4255-4259.

Wang and Huang (1989). Biochemistry 28:9508-9514.

-111-

Wartell, R.M., et al. (1990). Nucl. Acids Res. 18:2699-2705. Weber, J.L. (1990). Genomics 7:524-530. Weber and May (1989). Am. J. Hum. Genet. 44:388-396. 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, M.B., et al., (1992). Genomics 12:301-306. White and 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 265:2088. Wu, et al. (1989a). Genomics 4:560-569. Wu, et al. (1989b). J. Biol. Chem. 264:16985-16987. Wu, et al. (1991). J. Biol. Chem. 266:14338-14342. Zenke, et al. (1990). Proc. Natl. Acad. Sci. USA 87:3655-3659. List of Patents and Patent Applications: U.S. Patent No. 3,817,837 U.S. Patent No. 3,850,572 U.S. Patent No. 3,939,350 U.S. Patent No. 3,996,345 U.S. Patent No. 4,275,149

U.S. Patent No. 4,277,437

U.S. Patent No. 4,366,241

U.S. Patent No. 4,376,110

U.S. Patent No. 4,486,530

U.S. Patent No. 4,683,195

U.S. Patent No. 4,683,202

U.S. Patent No. 4,816,567

U.S. Patent No. 4,868,105

U.S. Patent No. 5,252,479

EPO Publication No. 225,807

European Patent Application Publication No. 0332435

Geysen, H., PCT published application WO 84/03564, published 13 September 1984

-112-

Hitzeman et al., EP 73,675A

PCT published application WO 93/07282



-113-

SEQUENCE LISTING

. --

(1) GENERAL INFORMATION:

 (i) APPLICANT: Shattuck-Eidens, Donna M. Simard, Jacques
 Emi, Mitsuru
 Nakamura, Yusuke

- (ii) TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer Susceptibility Gene Mutations
- (iii) NUMBER OF SEQUENCES: 65
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP(B) STREET: 1201 New York Avenue, N.W., Suite 1000
 - (C) CITY: Washington
 - (D) STATE: DC
 - (D) STATE: DC
 - (E) COUNTRY: USA
 - (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 #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US
 - (B) FILING DATE:
 - (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/348,824(B) FILING DATE: 29-NOV-1994

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/308,104(B) FILING DATE: 16-SEP-1994

- (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

. *•*



(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Ihnen, Jeffrey L.
 - (B) REGISTRATION NUMBER: 28,957
 - (C) REFERENCE/DOCKET NUMBER: 24884-109347

-114-. ..

- (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: 5913 base pairs
 - (B) TYPE: nucleic acid(C) STRANDEDNESS: double

 - (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: 120..5711

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

AGC	rcgc	rga (GACT	PCCTO	GG A	cccc	SCACO	C AGO	3CTG:	rggg	GTT	rctc)	AGA 1	raac'	IGGGCC	60
CCTC	GCGC	FCA C	GAG	SCCT	IC A	CCCT	CTGC	r cto	GGT	AAAG	TTC	ATTG	GAA (CAGA	AAGAA	119
	GAT							-		, e						167
Met 1	Asp	Leu	Ser	Ala 5	Leu	Arg	Val	Glu	Glu 10	Val	Gln	Asn	Val	Ile 15	Asn	•
GCT	ATG	CAG	ААА	ATC	TTA	GAG	TGT	CCC	ATC	TGT	CTG	GAG	TTG	ATC	AAG	215
Ala	Met	Gln	Lys 20	Ile	Leu	Glu	Суз	Pro 25	Ile	Cys	Leu	Glu	Leu 30	Ile	Lys	
GAA	CCT	GTC	TCC	ACA	AAG	TGT	GAC	CAC	ATA	TTT	TGC	AAA	TTT	TGC	ATG	263
Glu	Pro	Val 35		Thr	Lys	Суз	Asp 40	His	Ile	Phe	Cys	Lys 45	Phe	Cys	Met	
CTG	ААА	CTT	CTC	AAC	CAG	AAG	ААА	GGG	CCT	TCA	CAG	TGT	CCT	TTA	TGT	311
Leu	Lys 50	Leu	Leu	Asn	Gln	Lys 55	Lys	Gly	Pro	Ser	Gln 60	Cys	Pro	Leu	Сув	

-115-AAG AAT GAT ATA ACC AAA AGG AGC CTA CAA GAA AGT ACG AGA TTT AGT Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser CAA CTT GTT GAA GAG CTA TTG AAA ATC ATT TGT GCT TTT CAG CTT GAC Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp ACA GGT TTG GAG TAT GCA AAC AGC TAT AAT TTT GCA AAA AAG GAA AAT Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn AAC TCT CCT GAA CAT CTA AAA GAT GAA GTT TCT ATC ATC CAA AGT ATG Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met GGC TAC AGA AAC CGT GCC AAA AGA CTT CTA CAG AGT GAA CCC GAA AAT Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn CCT TCC TTG CAG GAA ACC AGT CTC AGT GTC CAA CTC TCT AAC CTT GGA Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly ACT GTG AGA ACT CTG AGG ACA AAG CAG CGG ATA CAA CCT CAA AAG ACG Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr TCT GTC TAC ATT GAA TTG GGA TCT GAT TCT TCT GAA GAT ACC GTT AAT Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn AAG GCA ACT TAT TGC AGT GTG GGA GAT CAA GAA TTG TTA CAA ATC ACC Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr CCT CAA GGA ACC AGG GAT GAA ATC AGT TTG GAT TCT GCA AAA AAG GCT Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala GCT TGT GAA TTT TCT GAG ACG GAT GTA ACA AAT ACT GAA CAT CAT CAA Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln CCC AGT AAT AAT GAT TTG AAC ACC ACT GAG AAG CGT GCA GCT GAG AGG Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg CAT CCA GAA AAG TAT CAG GGT AGT TCT GTT TCA AAC TTG CAT GTG GAG His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu CCA TGT GGC ACA AAT ACT CAT GCC AGC TCA TTA CAG CAT GAG AAC AGC Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser .

AGT TTA TTA CTC ACT AAA GAC AGA ATG AAT GTA GAA AAG GCT GAA TTC Ser Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe TGT AAT AAA AGC AAA CAG CCT GGC TTA GCA AGG AGC CAA CAT AAC AGA Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg TGG GCT GGA AGT AAG GAA ACA TGT AAT GAT AGG CGG ACT CCC AGC ACA Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr GAA AAA AAG GTA GAT CTG AAT GCT GAT CCC CTG TGT GAG AGA AAA GAA Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu TGG AAT AAG CAG AAA CTG CCA TGC TCA GAG AAT CCT AGA GAT ACT GAA Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu GAT GTT CCT TGG ATA ACA CTA AAT AGC AGC ATT CAG AAA GTT AAT GAG Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu TGG TTT TCC AGA AGT GAT GAA CTG TTA GGT TCT GAT GAC TCA CAT GAT Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp GGG GAG TCT GAA TCA AAT GCC AAA GTA GCT GAT GTA TTG GAC GTT CTA Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu AAT GAG GTA GAT GAA TAT TCT GGT TCT TCA GAG AAA ATA GAC TTA CTG Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu GCC AGT GAT CCT CAT GAG GCT TTA ATA TGT AAA AGT GAA AGA GTT CAC Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His TCC AAA TCA GTA GAG AGT AAT ATT GAA GAC AAA ATA TTT GGG AAA ACC Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr TAT CGG AAG AAG GCA AGC CTC CCC AAC TTA AGC CAT GTA ACT GAA AAT Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn CTA ATT ATA GGA GCA TTT GTT ACT GAG CCA CAG ATA ATA CAA GAG CGT Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg CCC CTC ACA AAT AAA TTA AAG CGT AAA AGG AGA CCT ACA TCA GGC CTT Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu

-116-

CAT CCT GAG GAT TTT ATC AAG AAA GCA GAT TTG GCA GTT CAA AAG ACT His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr CCT GAA ATG ATA AAT CAG GGA ACT AAC CAA ACG GAG CAG AAT GGT CAA Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln GTG ATG AAT ATT ACT AAT AGT GGT CAT GAG AAT AAA ACA AAA GGT GAT Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp TCT ATT CAG AAT GAG AAA AAT CCT AAC CCA ATA GAA TCA CTC GAA AAA Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys GAA TCT GCT TTC AAA ACG AAA GCT GAA CCT ATA AGC AGC AGT ATA AGC Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser AAT ATG GAA CTC GAA TTA AAT ATC CAC AAT TCA AAA GCA CCT AAA AAG Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys AAT AGG CTG AGG AGG AAG TCT TCT ACC AGG CAT ATT CAT GCG CTT GAA Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu . 620 CTA GTA GTC AGT AGA AAT CTA AGC CCA CCT AAT TGT ACT GAA TTG CAA Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln ATT GAT AGT TGT TCT AGC AGT GAA GAG ATA AAG AAA AAG TAC AAC Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Tyr Asn CAA ATG CCA GTC AGG CAC AGC AGA AAC CTA CAA CTC ATG GAA GGT AAA Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys GAA CCT GCA ACT GGA GCC AAG AAG AGT AAC AAG CCA AAT GAA CAG ACA Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr AGT AAA AGA CAT GAC AGC GAT ACT TTC CCA GAG CTG AAG TTA ACA AAT Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn GCA CCT GGT TCT TTT ACT AAG TGT TCA AAT ACC AGT GAA CTT AAA GAA Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu TTT GTC AAT CCT AGC CTT CCA AGA GAA GAA AAA GAA GAG AAA CTA GAA Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu

GeneDX 1010, pg. 122

-117-

			.'						-1	18-							
			GTG													2375	
Thr	Val	Lys	Val 740	Ser	Asn	Asn	Ala		Asp	Pro	Lys	Asp	Leu 750	Met	Leu		
			740					745					/50				
AGT	GGA	GAA	AGG	GTT	TTG	CAA	ACT	GAA	AGA	TCT	GTA	GAG	AGT	AGC	AGT	2423	
Ser	Gly	Glu	Arg	Val	Leu	Gln	Thr	Glu	Arg	Ser	Val	Glu	Ser	Ser	Ser		
		755					760					765					
		-	~~~	~~~	~~~		~ ~ ~								-	0471	
			GTA Val													2471	
	770	Leu	VAL	FIO	GTÀ	775	veh	TÀT	GTA	****	780	Gru	Jer		UG1		
TTA	CTG	GAA	GTT	AGC	ACT	CTA	GGG	AAG	GCA	AAA	ACA	GAA	CCA	AAT	AAA	2519	
	Leu	Glu	Val	Ser		Leu	Gly	Lys	Ala		Thr	Glu	Pro	Asn			
785					790					795					800		
TOT	GTG	AGT	CAG	TGT	GCA	GCA	ጥጥጥ	GAA	AAC	CCC	AAG	GGA	CTA	ATT	CAT	2567	
			Gln														
				805					810					815			
			AAA													2615	
GΙΥ	Cys	Ser	Lys 820	Asp	Asn	Arg	Asn	Asp 825	Thr	Glu	GIΥ	Pne	830	Tyr	Pro		
			020					040					030				
TTG	GGA	CAT	GAA	GTT	AAC	CAC	AGT	CGG	GAA	ACA	AGC	ATA	GAA	ATG	GAA	2663	
Leu	Gly	His	Glu	Val	Asn	His	Ser	Arg	Glu	Thr	Ser	Ile	Glu	Met	Glu		
		835					840					845					
33	200	~~~	CTT	~~~	000	~~~	-	mmc	~~~		202	mma	NAG	CUTUT	ጥሮ እ	2711	
			Leu													dia 7 ata sta	
U.L.U	850	044				855	-1-				860			, .,			
			TCA													2759	
-	Arg	Gln	Ser	Phe		Pro	Phe	Ser	Asn		Gly	Asn	Ala	Giu	880		
865					870					875					000		
GAA	TGT	GCA	ACA	TTC	TCT	GCC	CAC	TCT	GGG	TCC	TTA	AAG	AAA	CAA	AGT	2807	
			Thr														
				885					890					895			
					~ ~ ~	-	~ • •	~ ~ ~		~			~~~~	007	NAG	2855	'
			ACT												Lys	2033	
ETO.	цуа	VAT	900	FIIG	Gru	cya	Gru	905	цуа	94.4	414		910	,,	-4-		
			AAT													2903	
Asn	Glu		Asn	Ile	Lys	Pro		Gln	Thr	Val	Asn		Thr	Ala	Gly		
		915					920					925					
TTT	CCT	GTG	GTT	GGT	CAG	ААА	GAT	AAG	CCA	GTT	GAT	AAT	GCC	AAA	TGT	2951	
			Val														
	930			-		935		-			940						
•		.	.										mme	***	000	2999	
															GGC Gly	6333	
945	T 7 6	างร	ary	GTÅ	950	ALG.	E116	Cya	Led	955		9711	- 116	J	960		
				•													

									-	119-							
AAC	GAA	ACT	GGA	CTC	ATT	ACT	CCA	AAT	AAA	CAT	GGA	CTT	TTA	CAA	AAC	3(047
Asn	Glu	Thr	Gly	Leu	Ile	Thr	Pro	Asn	Lys	His	Gly	Leu	Leu	Gln	Asn		
				965					970		- -			975			
CCA	TAT	CGT	ATA	CCA	CCA	CTT	TTT	CCC	ATC	AAG	TCA	TTT	GTT	AAA	ACT	3(95
Pro	Tyr	Arg	Ile	Pro	Pro	Leu	Phe	Pro	Ile	Lys	Ser	Phe	Val	Lys	Thr		
			980					985		-			990				
AAA	TGT	AAG	AAA	AAT	CTG	CTA	GAG	GAA	AAC	TTT	GAG	GAA	CAT	TCA	ATG	3.1	L43
Lys	Cys	Lys	Lys	Asn	Leu	Leu	Glu	Glu	Asn	Phe	Glu	Glu	His	Ser	Met		
		995					1000	2				1005	5				
		•															
TCA	CCT	GAA	AGA	GAA	ATG	GGA	AAT	GAG	AAC	ATT	CCA	AGT	ACA	GTG	AGC	3	L91
Ser	Pro	Glu	Arg	Glu	Met	Gly	Asn	Glu	Asn	Ile	Pro	Ser	Thr	Val	Ser		
	1010)				1015	5				1020)					
			1														
ACA	ATT	AGC	CGT	AAT	AAC	ATT	AGA	GAA	AAT	GTT	TTT	AAA	GAA	GCC	AGC	32	239
Thr	Ile	Ser	Arg	Asn	Asn	Ile	Arg	Glu	Asn	Val	Phe	Lys	Glu	Ala	Ser		
1025	5				1030)				1035	5				1040		
TCA	AGC	AAT	ATT	AAT	GAA	GTA	GGT	TCC	AGT	ACT	AAT	GAA	GTG	GGC	TCC	32	287
Ser	Ser	Asn	Ile	Asn	Glu	Val	Gly	Ser	Ser	Thr	Asn	Glu	Val	Gly	Ser		
				1045	5				1050)				1055	5		
AGT	ATT	AAT	GAA	ATA	GGT	TCC	AGT	GAT	GAA	AAC	ATT	CAA	GCA	GAA	CTA	33	335
Ser	Ile	Asn	Glu	Ile	Gly	Ser	Ser	Asp	Glu	Asn	Ile	Gln	Ala	Glu	Leu		
	•		1060)				106	5				1070).			
GGT	AGA	AAC	AGA	GGG	CCA	AAA	TTG	AAT	GCT	ATG	CTT	AGA	TTA	GGG	GTT	33	383
Gly	Arg	Asn	Arg	Gly	Pro	Lys	Leu	Asn	Ala	Met	Leu	Arg	Leu	Gly	Val		
		1075	5				1080)				1085	5				
														-			
TTG	CAA	CCT	GAG	GTC	TAT	AAA	CAA	AGT	CTT	CCT	GGA	AGT	AAT	TGT	AAG	34	131
Leu	Gln	Pro	Glu	Val	Tyr	Lys	Gln	Ser	Leu	Pro	Gly	Ser	Asn	Cys	Lys		
	1090)				1095	5				1100) ¹ 1					
CAT	CCT	GAA	ATA	AAA	AAG	CAA	GAA	TAT	GAA	GAA	GTA	GTT	CAG	ACT	GTT	34	179
His	Pro	Glu	Ile	Lys	Lys	Gln	Glu	Tyr	Glu	Glu	Val	Val	Gln	Thr	Val		
1105	5				1110)				1115	5				1120		
TAA	ACA	GAT	TTC	TCT	CCA	TAT	CTG	ATT	TCA	GAT	AAC	TTA	GAA	CAG	CCT	35	527
Asn	Thr	Asp	Phe	Ser	Pro	Tyr	Leu	Ile	Ser	Asp	Asn	Leu	Glu	Gln	Pro		
				1125	5				1130)				113	5		
ATG	GGA	AGT	AGT	CAT	GCA	TCT	CAG	GTT	TGT	TCT	GAG	ACA	CCT	GAT	GAC	35	575
Met	Gly	Ser	Ser	His	Ala	Ser	Gln	Val	Cys	Ser	Glu	Thr	Pro	Asp	Asp		
			1140).				114	5				1150)			
CTG	TTA	GAT	GAT	GGT	GAA	ATA	AAG	GAA	GAT	ACT	AGT	TTT	GCT	GAA	AAT	30	523
Leu	Leu	Asp	Asp	Gly	Glu	Ile	Lys	Glu	Asp	Thr	Ser	Phe	Ala	Glu	Asn		
		1155	5	-			1160)				1165	5 .				
GAC	ATT	AAG	GAA	AGT	TCT	GCT	GTT	TTT	AGC	AAA	AGC	GTC	CAG	AAA	GGA	30	571
Asp	Ile	Lys	Glu	Ser	Ser	Ala	Val	Phe	Ser	Lys	Ser	Val	Gln	Lys	Gly		
	1170					1175					118						

-120-. ... GAG CTT AGC AGG AGT CCT AGC CCT TTC ACC CAT ACA CAT TTG GCT CAG Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln GGT TAC CGA AGA GGG GCC AAG AAA TTA GAG TCC TCA GAA GAG AAC TTA Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu TCT AGT GAG GAT GAA GAG CTT CCC TGC TTC CAA CAC TTG TTA TTT GGT Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly AAA GTA AAC AAT ATA CCT TCT CAG TCT ACT AGG CAT AGC ACC GTT GCT Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala ACC GAG TGT CTG TCT AAG AAC ACA GAG GAG AAT TTA TTA TCA TTG AAG Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys AAT AGC TTA AAT GAC TGC AGT AAC CAG GTA ATA TTG GCA AAG GCA TCT Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser CAG GAA CAT CAC CTT AGT GAG GAA ACA AAA TGT TCT GCT AGC TTG TTT Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe TCT TCA CAG TGC AGT GAA TTG GAA GAC TTG ACT GCA AAT ACA AAC ACC Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr CAG GAT CCT TTC TTG ATT GGT TCT TCC AAA CAA ATG AGG CAT CAG TCT Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser GAA AGC CAG GGA GTT GGT CTG AGT GAC AAG GAA TTG GTT TCA GAT GAT Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp GAA GAA AGA GGA ACG GGC TTG GAA GAA AAT AAT CAA GAA GAG CAA AGC Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser ATG GAT TCA AAC TTA GGT GAA GCA GCA TCT GGG TGT GAG AGT GAA ACA Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr AGC GTC TCT GAA GAC TGC TCA GGG CTA TCC TCT CAG AGT GAC ATT TTA Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu ACC ACT CAG CAG AGG GAT ACC ATG CAA CAT AAC CTG ATA AAG CTC CAG Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln

-121-CAG GAA ATG GCT GAA CTA GAA GCT GTG TTA GAA CAG CAT GGG AGC CAG Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln CCT TCT AAC AGC TAC CCT TCC ATC ATA AGT GAC TCT TCT GCC CTT GAG Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu GAC CTG CGA AAT CCA GAA CAA AGC ACA TCA GAA AAA GCA GTA TTA ACT Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr TCA CAG AAA AGT AGT GAA TAC CCT ATA AGC CAG AAT CCA GAA GGC CTT Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu TCT GCT GAC AAG TTT GAG GTG TCT GCA GAT AGT TCT ACC AGT AAA AAT Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn AAA GAA CCA GGA GTG GAA AGG TCA TCC CCT TCT AAA TGC CCA TCA TTA Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu GAT GAT AGG TGG TAC ATG CAC AGT TGC TCT GGG AGT CTT CAG AAT AGA Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg AAC TAC CCA TCT CAA GAG GAG CTC ATT AAG GTT GTT GAT GTG GAG GAG Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu CAA CAG CTG GAA GAG TCT GGG CCA CAC GAT TTG ACG GAA ACA TCT TAC Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr TTG CCA AGG CAA GAT CTA GAG GGA ACC CCT TAC CTG GAA TCT GGA ATC Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile AGC CTC TTC TCT GAT GAC CCT GAA TCT GAT CCT TCT GAA GAC AGA GCC Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala CCA GAG TCA GCT CGT GTT GGC AAC ATA CCA TCT TCA ACC TCT GCA TTG Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu AAA GTT CCC CAA TTG AAA GTT GCA GAA TCT GCC CAG AGT CCA GCT GCT Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Ser Pro Ala Ala GCT CAT ACT ACT GAT ACT GCT GGG TAT AAT GCA ATG GAA GAA AGT GTG Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val

		-							-1	22-						
									-							
AGC																5063
361	HIG	1635	-	PIO	Glu	reu	1640		ser	Thr	GIU	Arg 1649		Asn	гуа	
								•					•			
AGA																5111
Arg			Met	Val	Val		-	Leu	Thr	Pro			Phe	Met	Leu	
	1650)				1655	5				1660)				
GTG	TAC	AAG	TTT	GCC	AGA	AAA	CAC	CAC	ATC	ACT	TTA	ACT	AAT	CTA	ATT	5159
Val							- /									
1665					1670)				1675	5				1680	
2.07		~~~		> am			~				a b		a a		ABA	r007
ACT Thr																5207
	044		* * * *	1689				1160	1690		un P	5 Bate 14	QT W	1695		
TGT																5255
Сув	Giu	Arg	Thr 1700		Lys	Tyr	Phe	Leu 1709	T	Ile	Ala	GIY	GLY 1710		Trp	
			1100	,				1/05					τ / ± ζ			•
GTA	GTT	AGC	TAT	TTC	TGG	GTG	ACC	CAG	TCT	ATT	ААА	GAA	AGA	AAA	ATG	5303
Val	Val			Phe	Trp	Val			Ser	Ile	Lys		-	Lys	Met	
		1715	5				1720)				1725	5			
CTG	AAT	GAG	CAT	GAT	TTT	GAA	GTC	AGA	GGA	GAT	GTG	GTC	AAT	GGA	AGA	5351
Leu																
	1730)				1735	5				1740)				
	~~~	(1) N	aam		220	002	003	3 (13)	<b>333</b>	<b></b>	<u></u>	<b>~</b> 1 <b>~</b>	202	770	እጥር	5399
AAC Asn																
1745		•			1750				<u>u</u> _u	175					1760	
				•												
TTC																5447
Phe	Arg	GTY	Leu	GIU 1769		Cys	Сув	TYT	GIY 1770		Pne	Thr	Asn			
										1						
202							•			)				1775		
ACA	GAT	CAA	CTG	GAA	TGG	ATG	GTA	CAG			GGT	GCT	TCT			5495
Thr			Leu	Glu		10 A		Gln	CTG Leu	TGT			Ser	GTG Val	GTG	5495
				Glu		10 A			CTG Leu	TGT				GTG Val	GTG	5495
Thr	Asp	Gln	Leu 178(	Glu )	Trp	Met	Val	Gln 1789	CTG Leu 5	TGT Cys	Gly	Ala	Ser 1790	GTG Val	GTG Val	5495 5543
Thr AAG	Asp GAG	Gln CTT	Leu 1780 TCA	Glu ) TCA	Trp TTC	Met ACC	Val CTT	Gln 1789 GGC	CTG Leu 5	TGT Cys GGT	Gly	Ala CAC	Ser 1790 CCA	GTG Val ) ATT	GTG Val GTG	
Thr	Asp GAG	Gln CTT	Leu 1780 TCA Ser	Glu ) TCA	Trp TTC	Met ACC	Val CTT	Gln 1789 GGC Gly	CTG Leu 5	TGT Cys GGT	Gly GTC Val	Ala CAC	Ser 1790 CCA Pro	GTG Val ) ATT	GTG Val GTG	
Thr AAG Lys	Asp GAG Glu	Gln CTT Leu 1795	Leu 1780 TCA Ser	Glu TCA Ser	Trp TTC Phe	Met ACC Thr	Val CTT Leu 1800	Gln 1785 GGC Gly	CTG Leu 5 ACA Thr	TGT Cys GGT Gly	Gly GTC Val	Ala CAC His 1809	Ser 1790 CCA Pro	GTG Val ) ATT Ile	GTG Val GTG Val	5543
Thr AAG Lys GTT	Asp GAG Glu GTG	Gln CTT Leu 1795 CAG	Leu 1780 TCA Ser	Glu TCA Ser GAT	Trp TTC Phe GCC	Met ACC Thr TGG	Val CTT Leu 1800 ACA	Gln 1785 GGC Gly GAG	CTG Leu 5 ACA Thr GAC	TGT Cys GGT Gly AAT	Gly GTC Val GGC	Ala CAC His 1809 TTC	Ser 1790 CCA Pro S	GTG Val ATT Ile GCA	GTG Val GTG Val ATT	
Thr AAG Lys	Asp GAG Glu GTG	Gln CTT Leu 1795 CAG Gln	Leu 1780 TCA Ser	Glu TCA Ser GAT	Trp TTC Phe GCC	Met ACC Thr TGG	Val CTT Leu 1800 ACA Thr	Gln 1785 GGC Gly GAG	CTG Leu 5 ACA Thr GAC	TGT Cys GGT Gly AAT	Gly GTC Val GGC	Ala CAC His 1809 TTC Phe	Ser 1790 CCA Pro S	GTG Val ATT Ile GCA	GTG Val GTG Val ATT	5543
Thr AAG Lys GTT	Asp GAG Glu GTG Val	Gln CTT Leu 1795 CAG Gln	Leu 1780 TCA Ser	Glu TCA Ser GAT	Trp TTC Phe GCC	Met ACC Thr TGG Trp	Val CTT Leu 1800 ACA Thr	Gln 1785 GGC Gly GAG	CTG Leu 5 ACA Thr GAC	TGT Cys GGT Gly AAT	Gly GTC Val GGC Gly	Ala CAC His 1809 TTC Phe	Ser 1790 CCA Pro S	GTG Val ATT Ile GCA	GTG Val GTG Val ATT	5543 5591
Thr AAG Lys GTT Val GGG	Asp GAG Glu GTG Val 1810 CAG	Gln CTT Leu 1795 CAG Gln ATG	Leu 1780 TCA Ser CCA Pro TGT	Glu TCA Ser GAT Asp GAG	Trp TTC Phe GCC Ala GCA	Met ACC Thr TGG Trp 1813 CCT	Val CTT Leu 1800 ACA Thr 5 GTG	Gln 1785 GGC Gly GAG Glu GTG	CTG Leu ACA Thr GAC Asp ACC	TGT Cys GGT Gly AAT Asn CGA	Gly GTC Val GGC Gly 1820 GAG	Ala CAC His 1809 TTC Phe TGG	Ser 1790 CCA Pro CAT His GTG	GTG Val ATT Ile GCA Ala TTG	GTG Val GTG Val ATT Ile GAC	5543
Thr AAG Lys GTT Val GGG Gly	Asp GAG Glu GTG Val 1810 CAG Gln	Gln CTT Leu 1795 CAG Gln ATG	Leu 1780 TCA Ser CCA Pro TGT	Glu TCA Ser GAT Asp GAG	Trp TTC Phe GCC Ala GCA Ala	Met ACC Thr TGG Trp 1815 CCT Pro	Val CTT Leu 1800 ACA Thr 5 GTG	Gln 1785 GGC Gly GAG Glu GTG	CTG Leu ACA Thr GAC Asp ACC	TGT Cys GGT Gly AAT Asn CGA Arg	Gly GTC Val GGC Gly 1820 GAG Glu	Ala CAC His 1809 TTC Phe TGG	Ser 1790 CCA Pro CAT His GTG	GTG Val ATT Ile GCA Ala TTG	GTG Val GTG Val ATT Ile GAC Asp	5543 5591
Thr AAG Lys GTT Val GGG	Asp GAG Glu GTG Val 1810 CAG Gln	Gln CTT Leu 1795 CAG Gln ATG	Leu 1780 TCA Ser CCA Pro TGT	Glu TCA Ser GAT Asp GAG	Trp TTC Phe GCC Ala GCA	Met ACC Thr TGG Trp 1815 CCT Pro	Val CTT Leu 1800 ACA Thr 5 GTG	Gln 1785 GGC Gly GAG Glu GTG	CTG Leu ACA Thr GAC Asp ACC	TGT Cys GGT Gly AAT Asn CGA	Gly GTC Val GGC Gly 1820 GAG Glu	Ala CAC His 1809 TTC Phe TGG	Ser 1790 CCA Pro CAT His GTG	GTG Val ATT Ile GCA Ala TTG	GTG Val GTG Val ATT Ile GAC	5543 5591
Thr AAG Lys GTT Val GGG Gly 1825 AGT	Asp GAG Glu GTG Val 1810 CAG Gln GTA	Gln CTT Leu 1795 CAG Gln Met GCA	Leu 1780 TCA Ser CCA Pro TGT Cys CTC	Glu TCA Ser GAT Asp GAG Glu TAC	Trp TTC Phe GCC Ala GCA Ala 1830 CAG	Met ACC Thr TGG Trp 1813 CCT Pro	Val CTT Leu 1800 ACA Thr GTG Val CAG	Gln 1789 GGC Gly GAG Glu GTG Val GAG	CTG Leu ACA Thr GAC Asp ACC Thr CTG	TGT Cys GGT Gly AAT Asn CGA Arg 1833 GAC	Gly GTC Val GGC Gly 1820 GAG Glu 5 ACC	Ala CAC His 1809 TTC Phe TGG Trp TAC	Ser 1790 CCA Pro CAT His GTG Val CTG	GTG Val ) ATT Ile GCA Ala TTG Leu ATA	GTG Val GTG Val ATT Ile GAC Asp 1840 CCC	5543 5591
Thr AAG Lys GTT Val GGG Gly 1825 AGT	Asp GAG Glu GTG Val 1810 CAG Gln GTA	Gln CTT Leu 1795 CAG Gln Met GCA	Leu 1780 TCA Ser CCA Pro TGT Cys CTC	Glu TCA Ser GAT Asp GAG Glu TAC	Trp TTC Phe GCC Ala GCA Ala 1830 CAG Gln	Met ACC Thr TGG Trp 1813 CCT Pro	Val CTT Leu 1800 ACA Thr GTG Val CAG	Gln 1789 GGC Gly GAG Glu GTG Val GAG	CTG Leu ACA Thr GAC Asp ACC Thr CTG	TGT Cys GGT Gly AAT Asn CGA Arg 183 GAC Asp	Gly GTC Val GGC Gly 1820 GAG Glu 5 ACC	Ala CAC His 1809 TTC Phe TGG Trp TAC	Ser 1790 CCA Pro CAT His GTG Val CTG	GTG Val ) ATT Ile GCA Ala TTG Leu ATA	GTG Val GTG Val ATT Ile GAC Asp 1840 CCC Pro	5543 5591 5639

CAG ATC CCC CAC AGC CAC TAC TGA CTGCAGCCAG CCACAGGTAC AGAGCCACAG Gln Ile Pro His Ser His Tyr * 

-123-. •

GACCCCAAGA ATGAGCTTAC AAAGTGGCCT TTCCAGGCCC TGGGAGCTCC TCTCACTCTT CAGTCCTTCT ACTGTCCTGG CTACTAAATA TTTTATGTAC ATCAGCCTGA AAAGGACTTC TGGCTATGCA AGGGTCCCTT AAAGATTTTC TGCTTGAAGT CTCCCTTGGA AA 

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1864 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2: Met Asp Leu Ser Ala Leu Arg Val Glu Glu Val Gln Asn Val Ile Asn Ala Met Gln Lys Ile Leu Glu Cys Pro Ile Cys Leu Glu Leu Ile Lys Glu Pro Val Ser Thr Lys Cys Asp His Ile Phe Cys Lys Phe Cys Met Leu Lys Leu Leu Asn Gln Lys Lys Gly Pro Ser Gln Cys Pro Leu Cys Lys Asn Asp Ile Thr Lys Arg Ser Leu Gln Glu Ser Thr Arg Phe Ser Gln Leu Val Glu Glu Leu Leu Lys Ile Ile Cys Ala Phe Gln Leu Asp Thr Gly Leu Glu Tyr Ala Asn Ser Tyr Asn Phe Ala Lys Lys Glu Asn Asn Ser Pro Glu His Leu Lys Asp Glu Val Ser Ile Ile Gln Ser Met Gly Tyr Arg Asn Arg Ala Lys Arg Leu Leu Gln Ser Glu Pro Glu Asn Pro Ser Leu Gln Glu Thr Ser Leu Ser Val Gln Leu Ser Asn Leu Gly Thr Val Arg Thr Leu Arg Thr Lys Gln Arg Ile Gln Pro Gln Lys Thr 

Ser Val Tyr Ile Glu Leu Gly Ser Asp Ser Ser Glu Asp Thr Val Asn Lys Ala Thr Tyr Cys Ser Val Gly Asp Gln Glu Leu Leu Gln Ile Thr Pro Gln Gly Thr Arg Asp Glu Ile Ser Leu Asp Ser Ala Lys Lys Ala Ala Cys Glu Phe Ser Glu Thr Asp Val Thr Asn Thr Glu His His Gln Pro Ser Asn Asn Asp Leu Asn Thr Thr Glu Lys Arg Ala Ala Glu Arg His Pro Glu Lys Tyr Gln Gly Ser Ser Val Ser Asn Leu His Val Glu Pro Cys Gly Thr Asn Thr His Ala Ser Ser Leu Gln His Glu Asn Ser Ser Leu Leu Thr Lys Asp Arg Met Asn Val Glu Lys Ala Glu Phe Cys Asn Lys Ser Lys Gln Pro Gly Leu Ala Arg Ser Gln His Asn Arg Trp Ala Gly Ser Lys Glu Thr Cys Asn Asp Arg Arg Thr Pro Ser Thr Glu Lys Lys Val Asp Leu Asn Ala Asp Pro Leu Cys Glu Arg Lys Glu Trp Asn Lys Gln Lys Leu Pro Cys Ser Glu Asn Pro Arg Asp Thr Glu Asp Val Pro Trp Ile Thr Leu Asn Ser Ser Ile Gln Lys Val Asn Glu Trp Phe Ser Arg Ser Asp Glu Leu Leu Gly Ser Asp Asp Ser His Asp Gly Glu Ser Glu Ser Asn Ala Lys Val Ala Asp Val Leu Asp Val Leu Asn Glu Val Asp Glu Tyr Ser Gly Ser Ser Glu Lys Ile Asp Leu Leu Ala Ser Asp Pro His Glu Ala Leu Ile Cys Lys Ser Glu Arg Val His Ser Lys Ser Val Glu Ser Asn Ile Glu Asp Lys Ile Phe Gly Lys Thr Tyr Arg Lys Lys Ala Ser Leu Pro Asn Leu Ser His Val Thr Glu Asn 

-124-

Leu Ile Ile Gly Ala Phe Val Thr Glu Pro Gln Ile Ile Gln Glu Arg Pro Leu Thr Asn Lys Leu Lys Arg Lys Arg Arg Pro Thr Ser Gly Leu His Pro Glu Asp Phe Ile Lys Lys Ala Asp Leu Ala Val Gln Lys Thr . Pro Glu Met Ile Asn Gln Gly Thr Asn Gln Thr Glu Gln Asn Gly Gln Val Met Asn Ile Thr Asn Ser Gly His Glu Asn Lys Thr Lys Gly Asp Ser Ile Gln Asn Glu Lys Asn Pro Asn Pro Ile Glu Ser Leu Glu Lys Glu Ser Ala Phe Lys Thr Lys Ala Glu Pro Ile Ser Ser Ser Ile Ser Asn Met Glu Leu Glu Leu Asn Ile His Asn Ser Lys Ala Pro Lys Lys Asn Arg Leu Arg Arg Lys Ser Ser Thr Arg His Ile His Ala Leu Glu Leu Val Val Ser Arg Asn Leu Ser Pro Pro Asn Cys Thr Glu Leu Gln Ile Asp Ser Cys Ser Ser Ser Glu Glu Ile Lys Lys Lys Tyr Asn Gln Met Pro Val Arg His Ser Arg Asn Leu Gln Leu Met Glu Gly Lys Glu Pro Ala Thr Gly Ala Lys Lys Ser Asn Lys Pro Asn Glu Gln Thr Ser Lys Arg His Asp Ser Asp Thr Phe Pro Glu Leu Lys Leu Thr Asn Ala Pro Gly Ser Phe Thr Lys Cys Ser Asn Thr Ser Glu Leu Lys Glu Phe Val Asn Pro Ser Leu Pro Arg Glu Glu Lys Glu Glu Lys Leu Glu Thr Val Lys Val Ser Asn Asn Ala Glu Asp Pro Lys Asp Leu Met Leu Ser Gly Glu Arg Val Leu Gln Thr Glu Arg Ser Val Glu Ser Ser Ser 

-125-

•

Ile Ser Leu Val Pro Gly Thr Asp Tyr Gly Thr Gln Glu Ser Ile Ser Leu Leu Glu Val Ser Thr Leu Gly Lys Ala Lys Thr Glu Pro Asn Lys Cys Val Ser Gln Cys Ala Ala Phe Glu Asn Pro Lys Gly Leu Ile His Gly Cys Ser Lys Asp Asn Arg Asn Asp Thr Glu Gly Phe Lys Tyr Pro Leu Gly His Glu Val Asn His Ser Arg Glu Thr Ser Ile Glu Met Glu Glu Ser Glu Leu Asp Ala Gln Tyr Leu Gln Asn Thr Phe Lys Val Ser Lys Arg Gln Ser Phe Ala Pro Phe Ser Asn Pro Gly Asn Ala Glu Glu Glu Cys Ala Thr Phe Ser Ala His Ser Gly Ser Leu Lys Lys Gln Ser Pro Lys Val Thr Phe Glu Cys Glu Gln Lys Glu Glu Asn Gln Gly Lys Asn Glu Ser Asn Ile Lys Pro Val Gln Thr Val Asn Ile Thr Ala Gly Phe Pro Val Val Gly Gln Lys Asp Lys Pro Val Asp Asn Ala Lys Cys Ser Ile Lys Gly Gly Ser Arg Phe Cys Leu Ser Ser Gln Phe Arg Gly Asn Glu Thr Gly Leu Ile Thr Pro Asn Lys His Gly Leu Leu Gln Asn Pro Tyr Arg Ile Pro Pro Leu Phe Pro Ile Lys Ser Phe Val Lys Thr Lys Cys Lys Lys Asn Leu Leu Glu Glu Asn Phe Glu Glu His Ser Met Ser Pro Glu Arg Glu Met Gly Asn Glu Asn Ile Pro Ser Thr Val Ser Thr Ile Ser Arg Asn Asn Ile Arg Glu Asn Val Phe Lys Glu Ala Ser Ser Ser Asn Ile Asn Glu Val Gly Ser Ser Thr Asn Glu Val Gly Ser Ser Ile Asn Glu Ile Gly Ser Ser Asp Glu Asn Ile Gln Ala Glu Leu 

-126-

• •

Gly Arg Asn Arg Gly Pro Lys Leu Asn Ala Met Leu Arg Leu Gly Val Leu Gln Pro Glu Val Tyr Lys Gln Ser Leu Pro Gly Ser Asn Cys Lys His Pro Glu Ile Lys Lys Gln Glu Tyr Glu Glu Val Val Gln Thr Val Asn Thr Asp Phe Ser Pro Tyr Leu Ile Ser Asp Asn Leu Glu Gln Pro Met Gly Ser Ser His Ala Ser Gln Val Cys Ser Glu Thr Pro Asp Asp Leu Leu Asp Asp Gly Glu Ile Lys Glu Asp Thr Ser Phe Ala Glu Asn Asp Ile Lys Glu Ser Ser Ala Val Phe Ser Lys Ser Val Gln Lys Gly Glu Leu Ser Arg Ser Pro Ser Pro Phe Thr His Thr His Leu Ala Gln Gly Tyr Arg Arg Gly Ala Lys Lys Leu Glu Ser Ser Glu Glu Asn Leu Ser Ser Glu Asp Glu Glu Leu Pro Cys Phe Gln His Leu Leu Phe Gly Lys Val Asn Asn Ile Pro Ser Gln Ser Thr Arg His Ser Thr Val Ala Thr Glu Cys Leu Ser Lys Asn Thr Glu Glu Asn Leu Leu Ser Leu Lys Asn Ser Leu Asn Asp Cys Ser Asn Gln Val Ile Leu Ala Lys Ala Ser Gln Glu His His Leu Ser Glu Glu Thr Lys Cys Ser Ala Ser Leu Phe Ser Ser Gln Cys Ser Glu Leu Glu Asp Leu Thr Ala Asn Thr Asn Thr Gln Asp Pro Phe Leu Ile Gly Ser Ser Lys Gln Met Arg His Gln Ser Glu Ser Gln Gly Val Gly Leu Ser Asp Lys Glu Leu Val Ser Asp Asp Glu Glu Arg Gly Thr Gly Leu Glu Glu Asn Asn Gln Glu Glu Gln Ser 

. -

Met Asp Ser Asn Leu Gly Glu Ala Ala Ser Gly Cys Glu Ser Glu Thr Ser Val Ser Glu Asp Cys Ser Gly Leu Ser Ser Gln Ser Asp Ile Leu Thr Thr Gln Gln Arg Asp Thr Met Gln His Asn Leu Ile Lys Leu Gln Gln Glu Met Ala Glu Leu Glu Ala Val Leu Glu Gln His Gly Ser Gln Pro Ser Asn Ser Tyr Pro Ser Ile Ile Ser Asp Ser Ser Ala Leu Glu Asp Leu Arg Asn Pro Glu Gln Ser Thr Ser Glu Lys Ala Val Leu Thr Ser Gln Lys Ser Ser Glu Tyr Pro Ile Ser Gln Asn Pro Glu Gly Leu Ser Ala Asp Lys Phe Glu Val Ser Ala Asp Ser Ser Thr Ser Lys Asn Lys Glu Pro Gly Val Glu Arg Ser Ser Pro Ser Lys Cys Pro Ser Leu Asp Asp Arg Trp Tyr Met His Ser Cys Ser Gly Ser Leu Gln Asn Arg Asn Tyr Pro Ser Gln Glu Glu Leu Ile Lys Val Val Asp Val Glu Glu Gln Gln Leu Glu Glu Ser Gly Pro His Asp Leu Thr Glu Thr Ser Tyr Leu Pro Arg Gln Asp Leu Glu Gly Thr Pro Tyr Leu Glu Ser Gly Ile 1565 . Ser Leu Phe Ser Asp Asp Pro Glu Ser Asp Pro Ser Glu Asp Arg Ala Pro Glu Ser Ala Arg Val Gly Asn Ile Pro Ser Ser Thr Ser Ala Leu Lys Val Pro Gln Leu Lys Val Ala Glu Ser Ala Gln Ser Pro Ala Ala Ala His Thr Thr Asp Thr Ala Gly Tyr Asn Ala Met Glu Glu Ser Val Ser Arg Glu Lys Pro Glu Leu Thr Ala Ser Thr Glu Arg Val Asn Lys Arg Met Ser Met Val Val Ser Gly Leu Thr Pro Glu Glu Phe Met Leu 

-128-

Val Tyr Lys Phe Ala Arg Lys His His Ile Thr Leu Thr Asn Leu Ile Thr Glu Glu Thr Thr His Val Val Met Lys Thr Asp Ala Glu Phe Val Cys Glu Arg Thr Leu Lys Tyr Phe Leu Gly Ile Ala Gly Gly Lys Trp Val Val Ser Tyr Phe Trp Val Thr Gln Ser Ile Lys Glu Arg Lys Met Leu Asn Glu His Asp Phe Glu Val Arg Gly Asp Val Val Asn Gly Arg Asn His Gln Gly Pro Lys Arg Ala Arg Glu Ser Gln Asp Arg Lys Ile Phe Arg Gly Leu Glu Ile Cys Cys Tyr Gly Pro Phe Thr Asn Met Pro Thr Asp Gln Leu Glu Trp Met Val Gln Leu Cys Gly Ala Ser Val Val Lys Glu Leu Ser Ser Phe Thr Leu Gly Thr Gly Val His Pro Ile Val Val Val Gln Pro Asp Ala Trp Thr Glu Asp Asn Gly Phe His Ala Ile Gly Gln Met Cys Glu Ala Pro Val Val Thr Arg Glu Trp Val Leu Asp Ser Val Ala Leu Tyr Gln Cys Gln Glu Leu Asp Thr Tyr Leu Ile Pro Gln Ile Pro His Ser His Tyr * (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

(vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens -129-

. -



-130-

(vii) IMMEDIATE SOURCE: (B) CLONE: \$754 A

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

CTAGCCTGGG CAACAAACGA

(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
- (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE: (B) CLONE: s754 B

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

## GCAGGAAGCA GGAATGGAAC

(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
  - (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE: (B) CLONE: \$975 A
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TAGGAGATGG ATTATTGGTG

20

20

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs

-131-

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
   (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE: (B) CLONE: \$975 B

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

#### AGGCAACTTT GCAATGAGTG

(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
- (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
- (vii) IMMEDIATE SOURCE: (B) CLONE: tdj1474 A

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

#### CAGAGTGAGA CCTTGTCTCA AA

- (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

22

. 20



-132-. ...

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: 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

- (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
  - (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
  - (vii) IMMEDIATE SOURCE: (B) CLONE: tdj1239 A

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

#### AACCTGGAAG GCAGAGGTTG

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

23



-133-

(vii) IMMEDIATE SOURCE: (B) CLONE: tdj1239 B

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

TCTGTACCTG CTAAGCAGTG G

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 111 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
  - (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
- (ix) FEATURE:
  - (A) NAME/KEY: CDS

(B) LOCATION: 2..111

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

G GKC TTA CTC TGT TGT CCC AGC TGG AGT ACA GWG TGC GAT CAT GAG Xaa Leu Leu Cys Cys Pro Ser Trp Ser Thr Xaa Cys Asp His Glu 1865 1870 1875

GCT TAC TGT TGC TTG ACT CCT AGG CTC AAG CGA TCC TAT CAC CTC AGTAla Tyr Cys Cys Leu Thr Pro Arg Leu Lys Arg Ser Tyr His Leu Ser1880188518901895

CTC CAA GTA GCT GGA CT Leu Gln Val Ala Gly 1900

### (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 amino acids
    - (B) TYPE: amino acid
    - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein





46

94

111 .

21

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

Xaa Leu Leu Cys Cys Pro Ser Trp Ser Thr Xaa Cys Asp His Glu Ala151015Tyr Cys Cys Leu Thr Pro Arg Leu Lys Arg Ser Tyr His Leu Ser Leu202530

Gln Val Ala Gly 35

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1540 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

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

GAGGCTAGAG	GGCAGGCACT	TTATGGCAAA	CTCAGGTAGA	ATTCTTCCTC	TTCCGTCTCT	60
TTCCTTTTAC	GTCATCCGGG	GGCAGACTGG	GTGGCCAATC	CAGAGCCCCG	AGAGACGCTT	120
GGCTCTTTCT	GTCCCTCCCA	TCCTCTGATT	GTACCTTGAT	TTCGTATTCT	GAGAGGCTGC	180
TGCTTAGCGG	TAGCCCCTTG	GTTTCCGTGG	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	GGTCGCGAGG	540
GAAGCGCTGA	GGATCAGGAA	GGGGGCACTG	AGTGTCCGTG	GGGGAATCCT	CGTGATAGGA	600
ACTGGAATAT	GCCTTGAGGG	GGACACTATG	тстттааааа	CGTCGGCTGG	TCATGAGGTC	660
AGGAGTTCCA	GACCAGCCTG	ACCAACGTGG	TGAAACTCCG	TCTCTACTAA	AAATACNAAA	720

# -134-

ATTAGCCGGG CGTGGTGCCG CTCCAGCTAC TCAGGAGGCT GAGGCAGGAG AATCGCTAGA 780 ACCCGGGAGG CGGAGGTTGC AGTGAGCCGA GATCGCGCCA TTGCACTCCA GCCTGGGCGA 840 900 TGTATGAGAG GATGGGACCT TGTGGAAGAA GAGGTGCCAG GAATATGTCT GGGAAGGGGA 960 GGAGACAGGA TTTTGTGGGA GGGAGAACTT AAGAACTGGA TCCATTTGCG CCATTGAGAA 1020 AGCGCAAGAG GGAAGTAGAG GAGCGTCAGT AGTAACAGAT GCTGCCGGCA GGGATGTGCT 1080 TGAGGAGGAT CCAGAGATGA GAGCAGGTCA CTGGGAAAGG TTAGGGGCGG GGAGGCCTTG 1140 ATTGGTGTTG GTTTGGTCGT TGTTGATTTT GGTTTTATGC AAGAAAAAGA AAACAACCAG 1200 AAACATTGGR GAAAGCTAAG GCTACCACCA CCTACCCGGT CAGTCACTCC TCTGTAGCTT 1260 1320 TCTCTTYCTT GGAGAAAGGA AAAGACCCAA GGGGTTGGCA GCRATATGTG AAAAAATTCA 1380 GAATTTATGT TGTCTAATTA CAAAAAGCAA CTTCTAGAAT CTTTAAAAAT AAAGGACGTT GTCATTAGTT CTTTGGTTTG TATYATTCTA AAACCTTCCA AATCTTAAAT TTACTTTATT 1440 TTAAAATGAT AAAATGAAGT TGTCATTTTA TAAACCTTTT AAAAAGATAT ATATATATGT 1500 1540 TTTTCTAATG TGTTANAGTT CATTGGAACA GAAAGAAATG

-135-

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21715 base pairs
    - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (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

-136-

AGATAACTGG GCCCCTGCGC TCAGGAGGCC TTCACCCTCT GCTCTGGGTA AAGVVVVVAA 300 AAATAAAGGA CGCTGTCGAY TAGTTCTTCT GGTTTGTATT ATTCTMAAAC CTTCCAAATC 360 TTCAAATTTA CTTYATTTTA AAATGATGAA AATGAAGTTG TCATTTTATA AACCTTTTAA 420 AAAGATATAT ATATATGTTT TTCTAATGTG TTAAAGTTCA TTGGAACAGA AAGAAATGGA 480 TTTATCTGCT CTTCGCGTTG AAGAAGTACA AAATGTCATT AATGCTATGC AGAAAATCTT 540 AGAGTGTCCC ATCTGGTAAG TCAGCACAAG AGTGTATTAA 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 GGTTCTGCAC TGGGCAGGAT CTGTTAGATA 840 CTGCATTTGC TTTCTCCAGC TCTAAAVVVV VVVVVVAA ATGCTGATGA TAGTATAGAG 900 TATTGAAGGG ATCAATATAA TTCTGTTTTG ATATCTGAAA GCTCACTGAA GGTAAGGATC 960 GTATTCTCTG CTGTATTCTC AGTTCCTGAC ACAGCAGACA TTTAATAAAT ATTGAACGAA 1020 CTTGAGGCCT TATGTTGACT CAGTCATAAC AGCTCAMAGT TGAACTTATT CACTAAGAAT 1080 AGCTTTATTT YTAAATAAAT TATTGAGCCT CATTTATTTT CTTTTTCTCV VVVVTCTGGA 1140 GTTGATCAAG GAACCTGTCT CCACAAAGTG TGACCACATA TTTTGCAAVV VVVATGTGGC 1200 TCCATTATTA GCTTTTGTTT TTGTCCTTCA TAACCCAGGA AACACCTAAC TTTATAGAAG 1260 1320 CTTTACTTTC TTCAATTAAG TGAGAACGAA AATCCAACTC CATTTCATTC TTTCTCAGAG AGTATATAGT TATCAAAAGT TGGTTGTAAT CATAGTTCCT GGTAAAGTTT TGACATATAT 1380 TATCTTTTTT TTTTTTGAG ACAAGTCTCG CTCTGTCGCC CAGGCTGGAG TGCAGTGGCA 1440 TGAGGCTTGC TCACTGCACC TCCGCCCCCG AGTTCAGCGA CTCTVVVVV VVVVVTGA 1500 GATCTAGACC ACATGGTCAA AGAGATAGAA TGTGAGCAAT AAATGAACCT TAAATTTTTC 1560 1620 TACAGWGTGC GATCATGAGG CTTACTGTTG CTTGACTCCT AGGCTCAAGC GATCCTATCA 1680 1740 CCTCAGTCTC CAAGTAGCTG GACTGTAAGT GCACACCACC ATATCCAGCT AAATTTTGTG TTTTCTGTAG AGACGGGGTT TCGCCATGTT TCCCAGGCTG GTCTTGAACT TTGGGCTTAA 1800 CCCGTCTGCC CACCTAGGCA TCCCAAAGTG CTAGGATTAC AGGTGTGAGT CATCATGCCT 1860 GGCCAGTATT TTAGTTAGCT CTGTCTTTTC AAGTCATATA CAAGTTCATT TTCTTTAAGT 1920

TAGTTAACAA CCTTATATCA TGTATTCTTT TCTAGCAVVV VVVVVVVV TGTGATCATA 1980 ACAGTAAGCC ATATGCATGT AAGTTCAGTT TCATAGATCA TTGCTTATGT AGTTTAGGTT 2040 TTTGCTTATG CAGCATCCAA AAACAATTAG GAAACTATTG CTTGTAATTC ACCTGCCATT 2100 ACTITITAAA TGGCTCTTAA GGGCAGTTGT GAGATTATCT TTTCATGGCT ATTTGCCTTT 2160 TGAGTATTCT TTCTACAAAA GGAAGTAAAT TAAATTGTTC TTTCTTTCTT TATAATTTAT 2220 AGATVVVVA TTTTGCATGC TGAAACTTCT CAACCAGAAG AAAGGGCCTT CACAGTGTCC 2280 TTTATGTAAG AATGATATAA CCAAAAGGAT ATAANNNGAA NTGATGCTAG GTTGGAAGCA 2340 ACCACAGTAG GAAAAGTAGA AATTATTTAA TAACATAGCG TTCCTATAAA ACCATTCATC 2400 AGAAAAATTT ATAAAAGAGT TTTTAGCACA CAGTAAATTA TTTCCAAAGT TATTTTCCTG 2460 AAAGTTTTAT GGGCATCTGC CTTATACAGG TNTTGVVVVV VVVVVVGG TAGGCTTAAA 2520 TGAATGACAA AAAGTTACTA AATCACTGCC ATCACACGGT TATATACAGA TGTCAATGAT 2580 2640 GTATTGATTT AGAGGTTTTC TACTGTTGCT GCATCTTATT TTTATTTGTT TACATGTCTT TTCTTATTTT AGTGTCCTTA AAAGGTTGAT AATCACTTGC TGAGTGTGTT TCTCAAAVVV 2700 VVGAGCCTAC AAGAAAGTAC GAGATTTAGT CAACTTGTTG AAGAGCTATT GAAAATCATT 2760 TGTGCTTTTC AGCTTGACAC AGGTTTGGAG TATGCAAACA GCTATAATTT TGCAAAAAAG 2820 2880 GAAAATAACT CTCCTGAACA TCTAAAAGAT GAAGTTTCTA TCATCCAAAG TATGGGCTAC AGAAACCGTG CCAAAAGACT TCTACAGAGT GAACCCGAAA ATCCTTCCTT GCAGAAACCA 2940 3000 3060 GTGGAGTCYT TGGCTCGTSS YGGCGCCGAG GCTAGGAAGC AGCTCCTCCT GCGCTCGTAC 3120 GCCACGCTAG TAGGTAGCGT GGGCAKACAC GGCACGCGCA CCATGGCCAG GCTAATTTTT GTATTGTTVV VVVVVVVV VATGATGGAG ATCTTAAAAA GTAATCATTC TGGGGCTGGG 3180 CGTAGTAGCT TGCACCTGTA ATCCCAGCAC TTCGGGAGGC 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 3540 AGCTGGGCGC AGTGGCTCTT GCCTGTAATC CCAGCACTTT GGGAGGCCAA GGCGGGTGGA 3600 TCTCGAGTCC TAGAGTTCAA GACCAGCCTA GGCAATGTGG TGAAACCCCA TCGCTACAAA

-137-

3660 AAATACAAAA ATTAGCCAGG CATGGTGGCG TGCGCATGTA GTCCCAGCTC CTTGGGAGGC TGAGGTGGGA GGATCACTTG AACCCAGGAG ACAGAGGTTG CAGTGAACCG AGATCACGCC 3720 3780 AAGTAGTTCT CACAGTACCA GCATTCATTT TTCAAAAGAT ATAGAGCTAA AAAGGAAGGA 3840 AAAAAAAGT AATGTTGGGC TTTTAAATAC TCGTTCCTAT 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 GGTTGCAGTG TGCCGAGATC ACGCCACTGC ACTCTAGCCT 4200 GGCAACAGAG CAAGATTCTG TCTCAAAAAA AAAAAACATA TATACACATA TATCCCAAAG 4260 4320 ATATATATGT GATATATATA TAACATATAT ATATGTAATA TATATGTGAT ATATATATAA 4380 TATATATATG TAATATATAT GTGATATATA TATATACACA CACACACA TATATATGTA 4440 TGTGTGTGTA CACACACACA CACAAATTAG CCAGGCATAG TTGCACACGC TTGGTAGACC 4500 CAGCTACTCA GGAGGCTGAG GGAGGAGAAT CTCTTGAACT TAGGAGGCGG AGGTTGCAGT 4560 GAGCTGAGAT TGCGCCACTG CACTCCAGCC TGGGTGACAG AGCAGGACTC TGTACACCCC 4620 4680 CCAAAACAAA AAAAAAAGTT ATCAGATGTG ATTGGAATGT ATATCAAGTA TCAGCTTCAA 4740 AATATGCTAT ATTAATACTT CAAAAATTAC ACAAATAATA CATAATCAGG TTTGAAAAAAT TTAAGACAAC MSAARAAAAA AWYCMAATCA CAMATATCCC ACACATTTTA TTATTMCTMC 4800 4860 TMCWATTATT TTGWAGAGMC TGGGTCTCAC YCYKTTGCTW ATGCTGGTCT TTGAACYCCY 4920 KGCCYCAARC ARTCCTSCTC CABCCTCCCA ARGTGCTGGG GATWATAGGC ATGARCTACC 4980 CGCTCCCAGC CCMAGACAAT TTTAGTGTGW GCAAAATTCC TGGGCATTTT TTCMAGGCHD NMWAYAYVTW DSMTGCWHRT GAHBGGCTGA TGWGTCAATT TATTTTGYCC ATGGTGTCAA 5040 KTYTCTCTTC AGGAGGAAAA GCACAMGAAC TGGCCAAYAA TTGCTTGACT GTTCTTTACC 5100 AWACTGTTWA RCAGGAAACC AGTCTCAGTG T-CCAACTCT CTAACCTTGG AACT-GTGAG 5160 AACTCTGAGG ACAAAGCAGC GGATACAACC TCAAAAGACG TCTGTCTACA TTGAATTGGG 5220 TAAGGGTCTC AGGTTTTTTA AGTATTTAAT AATAATTGCT GGATTCCTTA TCTTATAGTT 5280

-138-

-139-

TTGCCAMAAA TCTTGGTCAT AATTTGTATT TGTGGTAGGC AGCWTTGGGA 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 AGAMTMTCTG GTTTTAAMTM TGTGTAAATA 5760 TGTTTTTGGA AAGATGGAGA GTAGCAATAA GAAAAAACAT GATGGATTGC TACAGTATTT 5820 AGTTCCAAGA TAAATTGTAC TAGATGAGGA AGCCTTTTAA GAAGAGCTGA ATTGCCAGGC 5880 GCAGTGGCTC ACGCCTGTAA TCCCAGCACT TTGGGAGGCC GAGGTGGGCG GATCACCTGA 5940 GGTCGGGAGT TCAAGACCAG CCTGACCAAC ATGGAGAAAC CCCATCTCTA CTAAAAAAAA 6000 AAAAAAAAAA ATTAGCCGGG GTGGTGGCTT ATGCCTGTAA TCCCAGCTAC TCAGGAGGCT 6060 GAGGCAGGAG AATCGCTTGA ACCCAGGAAG CAGAGGTTGC AGTGAGCCAA GATCGCACCA 6120 6180 GAATCTTGGC TGGGCAGGAT GGCTCGTGCC TGTAATCCTA ACGCTTTGGA AGACCGAGGC 6240 AGAAGGATTG GTTGAGTCCA CGAGTTTAAG ACCAGCCTGG CCAACATAGG GGAACCCTGT 6300 6360 CTCTATTTTT AAAATAATAA TACATTTTTG-GCCGGTGCGG TGGCTCATGC CTGTAATCCC 6420 AATACTTTGG GAGGCTGAGG CAGGTAGATC ACCTGAGGTC AGAGTTCGAG ACCAGCCTGG 6480 TGGTAGCACA TGCTTGTAAT CCCAGCTACT TGGGAGGCTG AGGCAGGAGA ATCGCTTGAA 6540 6600 CCAGGGAGGC GGAGGTTACA ATGAGCCAAC ACTACACCAC TGCACTCCAG CCTGGGCAAT 6660 AGAGTGAGAC TGCATCTCAA AAAAATAATA ATTTTTAAAA ATAATAAATT TTTTTAAGCT 6720 TATAAAAAGA AAAGTTGAGG CCAGCATAGT AGCTCACATC TGTAATCTCA GCAGTGGCAG AGGATTGCTT GAAGCCAGGA GTTTGAGACC AGCCTGGGCA ACATAGCAAG ACCTCATCTC 6780 6840 TACAAAAAAA TTTCTTTTTT AAATTAGCTG GGTGTGGTGG TGTGCATCTG TAGTCCCAGC 6900 TACTCAGGAG GCAGAGGTGA GTGGATACAT TGAACCCAGG AGTTTGAGGC TGTAGTGAGC 6960 TATGATCATG CCACTGCACT CCAACCTGGG TGACAGAGCA AGACCTCCAA AAAAAAAAA
AAAAGAGCTG CTGAGCTCAG AATTCAAACT GGGCTCTCAA ATTGGATTTT CTTTTAGAAT 7020 ATATTTATAA TTAAAAAGGA TAGCCATCTT TTGAGCTCCC AGGCACCACC ATCTATTTAT 7080 CATAACACTT ACTGTTTTCC CCCCTTATGA TCATAAATTC CTAGACAACA GGCATTGTAA 7140 AAATAGTTAT AGTAGTTGAT ATTTAGGAGC ACTTAACTAT ATTCCAGGCA CTATTGTGCT 7200 TTTCTTGTAT AACTCATTAG ATGCTTGTCA GACCTCTGAG ATTGTTCCTA TTATACTTAT 7260 TTTACAGATG AGAAAATTAA GGCACAGAGA AGTTATGAAA TTTTTCCAAG GTATTAAACC 7320 TAGTAAGTGG CTGAGCCATG ATTCAAACCT AGGAAGTTAG ATGTCAGAGC CTGTGCTTTT 7380 TTTTTGTTTT TGTTTTTGTT TTCAGTAGAA ACGGGGGGTCT CACTTTGTTG GCCAGGCTGG 7440 TCTTGAACTC CTAACCTCAA ATAATCCACC CATCTCGGCC TCCTCAAGTG CTGGGATTAC 7500 7560 AGGTGAGAGC CACTGTGCCT GGCGAAGCCC ATGCCTTTAA CCACTTCTCT GTATTACATA CTAGCTTAAC TAGCATTGTA CCTGCCACAG TAGATGCTCA GTAAATATTT CTAGTTGAAT 7620 ATCTGTTTTT CAACAAGTAC ATTTTTTTAA CCCTTTTAAT TAAGAAAACT TTTATTGATT 7680 TATTTTTTGG GGGGAAATTT TTTAGGATCT GATTCTTCTG AAGATACCGT TAATAAGGCA 7740 ACTTATTGCA GGTGAGTCAA AGAGAACCTT TGTCTATGAA GCTGGTATTT TCCTATTAG 7800 TTAATATTAA GGATTGATGT TTCTCTCTTT TTAAAAATAT TTTAACTTTT ATTTTAGGTT 7860 7920 CAGGGATGTA TGTGCAGTTT GTTATATAGG TAAACACACG ACTTGGGATT TGGTGTATAG 7980 ATTTTTTCA TCATCCGGGT ACTAAGCATA CCCCACAGTT TTTTGTTTGC TTTCTTTCTG 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 8280 GTAATATTGA CACAGCAGTA ATTGTGACTG ATGAAAATGT TCAAGGACTT CATTTTCAAC 8340 TCTTTCTTTC CTCTGTTCCT TATTTCCACA TATCTCTCAA GCTTTGTCTG TATGTTATAT 8400 AATAAACTAC AAGCAACCCC AACTATGTTA CCTACCTTCC TTAGGAATTA TTGCTTGACC 8460 CAGGTTTTTT TTTTTTTTTTTTTTTTTGGAGACG GGGTCTTGCC CTGTTGCCAG GATGGAGTGT 8520 AGTGGCGCCA TCTCGGCTCA CTGCAATCTC CAACTCCCTG GTTCAAGCGA TTCTCCTGTC 8580 TCAATCTCAC GAGTAGCTGG GACTACAGGT ATACACCACC ACGCCCGGTT AATTGACCAT 8640 TCCATTTCTT TCTTTCTCTC TTTTTTTTTT TTTTTTTGA GACAGAGTCT TGCTCTGTTG

-140-

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 ACCATTTCAT TTCAACTAGA AGTTTCTAAA GGAGAGAGCA 8940 GCTTTCACTA ACTAAATAAG ATTGGTCAGC TTTCTGTAAT CGAAAGAGCT AAAATGTTTG 9000 ATCTTGGTCA TTTGACAGTT CTGCATACAT GTAACTAGTG 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 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 GCCCACTTCA GCCTCCCAAA 9660 GTGCTGCAGT TACAGGCTTG AGCCACTGTG CCTGGCCTGC CCTTTACTTT TAATTGGTGT 9720 ATTTGTGTTT CATCTTTTAC CTACTGGTTT TTAAATATAG GGAGTGGTAA GTCTGTAGAT 9780 9840 AGAACAGAGT ATTAAGTAGA CTTAATGGCC AGTAATCTTT AGAGTACATC AGAACCAGTT TTCTGATGGC CAATCTGCTT TTAATTCACT CTTAGACGTT AGAGAAATAG GTGTGGTTTC 9900 TGCATAGGGA AAATTCTGAA ATTAAVVVVV VVVVVVVGA 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 AACTTGCATG TGGAGCCATG TGGCACAAAT -ACTCATGCC AGCTCATTAC 10320

-141-

AGCATGAGAA CAGCAGTTTA TTACTCACTA AAGACAGAAT GAATGTAGAA AAGGCTGAAT 10380 TCTGTAATAA AAGCAAACAG CCTGGCTTAG CAAGGAGCCA ACATAACAGA TGGGCTGGAA 10440 GTAAGGAAAC ATGTAATGAT AGGCGGACTC CCAGCACAGA AAAAAAGGTA GATCTGAATG 10500 CTGATCCCCT GTGTGAGAGAA AAAGAATGGA ATAAGCAGAA ACTGCCATGC TCAGAGAATC 10560 CTAGAGATAC TGAAGATGTT CCTTGGATAA CACTAAATAG CAGCATTCAG AAAGTTAATG 10620 AGTGGTTTTC CAGAAGTGAT GAACTGTTAG GTTCTGATGA CTCACATGAT GGGGAGTCTG 10680 AATCAAATGC CAAAGTAGCT GATGTATTGG ACGTTCTAAA TGAGGTAGAT GAATATTCTG 10740 GTTCTTCAGA GAAAATAGAC TTACTGGCCA GTGATCCTCA TGAGGCTTTA ATATGTAAAA 10800 GTGAAAGAGT TCACTCCAAA TCAGTAGAGA GTAATATTGA AGACAAAATA TTTGGGAAAA 10860 CCTATCGGAA GAAGGCAAGC CTCCCCAACT TAAGCCATGT AACTGAAAAT CTAATTATAG 10920 GAGCATTTGT TACTGAGCCA CAGATAATAC AAGAGCGTCC CCTCACAAAT AAATTAAAGC 10980 GTAAAAGGAG ACCTACATCA GGCCTTCATC CTGAGGATTT TATCAAGAAA GCAGATTTGG 11040 CAGTTCAAAA GACTCCTGAA ATGATAAATC AGGGAACTAA CCAAACGGAG CAGAATGGTC 11100 AAGTGATGAA TATTACTAAT AGTGGTCATG AGAATAAAAC AAAAGGTGAT TCTATTCAGA 11160 ATGAGAAAAA TCCTAACCCA ATAGAATCAC TCGAAAAAGA ATCTGCTTTC AAAACGAAAG 11220 CTGAACCTAT AAGCAGCAGT ATAAGCAATA TGGAACTCGA ATTAAATATC CACAATTCAA 11280 AAGCACCTAA AAAGAATAGG CTGAGGAGGA AGTCTTCTAC CAGGCATATT CATGCGCTTG 11340 AACTAGTAGT CAGTAGAAAT CTAAGCCCAC CTAATTGTAC TGAATTGCAA ATTGATAGTT 11400 GTTCTAGCAG TGAAGAGATA AAGAAAAAA AGTACAACCA AATGCCAGTC AGGCACAGCA 11460 GAAACCTACA ACTCATGGAA GGTAAAGAAC CTGCAACTGG AGCCAAGAAG AGTAACAAGC 11520 11580 CAAATGAACA GACAAGTAAA AGACATGACA GCGATACTTT CCCAGAGCTG AAGTTAACAA 11640 ATGCACCTGG TTCTTTTACT AAGTGTTCAA ATACCAGTGA ACTTAAAGAA TTTGTCAATC 11700 CTAGCCTTCC AAGAGAAGAA AAAGAAGAGA AACTAGAAAC AGTTAAAGTG TCTAATAATG 11760 CTGAAGACCC CAAAGATCTC ATGTTAAGTG GAGAAAGGGT TTTGCAAACT GAAAGATCTG 11820 TAGAGAGTAG CAGTATTTCA TTGGTACCTG GTACTGATTA TGGCACTCAG GAAAGTATCT CGTTACTGGA AGTTAGCACT CTAGGGAAGG CAAAAACAGA ACCAAATAAA TGTGTGAGTC 11880 11940 AGTGTGCAGC ATTTGAAAAC CCCAAGGGAC TAATTCATGG TTGTTCCAAA GATAATAGAA 12000

ATGACACAGA AGGCTTTAAG TATCCATTGG GACATGAAGT TAACCACAGT CGGGAAACAA

-142-

GCATAGAAAT GGAAGAAAGT GAACTTGATG CTCAGTATTT GCAGAATACA TTCAAGGTTT 12060 CAAAGCGCCA GTCATTTGCT CCGTTTTCAA 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 AAAACTAAAT GTAAGAAAAA TCTGCTAGAG GAAAACTTTG 12480 AGGAACATTC AATGTCACCT GAAAGAGAAAA TGGGAAATGA GAACATTCCA AGTACAGTGA 12540 GCACAATTAG CCGTAATAAC ATTAGAGAAA ATGTTTTTAA AGAAGCCAGC TCAAGCAATA 12600 TTAATGAAGT AGGTTCCAGT ACTAATGAAG TGGGCTCCAG TATTAATGAA ATAGGTTCCA 12660 GTGATGAAAA CATTCAAGCA GAACTAGGTA GAAACAGAGG GCCAAAATTG AATGCTATGC 12720 TTAGATTAGG GGTTTTGCAA CCTGAGGTCT ATAAACAAAG TCTTCCTGGA AGTAATTGTA 12780 AGCATCCTGA AATAAAAAAG CAAGAATATG AAGAAGTAGT TCAGACTGTT AATACAGATT 12840 TCTCTCCATA TCTGATTTCA GATAACTTAG AACAGCCTAT GGGAAGTAGT CATGCATCTC 12900 AGGTTTGTTC TGAGACACCT GATGACCTGT TAGATGATGG TGAAATAAAG GAAGATACTA 12960 GTTTTGCTGA AAATGACATT AAGGAAAGTT CTGCTGTTTT TAGCAAAAGC GTCCAGAAAG 13020 GAGAGCTTAG CAGGAGTCCT AGCCCTTTCA CCCATACACA TTTGGCTCAG GGTTACCGAA 13080 GAGGGGCCAA GAAATTAGAG TCCTCAGAAG AGAACTTATC 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 GCTTGTTTTC TTCACAGTGC AGTGAATTGG 13380 AAGACTTGAC TGCAAATACA AACACCCAGG ATCCTTTCTT GATTGGTTCT TCCAAACAAA 13440 TGAGGCATCA GTCTGAAAGC CAGGGAGTTG GTCTGAGTGA CAAGGAATTG GTTTCAGATG 13500 ATGAAGAAAG AGGAACGGGC TTGGAAGAAA ATAATCAAGA AGAGCAAAGC ATGGATTCAA 13560 13620 ACTTAGGTAT TGGAACCAGG TTTGTGTGTG TGGCCCAGTC TATTTATAGA AGTGAGCTAA ATGTTTATGC TTTTGGGGGAG CACATTTTAC AAATTTCCAA GTATAGTTAA AGGAACTGCT 13680

-143-

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 VVVVVVVVG GTGTCCCAAA 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 VVVVVVVV 14220 VNGNGAATGT AATCCTAATA TTTCNCNCCN ACTTAAAAGA ATACCACTCC AANGGCATCN 14280 CAATACATCA ATCAATTGGG GAATTGGGAT TTTCCCTCNC TAACATCANT GGAATAATTT 14340 CATGGCATTA ATTGCATGAA TGTGGTTAGA TTAAAAGGTG TTCATGCTAG AACTTGTAGT 14400 TCCATACTAG GTATTTCAAT CNTTGTTCCT CAAANAATTT GTATGATATA TTNTCATTTA 14460 ATGGAAAGCT TCTCAAAGTA TTTCATTTC 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 14820 GGACCTAGCT CCAACATTTT ATGATCCTTG CTCAGCACAT GGGTAATTAT GGAGCCTTGG TTCTTGTCCC TGCTCACAAC TAATATACCA GTCAGAGGGA CCCAAGGCAG TCATTCATGT 14880 TGTCATCTGA GAVVVVVVV VVVVVAAATT CTGCCTGATA TACTTGTTNC AAAACCAATT 14940 15000 TGTGTATCAT AGATTGATGC TTTTGAANAA AATCAGTATT CTAACCTGAA TTATCACTAT 15060 CAGAACAAAG CAGTAAAGTA GATTTGTTTT CTCATTCCAT TTAAAGCAGT ATTAACTTCA CAGAAAAGTA GTGAATACCC TATAAGCCAG AATCCAGAAG GCCTTTCTGC TGACAAGTTT 15120 15180 GAGGTGTCTG CAGATAGTTC TACCAGTAAA AATAAAGAAC CAGGAGTGGA AGTAAGAAAC 15240 ATCAATGTAA AGATGCTGTG GTATCTGACA TCTTTATTTA TATTGAACTC TGATTGTTAA 15300 TTTTTTTCAC CATACTTTCT CCAGTTTTTT TGCATACAGG CATTTATACA CTTTTATTGC TCTAGGATAC TTCTTTGTT TAATCCTATA TAGGVVVVVV VVVVVVGGA TAAGNTCAAG 15360

-144-

AGATATTTTG ATAGGTGATG CAGTGATNAA TTGNGAAAAT TTNCTGCCTG CTTTTAATCT 15420 15480 15540 TCCTCCCTTC CTTTTCTTTT CTTTCCTTTCC TTTCCTTTTC TTTCCTTTCT TTCCTTTCCT 15600 TTCTTTCTTG ACAGAGTCTT GCTCTGTCAC TCAGGCTGGA GTGCAGTGGC GTGATCTCGN 15660 CTCACTGCAA CCTCTGTCTC CCAGGTTCAA GCAATTTTCC TGCCTCAGCC TCCCGAGTAG 15720 CTGAGATTAC AGGCGCCAGC CACCACACCC AGCTACTGAC CTGCTTTTVV VAAACAGCTG 15780 GGAGATATGG TGCCTCAGAC CATCCCCATG TTATATGTCA ACCCTGACAT ATTGGCAGGC 15840 AACATGAATC CAGACTTCTA GGCTGTCATG CGGGCTCTTT CTTTGCCAGT CATTNCTGAT 15900 CTCTCTGACA TGAACTGTNT CAGGTATGCT TTGGCTGCCC AGCAAGTATG AAATGTCCTN 15960 TCACAAGGGG TGGCGATGGT TTTCTCCTTC CATTTATCTT TCTAGGTCAT CCCCTTCTAA 16020 ATGCCCATCA TTAGATGATA GGTGGTACAT GCACAGTTGC TCTGGGAGTC TTCAGAATAG 16080 AAACTACCCA TCTCAAGAGG AGCTCATTAA GGTTGTTGAT GTGGAGGAGC AACAGCTGGA 16140 AGAGTCTGGG CCACACGATT TGACGGAAAC ATCTTACTTG CCAAGGCAAG ATCTAGGTAA 16200 TATTTCATCT GCTGTATTGG ARCAMACACT YTGATTTTAC TCTGAATCCT ACATAAAGAT 16260 ATTCTGGTTA RCCAACTTTT AGATGTRCTA GTCTATCATG GACACTTTTG TTATACTTAA 16320 TTAAGCCCAC TTTAGAAAAA 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 CGGAGGGGGNG GGTGGGGGGNG AGCAGGNCAA AACACNGACC CAGCTGGGGT 16740 GGAAGGGAAG CCCACTCNAA AAAANNTTNV VVVVVVVV VVAGGGAACC CCTTACCTGG 16800 AATCTGGAAT CAGCCTCTTC TCTGATGACC CTGAATCTGA TCCTTCTGAA GACAGAGCCC 16860 CAGAGTCAGC TCGTGTTGGC AACATACCAT CTTCAACCTC TGCATTGAAA GTTCCCCAAT 16920 TGAAAGTTGC AGAATCTGCC CAGAGTCCAG CTGCTGCTCA TACTACTGAT ACTGCTGGGT 16980 ATAATGCAAT GGAAGAAAGT GTGAGCAGGG AGAAGCCAGA ATTGACAGCT TCAACAGAAA 17040

-145-

GGGTCAACAA AAGAATGTCC ATGGTGGTGT CTGGCCTGAC CCCAGAAGAA TTTGTGAGTG 17100 TATCCATATG TATCTCCCTA ATGACTAAGA CTTAACAACA TTCTGGAAAG AGTTTTATGT 17160 AGGTATTGTC AATTAATAAC CTAGAGGAAG AAATCTAGAA ANCAATCACA GTTCTGTGTA 17220 ATTTAATTTC GATTACTAAT TTCTGAAAAT TTAGAAYVVV VVVVVVVV NCCCNNCCCC 17280 CNAATCTGAA ATGGGGGGTAA CCCCCCCCA ACCGANACNT GGGTNGCNTA GAGANTTTAA 17340 TGGCCCNTTC TGAGGNACAN AAGCTTAAGC CAGGNGACGT GGANCNATGN GTTGTTTNTT 17400 GTTTGGTTAC CTCCAGCCTG GTGACAGANC AAGCTCTGTC TNGNAAAANT AGGNGAGGGN 17460 NAAGATTCAC NNTAAATAGT TCCAGGACAC GTGTAGAACG TGCAGGATTG CTACGTAGGT 17520 AAACATATGC CATGGTGGGA TAACTAGTAT TCTGAGCTGT GTGCTAGAGG TAACTCATGA 17580 TAATGGAATA TTTGATTTAA TTTCAGATGC TCGTGTACAA GTTTGCCAGA AAACACCACA 17640 TCACTTTAAC TAATCTAATT ACTGAAGAGA CTACTCATGT TGTTATGAAA ACAGATATAC 17700 17760 CCAAGAACCT TTACAGAATA CCTTKCATCT SCTGCATAAA ACCACATGAG GCGAGGCACA GTGGCGCATG CCTGTAWTCG CARCACTTTG GGAGGCCGAG GCGGGCAGAT ACCCCGAGRT 17820 TMSGAGTCGA GCCATCCTGG CCAGCATGGT GAAACCCCGT CTCTACTANN AAATGGNAAA 17880 ATTANCTGGG TGTGGTCGCG TGCNCCTGTA GTCCCAGCTA CTCGTGAGGC TGAGGCAGGA 17940 GAATCACTTG AACCGGGGGAA ATGGAGGTTT CAGTGAGCAG AGATCATNCC CCTNCATTCC 18000 AGCCTGGCGA CAGAGCAAGG CTCCGTCNCC NAAAAAATAA AAAAAAACGT GAACAAATAA 18060 GAATATTTGT TGAGCATAGC ATGGATGATA GTCTTCTAAT AGTCAATCAA TTACTTTATG 18120 AAAGACAAAT AATAGTTTTG CTGCTTCCTT ACCTCCTTTT GTTTTGGGTT AAGATTTGGA 18180 18240 GTGTGGGCCA GGCACVVVV VVVVVVGA TCTATAGCTA GCCTTGGCGT CTAGAAGATG GGTGTTGAGA AGAGGGAGTG GAAAGATATT TCCTCTGGTC TTAACTTCAT ATCAGCCTCC 18300 CCTAGACTTC CAAATATCCA TACCTGCTGG TTATAATTAG TGGTGTTTTC AGCCTCTGAT 18360 TCTGTCACCA GGGGTTTTAG AATCATAAAT CCAGATTGAT CTTGGGAGTG TAAAAAACTG 18420 AGGCTCTTTA GCTTCTTAGG ACAGCACTTC CTGATTTTGT TTTCAACTTC TAATCCTTTG 18480 18540 AGTGTKTCTC ATTCTGCAGA TGCTGAGTTT GTGTGTGAAC GGACACTGAA ATATTTTCTA GGAATTGCGG GAGGAAAATG GGTAGTTAGC TATTTCTGGK AAGKATAATA CTATWTCTCC 18600 CCYCCTCCCC TAACAMCTCA GAATTGSATT TTACACCTAA CATTTACAMC TAAGNYYYTK 18660 GCTGATGCTT GAGTCTGAGT CMCMAAAGKC CTTTAATTGT AATACTAACT ACTTTTATCT 18720

-146-

18780 TTAATATCAC TTTGTTCAGA TAAGCTGGTG ATGCTGGGAA AATGGTCTCT TTTATAACTA ATAGGACCTA ATCTGCTCCT AGCAATGTTA GCATATGAGC TAGGGATTTA TTTAATAGTC 18840 GGCAGGAATC CATGTGCARC AGNCAAACTT ATAATGTTTA AATTAAACAT CAACTCTGTC 18900 TCCAGAAGGA AACTGCTGCT ACAAGCCTTA TTAAAGGGCT GTGGCTTTAG AGGGAAGGAC 18960 CTCTCCTCTG TCATTCTTCC TGTGCTCTTT TGTGAATCGC TGACCTCTCT ATCTCCGTGA 19020 AAAGAGCACG TTCTTCTGCT GTATGTAACC TGTCTTTTCT ATGATCTCTV VVVVVVVVV 19080 VVNAAAAACG GGGNNGGGAN TGGGCCTTAA ANCCAAAGGG CNAACTCCCC AACCATTNAA 19140 AAANTGACNG GGGATTATTA AAANCGGCGG GAAACATTTC ACNGCCCAAC TAATATTGTT 19200 AAATTAAAAC CACCACCNCT GCNCCAAGGA GGGACCTTCT CCNNCAAANC TTATTAAAGG 19260 GCCTGNGCCT TAAAGAGGGA AGGACCCTCT CCTCTGTCAA TCTTCCTGTG CTCTNTTTGT 19320 GAATCGCTGA CCTCTCTATC TCCGTGAAAA AGAGCACGTT CTTCTGCTGT ATGTAACCTG 19380 TCTTTTCTAT GATCTCTTTA GGGGTGACCC AGTCTATTAA AGAAAGAAAA ATGCTGAATG 19440 19500 AGGTAAGTAC TTGATGTTAC AAACTAACCA GAGATATTCA TTCAGTCATA TAGTTAAAAA 19560 TGTATTTGCT TCCTTCCATC AATGCACCAC TTTCCTTAAC AATGCACAAA TTTTCCATGA TAATGAGGAT CATCAAGAAT TATGCAGGCC TGCACTGTGG CTCATACCTA TAATCCCAGC 19620 GCTTTGGGAG GCTGAGGCGC TTGGATCVVV VVVVVVVV TGCACCCCAA GACGGATGGG 19680 NNCTCCCCCN CCNCTATTTT GATTTAGAAA NTGGGTTCCC ATGNGTAAGC NGTGCGACCC 19740 CTGCCTCAAT GACTNCTCCT AATCTCCCAA ATTCTAGGAT AGNGGGTGAN CCCTCCNCTG 19800 19860 NCTGATCCNT AAAATGGACG TGTCTGCTNC NCCACTCCAT TGAAGGAGCT TCTCTTTCTC 19920 TTATCCTGAT GGGTGTGNTG TTTCTTTCAG CATGATTTTG AAGTCAGAGG AGATGTGGTC 19980 AATGGAAGAA ACCACCAAGG TCCAAAGCGA GCAAGAGAAT CCCAGGACAG AAAGGTAAAG 20040 CTCCCTCCCT CAAGTTGACN NAAATCTCAC CCCACCACTC TGTATTCCAC TCCCCTTTGC AGAGATGGGC CGCTTCATTT TGTAAGACTT ATTACATACA TACACAGTGC TAGATACTTT 20100 20160 CACACAGGTT CTTTTTTCAC TCTTCCATCC CAACCACATA AATAAGTATT GTCTCTACTT 20220 TATGAATGAT AAAACTAAGA GATTTAGAGA GGCTGTGTAA TTTGGATTCC CGTCTCGGGT 20280 TCAGATCVVV VVVVVVV ATCTTCAGGG GGCTAGAAAT CTGTTGCTAT GGGCCCTTCA 20340 CCAACATGCC CACAGGTAAG AGCCTCGGGA GAANCCAGAG TTCCAGCACC AGCCTTTGTC 20400 TTACATAGTG GAGTATTATA AGCAAGGTCC CACGATGGGG GTTCCTCAGA TTGCTGAAAT

-147-

GTTCTAGAGG CTATTCTATT TCTCTACCAC TCTCCAAACA AAACAGCACC TAAATGTTAT 20460 CCTATGGCAA AAAAAAACTA TACCTTGTCC CCCTTCTCAA GAGCATGAAG GTGGTTAATA 20520 GTTAGGATTC AGTATGTTAT GTGTTCAGAT GGCGTTGAGC TGCTGTTAGT GCCVVVVVV 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 TCACCCTTGG CACAGVVVVV VVVVVVCA 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 VVVVVVVV VCAGTAATCC TNAGAACTCA 21300 TACGACCGGG CCCCTGGAGT CGNTGNTTNG AGCCTAGTCC NGGAGAATGA ATTGACACTA 21360 21420 ATCTCTGCTT GTGTTCTCTG TCTCCAGCAA TTGGGCAGAT GTGTGAGGCA CCTGTGGTGA CCCGAGAGTG GGTGTTGGAC AGTGTAGCAC TCTACCAGTG CCAGGAGCTG GACACCTACC 21480 TGATACCCCA GATCCCCCAC AGCCACTACT GACTGCAGCC AGCCACAGGT ACAGAGCCAC 21540 AGGACCCCAA GAATGAGCTT ACAAAGTGGC CTTTCCAGGC CCTGGGAGCT CCTCTCACTC 21600 TTCAGTCCTT CTACTGTCCT 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

-148-

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

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

-149-. ...

#### GATAAATTAA AACTGCGACT GCGCGGCGTG

(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
- (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

### AAAAATAAAG GACGCTGTCG AYTAGTTCTT

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

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

#### AAGATATATA TATATGTTTT TCTAATGTGT

30

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

GTAAGTCAGC ACAAGAGTGT ATTAATTTGG

(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

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

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

TTATTGAGCC TCATTTATTT TCTTTTTCTC

(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

(vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens 30

30

,

-150-

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

-151-

## ATGTGGCTCC ATTATTAGCT TTTGTTTTTG

- (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
    - (D) IOPOLOGI: LINear
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (iii) HYPOTHETICAL: NO
  - (vi) ORIGINAL SOURCE:
     (A) ORGANISM: Homo sapiens

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

AGCTACTTTT TTTTTTTTTTTTTTTTTTTTTTGAGACAG

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

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

GTAAGTGCAC ACCACCATAT CCAGCTAAAT

- (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

30

. . . . . .

30

(ii) MOLECULE TYPE: DNA (genomic)

-152-. ....

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

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

#### TTGTTCTTTC TTTCTTTATA ATTTATAGAT

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

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

GATATAANNN GAANTGATGC TAGGTTGGAA

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

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

GATAATCACT TGCTGAGTGT GTTTCTCAAA

30

-153-

(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

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

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

GTAAGTGTTG AATATCCCAA GAATGAGAGT

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

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

AAACATAATG TTTTCCCTTG TATTTTACAG

(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
  - (D) IOFOLOGI: IIIIeal

(ii) MOLECULE TYPE: DNA (genomic)

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

30

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

-154-

### AAACCATTTG TTTTCTTCTT CTTCTTCTTC

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

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

#### TTGACTGTTC TTTACCAWAC TGTTWARCAG

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

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

#### GGTAAGGGTC TCAGGTTTTT TAAGTATTTA

(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

•

30



-155-. ---

(ii) MOLECULE TYPE: DNA (genomic)

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

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

TTGATTTATT TTTTGGGGGG AAATTTTTTA

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

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

GGTGAGTCAA AGAGAACCTT TGTCTATGAA

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

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

TTCTTATTAG GACTCTGTCT TTTCCCTATA

30

30

30

. -



(2) INFORMATION FOR SEQ ID NO:34:

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

-156-

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

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

GGTAATGGCA AAGTTTGCCA ACTTAACAGG

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

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGAGTACCTT GTTATTTTTG TATATTTTCA

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

GeneDX 1010, pg. 161

30

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

-157-

GTATTGGAAC CAGGTTTGTG TGTGTGGCCC

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

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

ACATCTGAAC CTCTGTTTTT GTTATTTAAG

(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

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

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

NGNGAATGTA ATCCTAATAT TTCNCNCCNA

(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
  - D/ IVFOLUGI: IIIICUL

30

30

(ii) MOLECULE TYPE: DNA (genomic)

-158-. --

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

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39: TCTTGGTACC ATNNATCGTT TTTGAATNAG

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

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

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

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

AGATTTGTTT TCTCATTCCA TTTAAAGCAG

30

30

(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
- (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

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

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

CGATGGTTTT CTCCTTCCAT TTATCTTTCT

- (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

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44: GTAATATTTC ATCTGCTGTA TTGGARCAMA

-160-

- (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
    - (D) IOFOLOGI: IIIIeal
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (iii) HYPOTHETICAL: NO
  - (vi) ORIGINAL SOURCE:
     (A) ORGANISM: Homo sapiens

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

TGGAAGGGAA GCCCACTCNA AAAAANNTTN

- (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
    - D) TOFODOGI: TIMBAL
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (iii) HYPOTHETICAL: NO
  - (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens

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

GTGAGTGTAT CCATATGTAT CTCCCTAATG

(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

30

30

. -

-161-

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

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

ATGATAATGG AATATTTGAT TTAATTTCAG

(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
- (vi) ORIGINAL SOURCE:
   (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TATACCCAAG AACCTTTACA GAATACCTTK

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

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

TAATCCTTTG AGTGTKTCTC ATTCTGCAGA

(2) INFORMATION FOR SEQ ID NO:50:

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

-162-. ...

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

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

KAAGKATAAT ACTATWTCTC CCCYCCTCCC

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

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

TAACCTGTCT TTTCTATGAT CTCTTTAGGG

(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

(vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens 30

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

-163-نيە ر

GTAAGTACTT GATGTTACAA ACTAACCAGA

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

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

TTATCCTGAT GGGTGTGNTG TTTCTTTCAG

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

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

GTAAAGCTCC CTCCCTCAAG TTGACNNAAA

(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



30

30



(ii) MOLECULE TYPE: DNA (genomic)

-164-

- (iii) HYPOTHETICAL: NO
  - (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens.

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

TAATTTGGAT TCCCGTCTCG GGTTCAGATC

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

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

GTAAGAGCCT CGGGAGAANC CAGAGTTCCA

(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

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

GTGATTTTAG ATCCTAAATG TCCATTTTAG

30

30



(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid

-165-

- (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
   (A) ORGANISM: Homo sapiens
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CAGAGCAAGA CCCTGTCTCA AAAACAAACA

(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
  - (D) IOFOLOGI. IIneat
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE: (A) ORGANISM: Homo sapiens

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

TTGAATGCTC TTTCCTTCCT GGGGANCCAG

- (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

30



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

-166-

TTTATCACTC ATTACCTGGT GCTTGAGTAG

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

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

CTAATCTCTG CTTGTGTTCT CTGTCTCCAG

- (2) INFORMATION FOR SEQ ID NO:62:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 42 amino acids(B) TYPE: amino acid
    - (C) STRANDEDNESS:
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (iii) HYPOTHETICAL: NO

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

Cys 1	Pro	Ile	Cys	Leu 5	Glu	Leu	Ile	Lys	Glu 10	Pro	Val	Ser	Thr	Lys 15	Cys
Asp	His	Ile	Phe 20	Cys	Lys	Phe	Cys	Met 25	Leu	Lys	Leu	Leu	Asn 30	Gln	Lys
Lys	Gly	Pro 35	Ser	Gln	Суз	Pro	Leu 40	Суз	Lys						

30

-167-

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

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

CysPro Ile CysLeu Glu Leu Leu LysGlu Pro Val Ser Ala AspCys151015Asn His Ser Phe Cys Arg Ala CysIle Thr Leu Asn Tyr Glu Ser Asn202530

Arg Asn Thr Asp Gly Lys Gly Asn Cys Pro Val Cys Arg354045

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

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

Cys Pro Ile Cys Leu Asp Met Leu Lys Asn Thr Met Thr Thr Lys Glu 1 5 10 15 Cys Leu His Arg Phe Cys Ser Asp Cys Ile Val Thr Ala Leu Arg Ser 20 25 30

Gly Asn Lys Glu Cys Pro Thr Cys Arg 35 40

.

(i) SEQUENCE CHARACTERISTICS:

(2) INFORMATION FOR SEQ ID NO:65:

- (A) LENGTH: 42 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO

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

Cys Pro Val Cys Leu Gln Tyr Phe Ala Glu Pro Met Met Leu Asp Cys 1 5 10 15 Gly His Asn Ile Cys Cys Ala Cys Leu Ala Arg Cys Trp Gly Thr Ala 20 25 30

-168-

.

Cys Thr Asn Val Ser Cys Pro Gln Cys Arg 35 40

# WHAT IS CLAIMED IS:

 An isolated DNA sequence consisting essentially of DNA coding for a mutant BRCA1 polypeptide, said DNA selected from the group consisting of the mutations set forth in Table 14.

-169-

- A nucleic acid probe complementary to human altered BRCA1 gene sequences, said altered BRCA1 gene sequences selected from the group consisting of the mutations set forth in Table 14.
- 3. 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 and selected from the group consisting of the mutations of claim 1.

4. The method of claim 3 wherein said expression product is mRNA of BRCA1.

- 5. The method of claim 4 wherein the alteration of the wild-type BRCA1 gene mRNA is detected by hybridization of the mRNA from said human sample to a BRCA1 gene probe.
- 6. The method of claim 3 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.
- 7. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected byhybridization of a BRCA1 gene probe to genomic DNA isolated from said sample.
  - 8. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by hybridization with an allele-specific probe.

30

5

10

15

- The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by amplifying all or part of the BRCA1 gene in said sample to produce an amplified sequence and sequencing the amplified sequence.

-170-

5

9.

10.

The method of claim 3 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.

11. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by molecularly cloning all or part of the BRCA1 gene in said sample to produce a cloned sequence and sequencing the cloned sequence.

12. The method of claim 3 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.

14. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by amplification of BRCA1 gene sequences in said sample and hybridization of the amplified sequences to nucleic acid probes which comprise non-wild-type BRCA1 gene sequences.

25 15. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by screening for a deletion mutation.

16. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by screening for a point mutation.

15

10

20

5

17. The method of claim 3 wherein the alteration of the wild-type BRCA1 gene is detected by screening for an insertion.

-171-

18. The method of claim 3 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.

# TITLE OF THE INVENTION

17q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE MUTATIONS

-172-

## ABSTRACT OF THE DISCLOSURE

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. Additionally, the invention relates to somatic mutations in the BRCA1 gene in ovarian cancer. Additionally, the invention relates to somatic mutations in the BRCA1 gene in other human cancers and their use in the diagnosis 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 the BRCA1 gene for mutations, which are useful for diagnosing the predisposition to breast and ovarian cancer.

10

5



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

FIGURE 1

8



FIGURE 2

GeneDX 1010, pg. 179

Schematic Map of STS', P1s, and Bac's in the Region


## 08 409305

;

•





:

FIGURE 5

.

62 63 64 65	BRCA1 RPT1 RIN1 RFP1 C3HC4 motif	CPICLELIKEPVSTK-CDHIFCKFCMLKLLNQKKGPSQCPLCK CPICLELLKEPVSAD-CNHSFCRACITLNYESNRNTDGKGNCPVCR CPICLDMLKNTMTTKECLHRFCSDCIVTALRSGNKECPTCR CPVCLQYFAEPMMLD-CGHNICCACLARCWGTACTNVSCPQCR CCC-HCCCCCC
----------------------	----------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

.*

. . .

SEQ ID NO:

٠.

1

8;

409305

•.•



FIGURE 6

18 409305

GeneDX 1010, pg. 183





.

# **# 409305**

÷



FIGURE 9A

۰.

3

¹⁸ 409305





FIGURE 9B

Ì

2

08 409305

•



PM1	4	5	6	7	8	9	
							A
					2 • .	Ser.	G
		1. Ca				•	A
			-		-		G

PM A							A
:: ::				M73			G
織	a 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1. A.					A
4				H.	6		G

D11.0-7

FIGURE 9C

I

			•		•		
					•	<b>10 m</b>	
						no en	<b>A</b> • •
	•					08 40	930c
				• ••			000
1	ccttttacgt	categgggag	actgggtggc	aatcgcagcc	cgagagacgc	atggctcttt	
						gcttagcggt	•
		tttccgtggc	-				
		gtgAGCTCGC					
		GCCCCTGCGC					
		cgctgtcgay					
		cttyattta					
						AAGAAATGGA	
		CTTCGCGTTG					
		ATCTGgtaag					
		aatgaacaga					
		attgtaattg					
		tetecateca	·				
		agcagggete					
						tagtatagag	
		atcaatataa					
						attgaacgaa	
		tatgttgact					
		ytaaataaat					
		GAACCTGTCT					
		gettttgttt					
		ttcaattaag					
		tatcaaaagt					
		ttttttgag					
		tcactgcacc					
		acatggtcaa					
1561	aacagetact	tttttttt	ttttttgaga	cagGGkCTTA	CTCTGTTGTC	CCAGCTGGAG	
1621	TACAGW-G-T	GCGATCATGA	GGCTTACTGT	TGC-TTGAC-	TCCTA-GGCT	CAAGCGATCC	
		GTCTC-CAAG					
		ttctgtagag					
		cgtctgccca					
		ccagtatttt					
		gttaacaacc					
		agtaagccat					
2041	tttaggtttt	tgettatgea	acycacycaa	acaattagga	aactattoct	totaattcac	
2101	ctoccattac	tttttaaatg	gcattaaaa	acaattataa	gattatett	tcatggctat	
						tctttcttta	
						AGGGCCTTCA	
						gatgctaggt.	•
						cctataaaac	
						tccaaagtta	
		agttttatgg					
2501	cectessa	ageceacyg	geatecgeet	tacacaggen	cacacoutta	tatacagatg	
2021 9201	gguudaaug	attgatttag	agecaleadd	ctattactac	atcttattt	tattotta	
6301 9611	-caacyacyc	cttattttag	tatattat	addttastas	tcacttooto	agtgtgttttc	
2041	Calgictet	GACCOTACAA	CANAGE ACA	CATTRACTCA	ACTTGTTGAA	GAGCTATTGA	
3761	A A A MOA MOMO	TGCTTTTCAG	OPTOR CRORE	CTTTCCARCINA	TGCAAACAGC	TATAATTTTG	
2021	CANALATIC	TOCTTTTCAG	CITCACACAG	TAAAACATCA	ACITTTCTATC	ATCCAAAGTA	
2041	TACCOMPACE	AAATAACTCT	AAAAAAA	TACACACITCA	ACCCGABABAT	CCTTCCTTGC	
2001	ACasagoott	tgttttcttc	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	tattatta	++++++++++	cttttttt	
2001 4741	toutasses	ggagtcyttg			tannaannan	ctectectec	
2001	Lewegagage	ggagtcyttg	gecegessyg	ycyccgagge	Layyaaycay		

÷

Figure 10A

GeneDX 1010, pg. 189

GeneDX 10	10. r	oa. 1	190
-----------	-------	-------	-----

08 409305

:

	3061	actestacae	cacgctagta	aataacataa	acakacacaa	cacacacaco	atoocaooo	
			attgttvvvv					
			tagtagettg					
			gtcaggagtt					
			aaattagctg					
			aagaattgct					
			ccagcetggg					
			ctgggcgcag					•
			tcgagtccta					
			atacaaaaat					
			aggtgggagg					
			cacgetecag					
			gtagttetca					
			aaaaaagtaa					
			ggggttttat					
			atggctgtaa					
			gtgagaccag					
			aggcgtggtg					
			ctgaaccagg					
			caacagagca					
			ctgggattac					
•			atatatgtga					
			tatatatgta					
			tgtgtgtaca					
			gctactcagg					
			gctgagattg					
	4621	tacacccccc	aaaacaaaaa	aaaaagttat	cagatgtgat	tggaatgtat	atcaagtatc	
			tatgctatat					
			aagacaacms					
			cwattatttt					
			ccycaarcar					
			ctcccagccc					
			wayayvtwds					
	5041	ggtgtcaakt	ytctcttcag	gaggaaaagc	acamgaactg	gccaayaatt	gettgactgt	
	5101	tetttaccaw	actgttwarc	agGAAACCAG	TCTCAGTGTC	C-AACTCTCT	AACCTTGGAA	
	5161	CT-GTGAGAA	CTCTGAGGAC	AAAGCAGCGG	ATACAACCTC	AAAAGACGTC	TGTCTACATT	
	5221	GAATIGggta	agggtctcag	gttttttaag	tatttaataa	taattgctgg	attecttate	
			gccamaaatc					
							tcccaccttg	
			ctttaaaaag					
			tgggaggcca					
			tggagaaact					
			gtaateccag					
			ttgcggtgaa					
	5701	cgaaactcca	tctaaaaaaa	aaaaagagag	caaaagaaag	amemeetgge	terestracta	
	5761	tgtaaatatg	ttttggaaa	gatggagagt	agcaataaga	aaaaacatga	Lygarlycia	
	5821	cagtattag	ttccaagata	aattgtacta	gatgaggaag	CELEELaaga	ayayuuyaau	
	5881	tgccaggcgc	agtggctcac	gcctgtaate	ccagcacttt	gggaggeega	Saraaa Saraa	
	5741	ccacctgagg	tcgggagttc	aagaccagcc	cgaccaacat	yyayaaaccc	cractacte	
	5001	aaaaaaaaaa	aaaaaaaaat	tageeggggt	ggtggettat	georgeate	traccaada	
	0001	ayyayyctga	ggcaggagaa	regettgaad	ccaygaagca	Jaggrideag	-343004434	

Figure 10B



.



ConoDV	1010	na	101
GeneDX	1010.	pų.	191

18 409305

:

6121	+		-		. Is a mainte ante anaiste		
		gcactccagc					
		atcttggctg					•
		aaggattggt					
		ctattttaa					
		tactttggga					
		aacctggtga gtagcacatg					
		agggaggcgg					
		agtgagactg					
		taaaaagaaa					
		gattgettga					
		caaaaaaatt					
		ctcaggaggc					
		tgatcatgcc					
		aagagetget					
		atttataatt					
		taacacttac					
		atagttatag					
		tettgtataa					
		tacagatgag					
		gtaagtggct					
		tttgtttttg					
		ttgaacteet					
		gtgagagcca					
		agettaacta					
		ctgtttttca					
		tttttgggg					
		TTATTGCAgg					
		aatattaagg					
		gggatgtatg					
		tttttcatc					
		tttctccctc					
		tggtattgga					
8101	tagtggagga	aataccettt	ttgtttgttg	gatttagtta	ttagcactgt	ccaaaggaat	
8161	ttaggataac	agtagaactc	tgcacatgct	tgcttctagc	agattgttct	ctaagttcct	
8221	catatacagt	aatattgaca	cagcagtaat	tgtgactgat	gaaaatgttc	aaggacttca	
8281	ttttcaactc	tttctttcct	ctgttcctta	tttccacata	tetetcaage	tttgtctgta	
8341	tgttatataa	taaactacaa	gcaaccccaa	ctatgttacc	taccttcctt	aggaattatt	
8401	gcttgaccca	ggttttttt	ttttttt	tggagacggg	gtcttgccct	gttgccagga.	
8461	tggagtgtag	tggcgccatc	tcggctcact	gcaatctcca	actccctggt	tcaagcgatt	
8521	ctcctgtctc	aatctcacga	gtagctggga	ctacaggtat	acaccaccac	gcccggttaa	
8581	ttgaccattc	catttcttc	tttctctctt	ttttttt.	tttttgaga	cagagtcttg	
8641	ctctgttgcc	caggetggag	tacagaggtg	tgatetcacc	tctccgcaac	gtctgcctcc	
8701	caggttgaag	ccatactcct	gcctcagcct	ctctagtagc	tgggactaca	ggcgcgcgcc	
8761	accacacccg	gctaatttt	gtattttag	tagagatggg	gtttcaccat	gttggccagg	
8821	ctggtcttga	actcatgacc	tcaagtggtc	caccegeete	agceteceaa	agtgctggaa	
		gagccaccgt					
		tttcactaac					
		cttggtcatt					
		ttccctataG					
9121	ACCAGGGATG	AAATCAGTTT	GGATTCTGCA	AAAAAGggta	atggcaaagt	ttgccaactt	

Figure 10C

					•••••••	
	C.					00
	The state					@ 409305
	•					C00003
	e e					
9181	aacaggcact	gaaaaqaqaq	tqqqtaqata	caqtactqta	attagattat	tetgaagaee
						ttctctgtca .
		tatggtctga				
		actgattgtt				
9421	ttgagatggg	gteteactet	gttgcccagg	ctggagtgca	gtgatgcaat	cttggctcac
		acctccaaag				
		gcatgccacc				
		accgaggetg				
		gctgcagtta				
						agtggtaagt
		aacagagtat				
9841	aaccagtttt	ctgatggcca	atctgctttt	aattcactct	tagacgttag	agaaataggt
9901	gtggtttctg	catagggaaa	attetgaaat	taavvvvvvv	vvvvvgatc	ctaagtggaa
9961	ataatctagg	taaataggaa	ttaaatgaaa	gagtatgagc	tacatettea	gtatacttgg
10021	tagtttatga	ggttagtttc	tctaatatag	ccagttggtt	gatttccacc	tccaaggtgt
10081	atgaagtatg	tattttta	atgacaattc	agtttktgag	taccttgtta	ttttgtata
10141	ttttcaGCTG	CTTGTGAATT	TTCTGAGACG	GATGTAACAA	ATACTGAACA	TCATCAACCC
		GATTTGAACA				
10261	TCAGGGTAGT	TCTGTTTCAA	ACTTGCATGT	GGAGCCATGT	GGCACAAAT-	ACTCATGCCA
		GCATGAGAAC				
		CTGTAATAAA				
		TAAGGAAACA				
		TGATCCCCTG				
		TAGAGATACT			and the second	
		GTGGTTTTCC				
		ATCAAATGCC				
		TTCTTCAGAG				
		TGAAAGAGTT				
		CCTATCGGAA				
		GAGCATTTGT				
		GTAAAAGGAG				
		CAGTTCAAAA				
		AAGTGATGAA				
		GAATGAGAAA				
		AGCTGAACCT				
		AAAAGCACCT				
		TGAACTAGTA				
and the second		TTGTTCTAGC				
						GGAGCCAAGA
		GCCAAATGAA				
		AAATGCACCT				
		TCCTAGCCTT				
		TGCTGAAGAC				
		TGTAGAGAGT				
		CTCGTTACTG				
		TCAGTGTGCA				
		AAATGACACA AAGCATAGAA				
		AAGCATAGAA TTCAAAGCGC				
		AACATTCTCT				
		TGAACAAAAAG				
7070T	CITTIGAAIG	I GAACAAAAG	GANGANAATC	ANUGAANUGAA	TOWATCIWAT	4.9 W M 4.9 M M M 4.9 M

.

÷.

Figure 10D

					-		
	12241	TACAGACAGT	TAATATCACT	GCAGGCTTTC	CTATGGTTGG	TCAGAAAGAT	AAGCCAGTTG
		ATAATGCCAA					
		GCAACGAAAC					
		TACCACCACT					
		AGGAAAACTT					
		CAAGTACAGT					
		GCTCAAGCAA					
		AAATAGGTTC					
		TGAATGCTAT					
		GAAGTAATTG					
		TTAATACAGA					
		GTCATGCATC					
•		AGGAAGATAC					
		GCGTCCAGAA		the second second second second			
		AGGGTTACCG					
		ATGAAGAGCT					
		AGTCTACTAG					
		TATTATCATT					
		CTCAGGAACA					
		GCAGTGAATT					
		CTTCCAAACA					
		TGGTTTCAGA	•				
		GCATGGATTC					
		tagaagtgag					
		ttaaaggaac					
		aagteettea			-		
		ctggtggact		-			
		ttatcactgg					
		gngnttggct					
		vvvggtgtcc					
		tgccaatgag					
		aacctctgtt					
		AAGCGTCTCT					
		CAGAGVVVVV					
		-tacc-actc					
÷		tenetaacat					
		aaggtgttca					
	14461	caaan-aatt	totatoatat	attntcattt	aatggaaa-g	cttctcaaag	tatttc-att
		tt-cttggta					
		AAGCTCCAGC					
		TTCTAACAGC					
		AGAACA-A-A					
	14761	caaaattctt	tcettceet	ttatctcctt	ctgaagagta	aggacctage	tccaacattt
	14821	tatgateett	geteageaca	toogtaatta	tggagcettg	gttettgtee	ctgctcacaa
	14881	ctaatatacc	agtcagaggg	acccaaggca	gtcattcatg	ttgtcatctg	agavvvvvvv
	14941	vvvvvaaat	tctgcctgat	atacttgttn	caaaaccaat	ttgtgtatca	tagattgatg
	15001	cttttgaana	aaatcagtat	tctaacctga	attatcacta	tcagaacaaa	gcagtaaagt
	15061	agatttgttt	tctcattcca	tttaaagcag	TATTAACTTC	ACAGA-AAAG	TAGTGAATAC
	15121	CCTATAAGCC	AGAATCCAGA	AGGCCTTTCT	GCTGACAAGT	TTGAGGTGTC	TGCAGATAGT
	15181	TCTACCAGTA	аааатааада	ACCAGGAGTG	GAAgtaagaa	acatcaatgt	aaagatgctg
	15241	tggtatctga	catctttatt	tatattgaac	tctgattgtt	aattttttc	accatacttt
		-					

Figure 10E

*le* 409305

:

0 01	40.40		40.4
GeneDX	1010.	pa.	194

*18* 409305

15301	ctccagtttt	tttgcataca	ggcatttata	cacttttatt	getetaggat	acttettte
	tttaatccta					
	tgcagtgatn		-	-		
	tnectcecte		-	-		
	tcentetnte					
	ttetttettt					
	ttgctctgtc					
	gtc-tc-cca				····	
	ggcgccagcc		· · · · · · · · · · · · · · · · · · ·			
	atggtgcctc					
	aatccagact		•	-		
15961	gacatgaact	gtntcaggta	tgetttgget	gcccagcaag	tatgaaatgt	cctntcacaa
	ggggtggcga					
	CATCA-TTA-					
16141	AACTACCCA-	TCTCAAGAGG	AGCT-CATTA	AGGTTGTTGA	TGTGGAGGAG	CAACAGCTGG
16201	AAGAGTCTGG	GCCACACGAT	TTGACGGAAA	CATCTTACTT	GCCAAGGCAA	GATCTAGgta
16261	atatttcatc	tgetgtattg	garcamacac	tytgatttta	ctctgaatcc	tacataaaga
16321	tatt-ctggt	tarccaactt	-ttagat-gt	rctagtctat	-catggacac	ttttgt-tat
16381	acttaattaa	gcccacttta	gaaaaatagc	tcaagtgtta	atcwaggttt	acttgwaaat
	tattgaaact					
16501	gwgggagtck	tggtgtactc	tamatgtatt	atttcaggcc	aggcatagtg	gctcacgcct
16561	ggtaatccca	gtayycmrga	gcccgaggca	ggtggagcca	gctgaggtca	ggagttcaag
16621	acctgtcttg	gccaacatgg	gngaaaccct	gtettettet	taaaaaanac	aaaaaaatt
16681	aactgggttg	tgettaggtg	natgccccgn	atcctagttn	ttcttgnggg	ttgagggagg
16741	agatcacntt	ggaccccgga	ggggngggtg	ggggngagca	ggncaaaaca	cngacccagc
	tggggtggaa					
	ACCTGGAATC					
16921	ACA-GAGCCC	C-A-GAGTCA	-GCTCGTGTT	GGC-AACA-T	ACCATCTTC-	AACCTCTGCA
16981	TTGAAAGTT-	CCCCAATTGA	AA-GTTGCAG	AATCTGCCCA	GAGTCCAGCT	GCTGCTCA-T
	ACTACTGATA					
	TTGACAGCTT					
17161	CCCAGAAGAA	TTTgtgagtg	tatccatatg	tatctcccta	atgactaaga	cttaacaaca
	ttctggaaag					
	ancaatcaca			-		
	*****					
	gggtn-gcnt					
	gng-acgtgg					
	aagetetgte					
	gtgtagaacg					
	tctgagctgt					
	TCGTGTACAA					
	CTACTCATGT					
	sctgcataaa					
	ggaggccgag					
	gaaaccccgt					
	gtcccagcta					
	cagtgagcag					
	naaaaataa					
	gtcttctaat					
	acctcctttt					
18301	tctat-aget	ageettggeg	tctagaagat	gggtgttgag	aagagggagt	ggaaagatat

Figure 10F

18361	tteetetaat	cttaacttca	tatcageete	cectagaett	ccaaatatcc	atacctocto	
		gtggtgtttt					
		tcttgggagt					
		ttttcaactt					
		CGGACACTGA					
		kaagkataat					
		acatttacam		7 .			
		taatactaac					
		aaatggtctc					
		gctagggatt					
		taaattaaac					
	and the second	ctgtggcttt					
		gctgacctct					
		Ctatgatete					
		ggcnaactcc					
		tcacngccca					
		ctccnncaaa	-				
		aatcttcctg					
		ttettetget					
		AAAGAAAGAA					
		cattcagtca					
		acaatgcaca					
		ggetcatace					
		vvtgcacccc					
		ccatgngtaa					
		atagngggtg					
		cact-ccatt					
		AGCATGATTT					
	A 1	GAGCAAGAGA					
		accccaccac					
		ttattacata	-				•
		cccaaccaca					
		gaggetgtgt	<del>-</del> ·				
		GGGGCTAGAA				and the second	
and the second		gagaanccag					
		cccacgatgg					
		actctccaaa					
		ccccttctc					
		atggcgttga					
		taccaagtgg					, 
		cttaaaatcc					
		ctcccattga					
		gttaagtatg					
		GAATGGATGG					
		GGCACAGVVV					
		tgaagtgaca					
		nccagGTGTC					
		CCATGtttat					
		atttgttgaa					
		tgtaaaagta					
		vvvvvvvvvv					

Figure 10G

08 409305

# 



## 08 409305

1.1

21661 21721	AGCTTACAAA	GTGGCCTTTC CTAAATATTT	CAGGCCCTGG TATGTACATC	GAGCTCCTCT AGCCTGAAAA	CACTCTTCAG	TCCTTCTACT	
21721	GTCCTGGCTA	CTAAATATTT	TATGTACATC	AGCCTGAAAA	GGACTTCTGG	CTATGCAAGG	
21661	AGCTTACAAA	GTGGCCTTTC	CAGGCCCTGG	GAGCTCCTCT	CACTCTTCAG	TCCTTCTACT	
21601	CCCACAGCCA	CTACTGACTG	CAGCCAGCCA	CAGGTACAGA	GCCACAGGAC	CCCAAGAATG	
21541	TGGACAGTGT	AGCACTCTAC	CAGTGCCAGG	AGCTGGACAC	CTACCTGATA	CCCCAGATCC	·
21481	ctctgtctcc	agCAATTGGG	CAGATGTGTG	AGGCACCTGT	GGTGACCCGA	GAGTGGGTGT	
21421	gtcgnt-gnt	tngagcctag	tccngg-ag-	aatgaattga	cactaatctc	tgettgtgtt	

Figure 10H



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

re. application of:

Donna M. SHATTUCK-EIDENS et al.

Serial No. (to be assigned)

Filed: herewith

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

STATEMENT PURSUANT TO 37 CFR §1.821(f)

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

Sir:

In the matter of the above-identified application, filed concurrently herewith, Applicants present a computer-readable copy of the Sequence Listing pursuant to 37 CFR §1.821(f). It is hereby certified that these copies are identical.

Respectfully submitted,

Jeff 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. 24884-109347-5 Dated: 24 March 1995 ^{1935.01}

## TEAM 5

## PAGE: 1

->

#### RAW SEQUENCE LISTING PATENT APPLICATION US/08/409,305

DATE: 09/21/95 TIME: 10:38:55

**INPUT SET: S6247.raw** # 3

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

SEQUENCE LISTING 1 2 ENTERFM 3 (1)General Information: 4 5 (i) APPLICANT: Shattuck-Eidens, Donna 6 Simard, Jacques 7 Emi, Mitsuru 8 Nakamura, Yusuke 9 10 (ii) TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer 11 Susceptibility Gene Mutations 12 13 (iii) NUMBER OF SEQUENCES: 65 14 (iv) CORRESPONDENCE ADDRESS: 15 16 (A) ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP 17 (B) STREET: 1201 New York Avenue, N.W., Suite 1000 18 (C) CITY: Washington 19 (D) STATE: DC 20 (E) COUNTRY: USA 21 (F) ZIP: 20005 22 23 (V) COMPUTER READABLE FORM: 24 (A) MEDIUM TYPE: Floppy disk 25 (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS 26 27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 28 29 (vi) CURRENT APPLICATION DATA: (A) APPLICATION NUMBER: US 30 31 (B) FILING DATE: 32 (C) CLASSIFICATION: 33 34 (vii) PRIOR APPLICATION DATA: (A) APPLICATION NUMBER: US 08/348,824 35 36 (B) FILING DATE: 29-NOV-1994 37 38 (vii) PRIOR APPLICATION DATA: 39 (A) APPLICATION NUMBER: US 08/308,104 40 (B) FILING DATE: 16-SEP-1994 41 42 (vii) PRIOR APPLICATION DATA: 43 (A) APPLICATION NUMBER: US 08/300,266 44 (B) FILING DATE: 02-SEP-1994 45 (vii) PRIOR APPLICATION DATA: 46

.

## RAW SEQUENCE LISTING PATENT APPLICATION US/08/409,305

.

DATE: 09/21/95 TIME: 10:38:57

							· ·					Il	<b>NPUT</b>	SET: Se	6247.raw
			APPLIC						/289	,221					
		(B)	FILING	J DAT	E: 1	2-A	UG-1	994							
	(vii)	PRIO	R APPL	ICATI	ON D	АТА	:								
		(A)	APPLIC	CATIO	N NU	MBE	R: U	5 08.	/246	,602					
		(B)	FILING	G DAT	E: 2	0-M	AY-1	994		-					
		1 7 7 0		~~~~											
(	V111)		RNEY/A												
			NAME:												
			REGIS: REFERI							1-10	0247				
		(0)	KEL EKI	SNCE/	DOCK		NOMD	ur:	4400	- I V	7341				
	(ix)	TELE	COMMUN	CATI	ON I	NFO	RMAT	TON:							
	<b>,</b> ,		TELEPH									,			
			TELEF												
(2)	INFO	RMATI	ON FOR	SEQ	ID N	0:1	:								
	(1)		ENCE CH												
			LENGTH TYPE:					rs							
			STRANI												
			TOPOLO				JIE								
		(2)	101 010			ar									
	(ii)	MOLE	CULE TY	PE:	CDNA										
	. ,														
	(iii)	HYPO	rhetic/	AL: N	0										
	(iv)	ANTI-	-SENSE:	NO NO											
	/ ż .	0070		NIDOR											
	(VI)		INAL SO ORGANI				aion	-							
		(A)	ORGANI	SM:	пошо	saj	oreu:	>							
	(ix)	FEAT	URE:												
	(/		NAME/F	(EY:	CDS										
			LOCATI			.57	11								
								_							
	(Xi)	SEQUI	ENCE DE	SCRI	PTIC	N: 5	SEQ :	ID NO	<b>D:l:</b>						
3.00	maaam	<b>a</b>		-	0000	<b></b>	-	1000	naca	amm	1000		T & • ~	Tagaaa	~
AGC	reger	GA GA	CTTCCTC	SG AC	CCCG	CAC	L AGC	actie.	reee	GTT	rere	AGA '	L'AAC'	raaacc	6
CCT	GCGCT		AGGCCTT	C AC	CCTC	TGC	r CTO	GGT	AAAG	TTC	ATTG	GAA	CAGA	AAGAA	11
	00001	0		. •		100						•••••			
АТС	GAT '	TTA TO	CT GCT	CTT	CGC	GTT	GAA	GAA	GTA	CAA	ААТ	GTC	ATT	ААТ	16
			er Ala												
1	-		5		~			10					15		
			AA ATC												21
Ala	Met	-	ys Ile	Leu	Glu	Cys		Ile	Cys	Leu	Glu			Lys	
		:	20				25					30			

GeneDX 1010, pg. 200

## RAW SEQUENCE LISTING PATENT APPLICATION US/08/409,305

DATE: 09/21/95 TIME: 10:39:00

														77	IPUT	SET:	S6247.ra	w
100	GAA	CCT	GTC	TCC	ACA	AAG	TGT	GAC	CAC	АТА	TTT	TGC	AAA					263
101	Glu	Pro	Val	Ser	Thr	Lys	Cys	Asp	His	Ile	Phe	Cys	Lys	Phe	Cys	Met		
102			35			-	-	40				-	45		-			
103																		
104	CTG	AAA	CTT	CTC	AAC	CAG	AAG	AAA	GGG	CCT	TCA	CAG	TGT	CCT	TTA	TGT		311
105	Leu	Lys	Leu	Leu	Asn	Gln	Lys	Lys	Gly	Pro	Ser	Gln	Cys	Pro	Leu	Cys		
106		50				•	55	-	-			60	-			-		
107																		
108`	AAG	AAT	GAT	ATA	ACC	AAA	AGG	AGC	CTA	CAA	GAA	AGT	ACG	AGA	TTT	AGT		359
109	Lys	Asn	Asp	Ile	Thr	Lys	Arg	Ser	Leu	Gln	Glu	Ser	Thr	Arg	Phe	Ser		
110	65					70					75					80		
111																		
112						CTA												407
113	Gln	Leu	Val	Glu	Glu	Leu	Leu	Lys	Ile	Ile	Cys	Ala	Phe	Gln	Leu	Asp		
114					85					90					95			
115																		
116	ACA	GGT	TTG	GAG	TAT	GCA	AAC	AGC	TAT	AAT	TTT	GCA	AAA	AAG	GAA	AAT		455
117	Thr	Gly	Leu	Glu	Tyr	Ala	Asn	Ser	Tyr	Asn	Phe	Ala	Lys	Lys	Glu	Asn		
118				100					105					110				
119																		
120						СТА												503
121	Asn	Ser	Pro	Glu	His	Leu	Lys	Asp	Glu	Val	Ser	Ile	Ile	Gln	Ser	Met		
122			115					120					125					
123																		
124						GCC												551
125	Gly	-	Arg	Asn	Arg	Ala		Arg	Leu	Leu	Gln		Glu	Pro	Glu	Asn		
126		130					135					140						
127		-		a. a	<i>.</i>	100	1.00			ama		-	mam		000			600
128						ACC												599
129		ser	Leu	GIN	GTU	Thr	Ser	Leu	Ser	var		Leu	ser	ASN	Leu			
130	145					150					155					160		
131 132	200	ama		2.00	000	200			010	000	2002	<b>73 3</b>	000	(1) N	220	100		647
132						AGG Arg												04/
133	TITT	Var	Arg	1111	165	ALA	THE	цуз	OTH	170	TTE	0111	FIO	GTU	175	7117		
135					100					1,0					1,0			
136	TOT	GTC	TAC	ልጥጥ	GAA	TTG	GGA	TCT	GAT	TCT	TCT	GAA	GAT	ACC	GTT	ልልጥ		695
137						Leu												
138			- 7 -	180			1		185				<u>F</u>	190				
139																		
140	AAG	GCA	ACT	TAT	TGC	AGT	GTG	GGA	GAT	CAA	GAA	TTG	TTA	CAA	ATC	ACC		743
141	Lys	Ala	Thr	Tyr	Cys	Ser	Val	Gly	Asp	Gln	Glu	Leu	Leu	Gln	Ile	Thr		
142	-		195	-	- ,			200	-				205					
143																		
144	CCT	CAA	GGA	ACC	AGG	GAT	GAA	ATC	AGT	TTG	GAT	TCT	GCA	AAA	AAG	GCT		791
145	Pro	Gln	Gly	Thr	Arg	Asp	Glu	Ile	Ser	Leu	Asp	Ser	Ala	Lys	Lys	Ala		
146		210				_	215					220						
147																		
148	GCT	TGT	GAA	TTT	TCT	GAG	ACG	GAT	GTA	ACA	AAT	ACT	GAA	CAT	CAT	CAA		839
149	Ala	Cys	Glu	Phe	Ser	Glu	Thr	Asp	Val	$\mathbf{Thr}$	Asn	Thr	Glu	His	His	Gln		
150	225					230					235					240		
151		•																
152	CCC	AGT	AAT	AAT	GAT	TTG	AAC	ACC	ACT	GAG	AAG	CGT	GCA	GCT	GAG	AGG	,	887

.

### RAW SEQUENCE LISTING PATENT APPLICATION US/08/409,305

DATE: 09/21/95 TIME: 10:39:03

														IN	IPIIT	SET:	S6247.raw
153 154 155	Pro	Ser	Asn	Asn	Asp 245	Leu	Asn	Thr	Thr	Glu 250	Lys	Arg	Ala				
156	CAT	CCA	GAA	AAG	TAT	CAG	GGT	AGT	тст	GTT	TCA	AAC	TTG	CAT	GTG	GAG	935
157 158 159										Val							
160	CCN	TOT	aac	202	አለመ	እሮም	ርአም	ann	NGC	TCA	ጥጥ እ	CNG.	0 <b>N</b> T	ava	220	NGC	983
161										Ser							505
162 163	110	C y B	275		AGII		111.5	280	Der	Der	Ded	OTH	285	910	AGII	561	
164	AGT	TTA	TTA	CTC	ACT	AAA	GAC	AGA	ATG	AAT	GTA	GAA	AAG	GCT	GAA	TTC	1031
165 166	Ser	Leu 290	Leu	Leu	Thr	Lys	Asp 295	Arg	Met	Asn	Val	Glu 300	Lys	Ala	Glu	Phe	
167	-																
168										GCA							1079
169 170	305	ASN	ràa	ser	Lys	310	PIO	στγ	Leu	Ala	Arg 315	ser	GTU	AIS	ASN	Arg 320	
171	305					210					315					320	
172	TGG	GCT	GGA	AGT	AAG	GAA	ACA	TGT	AAT	GAT	AGG	cgg	ACT	ccc	AGC	ACA	1127
173										Asp							
174	-		-		325			-		330	-	-			335		
175																	
176										CCC							1175
177	Glu	Lys	Lys		Asp	Leu	Asn	Ala	-	Pro	Leu	Cys	Glu	-	Lys	Glu	
178				340					345					350			
179				~ ~ ~				-				~~~		~ . ~			
180										GAG							1223
181 182	тгр	ASN	цуз 355	GTU	Lys	Leu	Pro	360	ser	Glu	ASN	bio	Arg 365	Asp	THE	GTU	
182			300					300					202				
184	GAT	GTT	ССТ	TGG	АТА	ACA	CTA	ААТ	AGC	AGC	ATT	CAG	ΑΑΑ	GTT	ААТ	GAG	1271
185										Ser							
186		370					375					380					
187																	
188	TGG	TTT	TCC	AGA	AGT	GAT	GAA	CTG	TTA	GGT	TCT	GAT	GAC	TCA	CAT	GAT	1319
189	Trp	Phe	Ser	Arg	Ser	Asp	Glu	Leu	Leu	Gly	Ser	Asp	Asp	Ser	His	Asp	
190	385					390					395					400	
191		~ ~ ~		<i>.</i>							<b>~</b> • • •		-	~~~	000		10/7
192 193										GCT							1367
193	GTÀ	Gru	ser	Gru	405	ASN	ATa	Lys	var	Ala 410	Asp	var	Leu	Asp	415	Leu	
194					405					410					415		
196	ААТ	GAG	GTA	GAT	GAA	TAT	тст	GGT	TCT	TCA	GAG	AAA	АТА	GAC	TTA	CTG	1415
197										Ser							
198				420		-	_	-	425			-		430			
199																	
200										TGT							1463
201	Ala	Ser		Pro	His	Glu	Ala			Cys	Lys	Ser		Arg	Val	His	
202			435					440	•				445				
203	-			-													
204										GAC							1511
205	ser	Lys	Ser	va⊥	GIU	ser	ASN	тте	GTU	Asp	гλг	тте	РПе	σтУ	гàг	Tur	

GeneDX 1010, pg. 202

## RAW SEQUENCE LISTING PATENT APPLICATION US/08/409,305

•

•

DATE: 09/21/95 TIME: 10:39:05

INPUT SET: S6247.raw

																041.0044/	.1417
206		450					455	•				460					
207 208	ጥልጥ	caa	AAG	AAG	GCA	AGC	CTC	ccc	AAC	ጥጥል	AGC	САТ	GTA	ACT	GAA	ልልጥ	1559
209												His					1007
210	465	9	1-	-1-		470					475					480	
211											_						
212	CTA	ATT	ATA	GGA	GCA	TTT	GTT	ACT	GAG	CCA	CAG	ATA	ATA	CAA	GAG	CGT	1607
213	Leu	Ile	Ile	Gly	Ala	Phe	Val	Thr	Glu	Pro	Gln	Ile	Ile	Gln	Glu	Arg	
214				-	485					490					495		
215																	
216												CCT					1655
217	Pro	Leu	Thr		Lys	Leu	Lys	Arg	-	Arg	Arg	Pro	Thr		Gly	Leu	
218				500					505					510			
219	<b>a</b> 2 <b>m</b>		~~~	~ ~ ~		3 10 0	110			<b></b>	mma			~ ~ ~ ~	220	A CIT	1 7 6 2
220 221												GCA					1703
221	пта	PIO	515	Asp	Pne	тте	гуs	520	ALd	Азр	Leu	Ala	525	GTU	uys	1111	
223			212					520					525				
224	ССТ	GAA	ATG	АТА	ААТ	CAG	GGA	ACT	AAC	CAA	ACG	GAG	CAG	ТАА	GGT	CAA	1751
225												Glu					
226		530					535					540			-		
227																	
228	GTG	ATG	AAT	ATT	ACT	AAT	AGT	GGT	CAT	GAG	AAT	AAA	ACA	AAA	GGT	GAT	1799
229	Val	Met	Asn	Ile	Thr	Asn	Ser	Gly	His	Glu	Asn	Lys	Thr	Lys	Gly	Asp	
230	545					550					555					560	
231																	
232												GAA			-		1847
233	Ser	Ile	GIn	Asn		Lys	Asn	Pro	Asn		11e	Glu	ser	Leu		Lys	
234 235					565					570					575		
235	GAA	ምረሞ	COT	ጥጥረግ	***	ACG	אגא	0.CTT	(1) N	CCT	מידים	AGC	NGC	አርሞ	ልሞል	ACC	1895
237												Ser					1095
238		~~~		580	-1-		-1-2		585					590			
239																	
240	AAT	ATG	GAA	CTC	GAA	TTA	AAT	ATC	CAC	AAT	TCA	AAA	GCA	CCT	AAA	AAG	1943
241	Asn	Met	Glu	Leu	Glu	Leu	Asn	Ile	His	Asn	Ser	Lys	Ala	Pro	Lys	Lys	
242			595					600					605				
243							_										
244												ATT					1991
245	Asn	-	Leu	Arg	Arg	Lys			Thr	Arg	His	Ile	His	Ala	Leu	GLU	
246 247		610					615					620					
247	ሮሞእ	C T T A	ame	አረሞ	202	እእጥ	ሮሞል	vaa	003	COT	አአጥ	TGT	አርሞ	GVV	ጥጥር	C & A	2039
240												Cys				_	2037
250	625	Vur	101	001	nr 9	630	Jeu	501			635	0,0	1111	01u	Lea	640	
251						•										. = =	
252	ATT	GAT	AGT	TGT	TCT	AGC	AGT	GAA	GAG	АТА	AAG	AAA	AAA	AAG	TAC	AAC	2087
253												Lys					
254		-		-	645					650	-	-	~	-	655		
255																	
256												CTC					2135
257	Gln	Met	Pro		Arg	His	Ser	Arg		Leu	Gln	Leu	Met		Gly	Lys	
258				660					665					670			

•

## SEQUENCE VERIFICATION REPORT PATENT APPLICATION US/08/409,305

DATE: 09/21/95 TIME: 10:39:08

INPUT SET: S6247.raw

Line

30 926 Error

Wrong application Serial Number Stop Codon at end of sequence removed - no error Original Text

#### (A) APPLICATION NUMBER: US



UNITED STATE 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

08/409,305 03/24/95

#### SHATTUCK-EIDENS

24884-109347

o #4

0252/0927 VENABLE BAETJER HOWARD & CIVILETTI 1201 NEW YORK AVE SUITE 1000 WASHINGTON DC 20005

> 0000 DATE MAILED

Ð

#### NOTICE TO FILE MISSING PARTS OF APPLICATION 09/27/95 **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 \$ 30 for large entities or _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  $\Box$  large entity,  $\Box$  small entity (verified statement filed), is  $\underline{\$}_{\Box}$ 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 ☐ Targe entity □ small entity, must submit \$_______ to complete the basic filing fee. _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 the newly submitted items.

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.  $\Box$  The signature(s) to the oath or declaration is/are:  $\Box$  missing;  $\Box$  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.  $\Box$  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 _under 37 CFR 1.17(k), unless this fee has translation of the application and a fee of \$_ already been paid.
- 8. 🗆 A \$ processing fee is required since your check was returned without payment. (37 CFR 1.21(m)).
- 9. 🗆 Your filing receipt was mailed in error because your check was returned without payment.
- 10.  $\Box$  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 to Box Missing Part and refer any questions to the Customer Service Center at (703) 308-1202.

#### A copy of this notice <u>MUST</u> be returned with the response.^{1010, pg. 205} OFFICE COPY

#### 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 MUTATIONS

for which an application for Letters Patent of the United States was filed on 24 March 1995 and assigned Serial No. 08/409,305.

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:

#### NONE

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:

Ser. No. 08/348,824, filed 29 November 1994
Ser. No. 08/308,104, filed 16 September 1994
Ser. No. 08/300,266, filed 2 September 1994
Ser. No. 08/289,221, filed 12 August 1994
Ser. No. 08/246,002, filed 20 May 1994

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.

¥9

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; E. Brendan Magrab, Registration No. 36,205; and Stephen A. Saxe, Registration No. 38,609; 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	ー っ こ Donna M. SHATTUCK-EIDENS
Inventor's Signature	Donina M. Shuthuck - Eidus
Date of Signature	Oct 20, 1995
Residence Address	1874 Texas Street Salt Lake City, Utah 84108
Post Office Address	SAME
Country of Citizenship	USA
Full Name of Second Inventor	Jacques SIMARD
Inventor's Signature	·
Date of Signature	· · · · · · · · · · · · · · · · · · ·
Residence Address	4808 Pilet St-Augustin de Desmaures Quebec, Canada, G3A 1Z9
Post Office Address	SAME
Country of Citizenship	CANADA

4

-2-

Full Name of Third Inventor	Mitsuru EMI
Inventor's Signature	
Date of Signature	
Residence Address	
Post Office Address	SAME
Country of Citizenship	JAPAN
Full Name of Fourth Inventor	Yusuke NAKAMURA
Inventor's Signature	
Date of Signature	<u> </u>
Residence Address	
Post Office Address	SAME
Country of Citizenship	JAPAN

,

-3-



#### DECLARATION AND POWER OF ATTORNEY

As below-named inventor, I hereby declare that my residence, post office address and ship 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 MUTATIONS

for which an application for Letters Patent of the United States was filed on 24 March 1995 and assigned Serial No. 08/409,305.

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:

#### NONE

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:

Ser. No. 08/348,824, filed 29 November 1994
Ser. No. 08/308,104, filed 16 September 1994
Ser. No. 08/300,266, filed 2 September 1994
Ser. No. 08/289,221, filed 12 August 1994
Ser. No. 08/246,002, filed 20 May 1994

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; E. Brendan Magrab, Registration No. 36,205; and Stephen A. Saxe, Registration No. 38,609; 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	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 Second Inventor	2 2 D Jacques SIMARD
Inventor's Signature	loca sking
Date of Signature	October 23 /995
Residence Address	4808 Pilet St-Augustin de Desmaures Quebec, Canada, G3A 1Z9 CAX
Post Office Address	SAME
Country of Citizenship	CANADA

-2-

Full Name of Third Inventor	Mitsuru EMI
Inventor's Signature	
Date of Signature	
Residence Address	
Post Office Address	SAME
Country of Citizenship	JAPAN
Full Name of Fourth Inventor	Yusuke NAKAMURA
Inventor's Signature	
Date of Signature	
Residence Address	
Post Office Address	SAME
Country of Citizenship	JAPAN

-3-

'96 01/05 FRI 10:25 FAX

JAN 03'96 10:51 FR

HABLE BAETJER

#### 202 962 _ 20 TO

Ø 003

P.02



#### 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 MUTATIONS

for which an application for Letters Patent of the United States was filed on 24 March 1995 and assigned Serial No. 08/409,305.

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:

#### NONE

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:

Ser. No. 08/348,824, filed 29 November 1994
Ser. No. 08/308,104, filed 16 September 1994
Ser. No. 08/300,266, filed 2 September 1994
Ser. No. 08/289,221, filed 12 August 1994
Ser. No. 08/246,002, filed 20 May 1994

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.

#### '96 01/05 FRI 10:27 FAX

JAN 03'96 10:52 FR

NABLE BAETJER

202 962 20 TO

Ø004

-2-

I hereby appoint Jeffrey L. Ihnen, Registration No. 28,957; Fred W. Hathaway, Registration No. 32,236; E. Brendan Magrab, Registration No. 36,205; and Stephen A. Saxe, Registration No. 38,609; 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	Donna M. SHATTUCK-EIDENS
Inventor's Signature	
Date of Signature	<b></b>
Residence Address	1874 Texas Street Salt Lake City, Utah 84108
Post Office Address	SAME
Country of Citizenship	USA
Full Name of Second Inventor	Jacques SIMARD
Inventor's Signature	
Date of Signature	••••••••••••••••••••••••••••••••••••••
Residence Address	4808 Pilet St-Augustin de Desmaures Quebec, Canada, G3A 1Z9
Post Office Address	SAME
Country of Citizenship	CANADA

P.03

96 01/05 FRI 10:28 FAX

২৩২ ৬

-3-

196

Azamino

Aoba

JPX

Japan JPX

8300 10

2005

P.(

Full Name of Third 300 Inventor Mitsuru EMI miterim Inventor's Signature 1/18/96 Date of Signature 3-20-208, Sengoku , Bunkyo-ku **Residence** Address 112, Tokyo Post Office Address SAME Country of Citizenship JAPAN 400 Full Name of Fourth Inventor Yusuke NAKAMURA Inventor's Signature 2 ł 12 Date of Signature **Residence** Address 3-4-1-401 Yokohama 225, Japan Post Office Address SAME

Country of Citizenship

JAPAN

GeneDX 1010, pg. 214



COMMERCE

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

Ð

FILING DATE

08/409,305

SHATTUCK-EIDENS

24884-109347

ATTY. DOCKET NO/TITLE

0252/0927

FIRST NAMED APPLICANT

VENABLE BAETJER HOWARD & CIVILETTI 1201 NEW YORK AVE SUITE 1000 DC 20005 WASHINGTON

03/24/95

## DATE MAILED: 0000

#### NOTICE TO FILE MISSING PARTS OF APPLICATION 09/27/95 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 PARMENT OF A SURCHARGE for items 1 and 3-6 only of \$______for large entities or THE PAYMENT OF A SURCHARGE for items 1 and 3-6 only of \$_ for large entities or IK 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 <u>860</u> 

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: Emissing 🗆 insufficient. Applicant as a E large entity 🗆 small 50 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.

 $\Box$  does not cover the newly submitted items.

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.  $\Box$  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(s) to the oath or declaration is/are: missing; 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.  $\Box$  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 since your check was returned without payment. (37 CFR 1.21(m)).
- 9. I Your filing receipt was mailed in error because your 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 to Box Missing Part and refer any questions to the Customer Service Center 110 101 750 OOCH a64703) 369 1208 02/05/96 08409305 22-0261

Met 1009 02/05/98 08409305 22-0265 110 A copy of this notice MUST be returned with



## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Donna M. SHATTUCK-EIDENS et al.

Serial No. 08/409,305

ATTN: Box Missing Part

Filed: 24 March 1995

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

#### TRANSMITTAL OF FORMAL DOCUMENTS

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

Sir:

In response to the "Notice to File Missing Parts -- Filing Date Granted" mailed for the above-identified application on 27 September 1995, transmitted herewith are the following:

- Declaration and Power of Attorney documents (two individually signed by coinventors Shattuck-Eidens and Simard, and one facsimile signed by coinventors Emi and Nakamura);
- 2. Assignment documents (one from Shattuck-Eidens to Myriad Genetics, Inc., one from Simard to Centre de Recherche du Chul, and one facsimile from Emi and Nakamura to Cancer Institute); and
- 3. A copy of Form PTO-1449.

Please charge the combined filing fee (\$750), surcharge (\$130) and assignments recordation fee (\$120) to our Deposit Account 22-0261. Any discrepancies in fee payment should also be charged to said Deposit Account. This transmittal letter is accompanied by a Petition for Extension of Time, with appropriate fee, and is submitted in triplicate.

Respectfully submitted,

Juffing Threes ev L. Knen, Reg. No. 28,957

VENABLE, BAETJER, HOWARD & CIVILETTI (202) 962-4810 Atty. Docket No. 24884-109347 Dated: 23 January 1996


# IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

pplication of:

Donna M. SHATTUCK-EIDENS et al.

Serial No. 08/409,305

ATTN: Box Missing Part

16

Filed: 24 March 1995

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

## **PETITION FOR EXTENSION**

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

Sir:

In response to the "Notice to File Missing Parts -- Filing Date Granted" mailed for the above-identified application on 27 September 1995, and concurrently with the documents filed in response thereto, petition is hereby made for an extension of three (3) months in which to respond.

Please charge the requisite three-month extension fee of \$900.00, and any noted discrepancies in fee payment, to Deposit Account No. 22-0261. This Petition is submitted in triplicate.

Respectfully submitted,

Ihnen, Reg. No. 28,957

VENABLE, BAETJER, HOWARD & CIVILETTI Washington, D.C. 20005 (202) 962-4810

Atty. Docket No. 24884-109347 Dated: 23 January 1996

MG11007 02/05/96 08409305 22-0261 110 117 900.00CH

Start a	ENT OF CO	AVALANCE
· UMITED	TATES OF	S. C.
	1168 0	

PTO-90C (REV. 2/95)

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

	APPLICATION NO.	FILING DATE	FIRST NAMED INVENT	OR		ATTORNEY DOCKET NO.
	08/409,	305 03/2	4/95 SHATTUCK-EIDENS	3	Þ	24884-109347
<b> </b>			18M2/0703	¬ [	REES,	KAMINER
		BAETJER H				PAPER NUMBER
	WASHING	W YORK AVE Ton DC 20	NUE NW SUITE 1000 005		1807	7
				D	ATE MAILED:	07/03/96

Please find below and/or attached an Office communication concerning this application or proceeding.

**Commissioner of Patents and Trademarks** 

		·		······
	Application No. 08/409,305	Applicant(s)	SHATTUCK-EIDE	NS ET AL.
Office Action Summary	Examiner	_ <u>_</u>	Group Art Unit	
	Dianne Re	es 	1807	
Responsive to communication(s) filed on Jan 23, 1996	5			
This action is FINAL.				
Since this application is in condition for allowance exce in accordance with the practice under Ex parte Quayle	•		on as to the me	rits is closed
A shortened statutory period for response to this action is is longer, from the mailing date of this communication. F application to become abandoned. (35 U.S.C. § 133). E 37 CFR 1.136(a).	ailure to respond with	nin the perio	d for response v	will cause the
Disposition of Claims				
X Claim(s) <u>1-18</u>	<u>·</u>	is/	are pending in t	he application.
Of the above, claim(s)		is/are	e withdrawn fro	m consideration.
Claim(s)		· · · · · · · · · · · · · · · · · · ·	is/are allowe	d.
X Claim(s) <u>1-18</u>				
Claim(s)				
Claims	are sul	bject to rest	riction or electio	on requirement.
<ul> <li>See the attached Notice of Draftsperson's Patent D</li> <li>The drawing(s) filed on is/are</li> <li>The proposed drawing correction, filed on</li> <li>The specification is objected to by the Examiner.</li> <li>The oath or declaration is objected to by the Examine</li> <li>Priority under 35 U.S.C. § 119</li> <li>Acknowledgement is made of a claim for foreign pr</li> <li>All Some* None of the CERTIFIED cog</li> <li>received.</li> <li>received in Application No. (Series Code/Serie</li> <li>received in this national stage application from * Certified copies not received:</li> <li>Acknowledgement is made of a claim for domestic</li> </ul> Attachment(s) <ul> <li>Notice of References Cited, PTO-892</li> <li>Information Disclosure Statement(s), PTO-1449, Pa</li> <li>Interview Summary, PTO-413</li> <li>Notice of Draftsperson's Patent Drawing Review, P</li> <li>Notice of Informal Patent Application, PTO-152</li> </ul>	e objected to by the E is ner. iority under 35 U.S.C pies of the priority do al Number) m the International Bu priority under 35 U.S	Examiner. approved [ C. § 119(a)- ocuments ha ureau (PCT 1	(d). ve been  Rule 17.2(a)).	
SEE OFFICE ACTION	ON THE FOLLOWING	PAGES		

PTO-326 (Rev. 9-95)

Office Action Summary

•

____

Part III DETAILED ACTION

#### Specification

1. The disclosure is objected to because of the following informalities: The specification refers to "our Utah collaborators" without providing names ( a reference or a footnote for example) to indicate who these collaborators are. Appropriate correction is required.

2. The following is a quotation of the first paragraph of 35 U.S.C. § 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1-7, and 9-18 are rejected under 35 U.S.C. § 112, first paragraph, as the disclosure is enabling only for claims directed to defined sequences, and to methods for detection of susceptibility to breast and ovarian cancer caused by alterations at specific target sequences within the BRCA1 gene limited to using assays which employ allele specific probes. See M.P.E.P. §§ 706.03(n) and 706.03(z).

Factors to be considered in determining whether a disclosure meets the enablement requirement of 35 USC 112, first paragraph,

-2-

have been described in *In Re Colianni*, 195 USPQ 150 (CCPA 1977) and have been adopted by the Board of Patent Appeals and Interferences in *Ex Parte Forman*, 230 USPQ 546 (BPAI 1986). Among these factors are: the nature of the invention, the state of the prior art, the predictability or lack thereof in the art, the breadth of the claims, the amount of direction or guidance present, and the presence or absence of working examples.

The claims are drawn to isolated DNA sequences "consisting essentially of DNA encoding for a mutant BRCA1 polypeptide said DNA selected from the group consisting of mutations set forth in Table 14. Further claims are drawn to methods of "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 and selected from the group consisting of the mutations set forth in Tables 11, 12, 12A and 14. "

The product claims thus encompass sequences which differ in an unspecified manner from those disclosed in Table 14. As most of the mutations in Table 14 are a single base change, the sequences may encompass additional mutations to the ones disclosed, different mutations at the same sites as those disclosed in Table 14 and might even encompass wild type sequences. The sequences are disclosed as having utility in a

GeneDX 1010, pg. 221

-3-

specific diagnostic test to determine predisposition for breast and ovarian cancer.

A variety of assays are set forth and claimed to detect polymorphisms including detecting BRCA 1 mRNA, such as by hybridization with a BRCA1 gene probe, by determining shifts in electrophoretic mobility of single stranded DNA, by general hybridization assays which may further include the use of allele specific oligos as probes, amplification assays which may be further combined with hybridization assays with probes directed against wild type or mutant sequences, cloning of BRCA1 sequences, detection of mismatches, and in situ hybridization. The claims are thus drawn to the detection of specific polymorphisms but the majority of assays claimed provide only general data which will allow one to determine that there is "a polymorphism" without any predictably certainty as to what the nature of the polymorphism is (as discussed above the product claims are further ambiguous as to what the nature of the polymorphism even is). It is highly unpredictable that any unselected polymorphism will be enabled as a screening tool, as recited by the claims- "said alteration indicating a predisposition to said (breast and ovarian) cancer". The large size of the BRCA1 gene makes it a target for many mutations and there is no apparent clustering for these mutations or association with the severity of breast or ovarian cancer (See Shattuck Eidens et al. JAMA 273:535541, 1995). It is further

GeneDX 1010, pg. 222

-4-

likely that many polymorphisms in fact will be innocuous and not predispose one to any risk. It is not predictable that any given polymorphism which is linked to breast cancer will have utility as a screening tool for ovarian cancer as although the art teaches that there is a higher overall rate of ovarian cancer in individuals with BRCA1 mutations, the incidence of ovarian cancer varies unpredictably with the nature of the mutant allele (see Merajver et al,. Nature Genetics 9: 439-443). The art teaches that "the molecular basis for heterogeneity of penetrance, organ of origin of the cancer, age of onset and laterality of the lesions has not yet been elucidated (see page 439). The specification raises relevant criteria for the assessment of a predisposing mutation:

1) The variant must 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." Thus detection assays which are not geared to the detection of specific mutations (those for which these criteria have been met) will not predictably provide useful markers for risk of breast cancer of ovarian cancer.

It is therefore the position of the examiner that it would require undue experimentation to make and use the isolated DNA sequences as broadly claimed.

GeneDX 1010, pg. 223

-5-

Claim Rejections - 35 USC § 112

3. Claims 1-18 are rejected under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

The following phrases render the claims vague and indefinite:

a) In claim 1, the recitation of "An isolated DNA sequence consisting essentially of DNA encoding for a mutant BRCA1 polypeptide " renders the claim vague and indefinite in that it is unclear how "essentiallY" is defined, and therefore it is not clear what the metes and bounds of the DNA sequences being claimed are. Further the recitation of "said DNA selected from the group consisting of mutations" in that it is unclear how the "sequences" may be selected from a group of mutations per se. The claim might be amended to recite --said DNA selected from a group of DNA sequences consisting of mutations--.

b) Claim 2 is indefinite in the recitation of "A nucleic acid probe complementary to human altered BRCA1 sequences" in that it is not clear what the extent of complementarity or specificity of these probes are. Do the probes distinguish mutant vs. wild type sequences for example? Does complementary refer to

-6-

full complementarity? The claim should provide some functional language to define the probes or more specifically recite the sequences encompassed by the probes.

c) Claim 3 is indefinite in the recitation of "the wild type BRCA1 gene" as the term lacks proper antecedent basis in the claim.

d) In claim 6, the recitation of "wherein the alterations in the germline sequence is detected by observing shifts in electrophoretic mobility of single stranded DNA on non-denaturing polyacrylamide gels, renders the claim indefinite because it is not clear where the single stranded DNA is derived from in this claim and further what the point of comparison is to allow one to determine a "shift in electrophoretic mobility".

e) Claim 8 is indefinite in that the claim recites --wherein the alteration of the wild type BRCA 1 gene by hybridization with an allele specific probe" as the claim does not recite what the probe is to be hybridized to.

d) Claim 10 is indefinite in the recitation that "the alteration of the wild type BRCA1 gene is detected by amplifying all or part of the BRCA1 gene for a specific BRCA1 allele-- in that it is unclear if the claim intends that only the mutant allele is amplified or if both wild type or mutant allele may be amplified in the method. Clarification is requested.

e) Claim 14 is indefinite in that it is unclear where the non-wild type sequences are derived from (are these allele

specific oligos or simply nucleic acids isolated from affected individuals?).

Double Patenting

A rejection based on double patenting of the "same invention" type finds its support in the language of 35 U.S.C. 101 which states that "whoever invents or discovers any new and useful process ... may obtain <u>a</u> patent therefor ..." (Emphasis added). Thus, the term "same invention," in this context, means an invention drawn to identical subject matter. *Miller v. Eagle Mfg. Co.*, 151 U.S. 186 (1894); *In re Ockert*, 245 F.2d 467, 114 USPQ 330 (CCPA 1957); and *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970).

The non-statutory double patenting rejection, whether of the obvious-type or non-obvious-type, is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent. In re Thorington, 418 F.2d 528, 163 USPQ 644 (CCPA 1969); In re Vogel, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); In re Van Ornam, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); In re Longi, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); and In re Goodman, 29 USPQ2d 2010 (Fed. Cir. 1993).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321 (b) and (c) may be used to overcome an actual or provisional rejection based on a non-statutory double patenting ground provided the conflicting application or patent is shown to be commonly owned with this application. See 37 CFR 1.78 (d).

Effective January 1, 1994, a registered attorney or agent of record may sign a Terminal Disclaimer. A Terminal Disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

Claims 3-18 are provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1-16 of copending application Serial No. 08483553. Although the conflicting claims are not identical, they are not patentably distinct from each other because the methods recited in 08409305 meet all of the limitations of claims 1-19 in 08483553.

Claims 3-18 are provisionally rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1-28 of copending application Serial No. 08487002. Although the conflicting claims are not identical, they are not patentably distinct from each other because the methods recited in 08409305 meet all of the limitations of claims 1-28 of 08487002.

This is a *provisional* obviousness-type double patenting rejection because the conflicting claims have not in fact been patented.

The examiner notes that the above rejections are not an exhaustive list of examples of double patenting related to the above application and reminds the applicant that as the prosecution advances, conflicting claims in all related applications will have to be either cancelled, or changed in scope (with the filing of a terminal disclaimer).

5. The following references are cited as relevant intervening art:

Castilla et al. Nature Genetics 8: 387-391, Dec, 1994, teaches mutations in BRCA1 in families with early onset breast and ovarian cancer.

Friedman et al. Nature Genetics 8: 1-6, Dec 1994 teaches the confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families.

Simard et al. Nature Genetics 8: 392-397 teaches BRCA mutations in Canadian breast and ovarian cancer families.

Goldgar et al. J. of the National Cancer Institute 86: 200-209, Feb 2, 1994, teaches a large kindred with 17q Linked Breast and Ovarian Cancer.

Futreal et al. Science 266 120-122, 1994, teaches BRCA mutations in Primary Breast and Ovarian Carcinomas.

-9-

Hosking et al. Nature Genetics 9: 343-344, April 1995 teaches a somatic BRCA1 mutation in an ovarian tumor. Holt et al.WO 95/19369 teaches a method detecting and

diagnosing preinvasive breast cancer by identifying differentially expressed genes in early, pre invasive breast cancer tissue and methods for screening for compounds which induce the expression of BRCA1.

6. No claims are allowed.

7. Papers related to this application may be submitted to Group 1800 by facsimile transmission via the P.T.O. Fax Center located in Crystal Mall 1. The CM1 Fax Center number is (703) 305-7401. Please note that the faxing of such papers must conform with the notice to Comply published in the Official Gazette, 1096 OG 30 (Nov 15, 1989).

An inquiry regarding this communication should be directed to examiner Dianne Rees, Ph.D., whose telephone number is (703) 308-6565. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, W. Gary Jones, can be reached on (703) 308-1156.

Calls of a general nature may be directed to the Group receptionist who may be reached at (703) 308-0196.

Dianne Rees

June 18, 1996

W. GARY JONES SUPERVISORY PATENT EXAMINER GROUP 1800

-10-

				Application No. 08/409 305	Applicant(s)	SKOLNICK	ET AL.	
		Notice of Refere	nces Cited	Examiner		Group Art Unit		
		· · · · · · · · · · · · · · · · · · ·		Dianne Re	98S	1807	F	Page 1 of 3
<u> </u>		·····	U.S	. PATENT DOCUMENTS				
*		DOCUMENT NO.	DATE	NAME	I 		CLASS	SUBCLASS
_	A				<u></u>			
 	В							
	C							
	D							
	E			· ·				
	F							
	Ģ							
	н							
	1					-1		
<u> </u>	J							
-	к				· · · · · · · · · · · · · · · · · · ·			
	L				••••••••••••••••••••••••••••••••••••••			
┝	M						· · ·	
-	<u> </u>		FORE	IGN PATENT DOCUMENTS				<u> </u>
*	<u> </u>	DOCUMENT NO.	DATE	COUNTRY	NAME		CLASS	SUBCLASS
*	N	- WO 95/19369		РСТ	HOLT ET	AL.		
	0			<del>\</del>	cted in 1487	7002		
	Р							
	۹							
	R			· ·				
	s							
	т							
	<u> </u>	hr	NO	N-PATENT DOCUMENTS			<u></u>	
*			DOCUMENT (Including Aut	thor, Title, Source, and Pertinent	Pages)			DATE
[		DORLAND'S MEDICAL DIC	TIONARY	1 m '487002	·			
*	U		.61760	40F002				
		GOLDBERG ET AL. CLINIC	AL CHEMISTRY 39: 2	360-2374				
*	v			cited in '487007	2			1993
		WOOSTER ET AL., SCIENC	CE 265:2088-2090			<u> </u>		
×	w			cited in 1487 or	72			1994
		SATO ET AL., CANCER RE	S. 52: 1643-1646		<u></u>			
×	X			cited in '48700'	2			1992

* A copy of this reference is not being furnished with this Office action. (See Manual of Patent Examining Procedure, Section 707.05(a).)

Notice of References Cited

		<u> </u>	<u></u>		Application No. 08/40930	25	Applicant(s)	SKOLNI			<u></u>
		Notice of Refere	nces Cite	ed 🛛	Examiner	,15		Group Art U		L.	
					1	nne Rees	3 .	1807		P	age 2 of 3
				U.S. F	PATENT DOCUMEN	NTS		<u> </u>			
*		DOCUMENT NO.	DATE			NAME			CLA	ss	SUBCLASS
	A										
	B		<u></u>								
	c		····								
	D										<u> </u>
	E			1							
	F					<u>,</u>					
	G							···			
	н						<del></del> <del></del>				
	1							·····			
	J				<u></u>					-	
	ĸ		·			<u> </u>					
	[]			· · · · · · · · · · · · · · · · · · ·							
	L			+							
	м		····-					- <u></u>			
	<u> </u>			FUREIG	N PATENT DOCUM	ENIS					
										1	
*		DOCUMENT NO.	DATE	c	OUNTRY		NAME		CLA	ss	SUBCLASS
*	N	DOCUMENT NO.	DATE	c	OUNTRY		NAME		CLA	SS	SUBCLASS
*	N O	DOCUMENT NO.		с	OUNTRY				CLA	ss	SUBCLASS
+				с	OUNTRY				CLA	SS	SUBCLASS
*	0	DOCUMENT NO.		c	OUNTRY				CLA	SS	SUBCLASS
*	O P			c					CLA	SS	SUBCLASS
*	0 P Q			c					CLA	SS	SUBCLASS
*	O P Q R			c		· · · · · · · · · · · · · · · · · · ·			CLA		SUBCLASS
*	O P Q R S								CLA	SS	SUBCLASS
*	O P Q R S			NON-1		NTS			CLA	SS	SUBCLASS
*	O P Q R S T		DOCUMENT (In	NON-I	PATENT DOCUME	NTS ertinent Pa			CLA	SS .	DATE
	O P Q R S		DOCUMENT (In	NON-I		NTS ertinent Pa			CLA		
*	O P Q R S T		DOCUMENT (Inc 250: 1233-12	NON-I	PATENT DOCUMEN or, Title, Source, and Pa Cited m ¹⁴ 8	NTS ertinent Pe			CLA		DATE
*	O P Q R S T	MALKIN ET AL., SCIENCE	DOCUMENT (Inc 250: 1233-12	NON-I	PATENT DOCUME	NTS ertinent Pe			CLA		DATE
*	O P Q R S T U	MALKIN ET AL.,SCIENCE BOYD, NATURE GENETIC	DOCUMENT (Inc 250: 1233-12 S 9: 335-336	NON-I cluding Autho	PATENT DOCUMEN r, Title, Source, and Pa Cited m '48 Cited m '48	NTS ertinent Pe			CLA		DATE 1990
*	O P Q R S T U	MALKIN ET AL., SCIENCE	DOCUMENT (Inc 250: 1233-12 S 9: 335-336	NON-I cluding Autho	PATENT DOCUMEN r, Title, Source, and Pa Cited m "48 cited m 3	NTS ertinent Pe 700 7 1487	nges)			SS	DATE 1990
* * ×	O P Q R S T U V	MALKIN ET AL.,SCIENCE BOYD, NATURE GENETIC MERAJVER ET AL, NATU	DOCUMENT (Inc 250: 1233-12 S 9: 335-336 RE GENETICS	NON-1 cluding Autho 38 9;439-443	PATENT DOCUMEN r, Title, Source, and Pa Cited m '48 Cited m '48	NTS ertinent Pe 700 7 1487	nges)			SS	DATE 1990 1995
* * ×	O P Q R S T U V	MALKIN ET AL.,SCIENCE BOYD, NATURE GENETIC	DOCUMENT (Inc 250: 1233-12 S 9: 335-336 RE GENETICS	NON-1 cluding Autho 38 9;439-443	PATENT DOCUMEN r, Title, Source, and Pa Cited m "48 cited m 3	NTS ertinent Pe 1700 7 1487 1487	nges)			SS	DATE 1990 1995

* A copy of this reference is not being furnished with this Office action. (See Manual of Patent Examining Procedure, Section 707.05(a).)

...

GeneDX 1010, pg. 230

		<u> </u>		Application No. 08/4ാറ്റുക്ക്	Applicant(s) SKOLNIC	CK ET AL.	
		Notice of Refere	ences Cited	Examiner Dianne Re	Group Art U es 1807		age 3 of 3
		<u></u>	U.S	S. PATENT DOCUMENTS			<u> </u>
$\square$		DOCUMENT NO.	DATE	NAME	· · · · ·	CLASS	SUBCLASS
[	A				•		
	B						
	с						ang ang tan
	D						
	E			, 4), 8,	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		
	F	······					
<u> </u>	G			,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,			
-	н		······				
	1			· · · · · · · · · · · · · · · · · · ·			
}	L						
	к			·····			
	L	······································	<u></u>				<u>_</u>
	м				<u></u>		
		<u> </u>	FORE	IGN PATENT DOCUMENTS	<u>, , , , , , , , , , , , , , , , , , , </u>		
		DOCUMENT NO.	DATE	COUNTRY	NAME	CLASS	SUBCLASS
	N						
	0						
	р						
	٥						
	R						
	s						
	т						
			NO	N-PATENT DOCUMENTS			
			DOCUMENT (Including Au	thor, Title, Source, and Pertinent	Pages)		DATE
	υ	FRIEDMAN ET AL., NATU	JRE GENETICS 8: 1-6	cited in 148700-	2		1004
*	Ū				•		1994
		GOLDGAR ET AL. J. OF	NAT. CANCER INST, 8				
×	V			cited m	1487002		1994
	w						
	**						
	x						
	l						

U. S. Patent and Trademark Office PTO-892 (Rev. 9-95)

•• • • • • • •

Part of Paper No. ____

GeneDX 1010, pg. 231

Form PTO 948 (Rev. 10-94)

U.S. DEPARTMENT OF COMMERCE - Patent and Trademark Office

Application No

## -NOTICE OF DRAFTSPERSON'S PATENT DRAWING REVIEW

PTO Draftpersons review all originally filed drawings regardless of whether they are designated as formal or informal. Additionally, patent Examiners will review the drawings for compliance with the regulations. Direct telephone inquiries concerning this review to the Drawing Review Branch, 703-305-8404.

the Drawing Review Branch, 703-303-6404.	
271/05	
The drawings filed (insert date) 3 24 19 Dare	View and enlarged view not labled separatly or properly.
A not objected to by the Draftsperson under 37 CFR 1.84 or 1.152.	Fig(s)
B objected to by the Draftsperson under 37 CFR 1.84 or 1.152 as	5 · · ·
indicated below. The Examiner will require submission of new, corrected	Sectional views. 37 CFR 1.84 (h) 3
drawings when necessary. Corrected drawings must be submitted	Hatching not indicated for sectional portions of an object.
according to the instructions on the back of this Notice.	Fig(s)
according to the instructions on the back of this Notice.	Cross section not drawn same as view with parts in cross section
1 DPAWINGS 37 CEP 1 84(a): Acceptable cotogories of drawinger	with regularly spaced parallel oblique strokes. Fig(s)
<ol> <li>DRAWINGS. 37 CFR 1.84(a): Acceptable categories of drawings: Black ink. Color.</li> </ol>	8. ARRANGEMENT OF VIEWS. 37 CFR 1.84(i)
Not black solid lines. Fig(s)	Words do not appear on a horizontal, left-to-right fashion when
Color drawings are not acceptable until petition is granted.	page is either upright or turned so that the top becomes the right
Fig(s)	side, except for graphs. Fig(s)
2. PHOTOGRAPHS. 37 CFR 1.84(b)	9. SCALE. 37 CFR 1.84(k)
Photographs acceptable until petition is granted.	Scale not large enough to show mechanism with crowding
Fig(s)	when drawing is reduced in size to two-thirds in reproduction.
Photographs hot properly mounted (must use brystol board or	Fig(s)
photographic double-weight paper). Fig(s)	Indication such as "actual size" or scale 1/2" not permitted.
Poor quality (half-tone). Fig(s)	Fig(s)
3. GRAPHIC FORMS. 37 CFR 1.84 (d)	10. CHARACTER OF LINES, NUMBERS, & LETTERS. 37 CFR
Chemical or mathematical formula not labeled as separate figure.	1. CHARACTER OF EINES, NUMBERS, & EETTERS. 57 CHR
Fig(s)	Lines, numbers & letters not uniformly thick and well defined,
Group of waveforms not presented as a single figure, using	clean, durable, a/d black (except for color drawings).
common vertical axis with time extending along horizontal axis.	Fig(s)
Fig(s)	
Individuals waveform not identified with a separate letter	11. SHADING. 37 CFR 1.84(m)
designation adjacent to the vertical axis. Fig(s)	Solid black shading areas not permitted.
4. TYPE OF PAPER. 37 CFR 1.84(c)	Fig(s)
Paper not flexible, strong, white, smooth, nonshiny, and durable.	Shade lines, pale, rough and blurred. Fig(s)
Sheet(s)	12. NUMBERS, LETTERS, & REFERENCE CHARACTERS. 37 CFR
Erasures, alterations, overwritings, interlineations, cracks, creases,	(()) + (())
and folds copy machine marks not accepted. Fig(s)	Yumbers and reference characters not plain and legible. 37 CFR
Mylar, velum paper is not acceptable (too thin). Fig(s)	1.84(p)(1) Fig(s) 1-0
5. SIZE OF PAPER. 37 CFR 1.84(f): Acceptable sizes:	Numbers and reference characters not oriented in same direction
21.6 cm. by $35.6$ cm. (8 1/2 by 14 inches)	as the view. 37 CFR 1.84(p)(1) Fig(s)
21.6  cm. by  33.1  cm.  (8 1/2  by  13  inches)	English alphabet not used. 37 CFR 1.84(p)(2)
21.6 cm. by 27.9 cm. (8 1/2 by 11 inches)	( /Tig(s)
21.0 cm. by 29.7 cm. (DIN size A4)	Numbers, letters, and reference characters do not measure at least
All drawing sheets not the same size. Sheet(s)	.32 cm. (122 nch) in height. 37 CFR(p)(3)
Drawing sheet not an acceptable size. Sheet(s)	Fig(s)
6. MARGINS. 37 CFR 1.84(g): Acceptable margins:	13. LEAD LINES. 37 CFR 1.84(q)
	Lead lines cross each other. Fig(s)
Paper size	Lead lines missing. Fig(s)
21.6 cm. X 35.6 cm. 21.6 cm. X 33.1 cm. 21.6 cm. X 27.9 cm. 21.0 cm. X 29.7 cm.	14. NUMBERING OF SHEETS OF DRAWINGS. 37 CFR 1.84(t)
(8 1/2 X 14 inches) (8 1/2 X 13 inches) (8 1/2 X 11 inches) (DIN Size A4)	Sheets not numbered consecutively, and in Arabic numerals,
T 5.1 cm. (2") 2.5 cm. (1") 2.5 cm. (1") 2.5 cm. L .64 cm. (1/4") .64 cm. (1/4") 2.5 cm.	
<b>R</b> .64 cm, $(1/4")$ .64 cm, $(1/4")$ .64 cm, $(1/4")$ .64 cm, $(1/4")$	beginning with number 1. Sheet(s) 15. NUMBER OF VIEWS. 37 CFR 1.84(u)
B .64 cm. (1/4") .64 cm. (1/4") .64 cm. (1/4") 1.0 cm.	Views not numbered consecutively, and infrabic numerals,
Margins do not conform the chart above.	
$\underbrace{ Sheet(s) }_{Top}(t) \underbrace{ Hett(L) }_{Right'(R)} \underbrace{ Bottom(B)}_{Bottom(B)} \underbrace{ Hett(L) }_{Right'(R)} \underbrace{ Hett(B)}_{Right'(R)} \underbrace$	S M. (The Fight)
	16: CORRECTIONS -37 CFR 1.84(w)
7. VIEWS. 37 OFR 1.84(b)	
<b>REMINDER:</b> Specification may require revision to correspond to	Corrections not made from prior PTO-948.
drawing changes.	Fig(s)
All views not grouped together. Fig(s)	17. DESIGN DRAWING. 37 CFR 1.152
Views connected by projection lines or lead lines.	Surface shading shown not appropriate. Fig(s)
Fig(s)	Solid black shading not used for color contrast.
Partial views. 37 CFR 1.84(h) 2	Fig(s)
COMMENTS:	
	ĺ
	(j.)
	х.
	GeneDX 1010, bg, 232
ATTACHMENT TO PAPER NO. REV	TEWERDATEDATE

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re. patent application of:

Donna M. SHATTUCK-EIDENS et al.

Serial No. 08/409,305

Filed: 24 March 1995

Examiner: D. Rees

Group Art Unit: 1807

For: 17Q-LINKED BREAST AND OVARIAN CANCER SUSCEPTIBILITY GENE MUTATIONS

## EXPRESS ABANDONMENT

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

Sir:

Applicants hereby expressly abandon the above-identified application, as of the date given hereinbelow.

Respectfully submitted

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. 24884-109347-04 Dated: 09 September 1996

** TOTAL PAGE.003 **

SEP 11'96 12:33 FR VENABLE POETJER

202 962 8300 TO 347#17033057401# P.01/03

### VENABLE, BAETJER, HOWARD & CIVILETTI, LLP 1201 New York Avenue, N.W., Suite 1000 Washington, D.C. 20005 Telephone No. (202) 962-4800 Telecopy No. (202) 962-8300

## FACSIMILE COVER SHEET

DATE: 11 September 1996 OUR FILE NO .: 24884-109347 IMMEDIATELY SEND: USER ID NO .: 2260 OFFICIAL TO: Examiner Dianne Rees, Group Art Unit 1807 **U.S. Patent and Trademark Office** FACSIMILE NO .: (703) 305-7401 **TELEPHONE NO.: (703) 308-6565** FROM: Jeffrey L. Ihnen TELEPHONE NO.: (202) 962-4810 TOTAL NO. OF PAGES: ______ (excluding cover sheet)

MESSAGE: Re. U.S. Serial No. 08/409,305; Shattuck-Eidens et al Our Ref.: 24884-109347-04

## SEE ATTACHED EXPRESS ABANDONMENT AND FILING RECEIPT

FAX CENTER RECEIVED

SEP 1 1 1996

GROUP 1800

#### CONFIDENTIALITY NOTICE

"WARNING: UNAUTHORIZED INTERCEPTION OF THIS TELEPHONIC COMMUNICATION COULD BE A VIOLATION OF FEDERAL AND STATE LAW." The documents accompanying this facsimile transmission contain confidential information belonging to the sender which is legally privileged. The information is intended only for the use of the individual or entity named above. If you are not the intended recipient, you are hereby notified that any disclosure, copying, distribution or the taking of any action in reliance on the contents of this facsimiled information is strictly prohibited. If you have received this facsimile in error, please immediately notify us by telephone to arrange for return of the original documents to us.

If there are any problems with the transmission, please call as soon as possible (202) 962-4994. Thank you!

SEP 11'96 12:34 FR VENABLE BAETJER 202 962 8300 TO 347#17033057401# P.02/03



# **Receipt for Documents Filed**

IN RE. XX PATENT TM INVENTOR/APPLICANT/OWNER: Donna M. SHATTUCK-EIDENS et al. REG. NO. SERIAL NO. 08/409,305 COUNTRY: PAT. NO. APPLICATION FILING DATE 24 March 19955 FOR: MARK MARK 17q-LINKED BREAST AND OVARIAN CANCER SUSCRPTIBILITY GENE MUTATIONS

## **DOCUMENTS FILED:**

EXPRESS ABANDONMENT

.

≠ d t t the second states and the se

DATE FILED: 10 September 1996

JLI:ch

DUE DATE: (NONE)

ATTY/SEC.

CLIENT/ MATTER NO. 24884-109347-04

CLIENT NAME: Myriad

Venable, Bactjer, Howard & Civiletti, LLP - 1201 New York Avenue, N.W., Suite 1000, - Washington, D.C. 20005 Telephone No. (202) 962-4800 - Facsimile: (202) 962-8300 18407.01 DC2DOCS 



### UNITED STATES DEPARTMENT OF COMMERCE Patent and Trademark Office

Address: COMMISSIONER OF PATENTS AND TRADEMARKS Washington, D.C. 20231

SERIAL NUMBER	FILING DATE	FIRST NAMED APPLICANT	ATTOR	NEY DOCKET NO.
08/40	19,305 037:	24/95 SHATTUCK-EIDENS	D	24884-1093
<u> </u>	· · · · · · · · · · · · · · · · · · ·	18M2/0919 <b>7</b>	EEES D	liner
JEFFF	EY L IHNEN			
		HOWARD & CIVILETTI	ART UNIT	PAPER NUMBER
		ENUE NW SUITE 1000		
WASHI	NGTON DC 2	0005	1807	9
		L	DATE MAILED:	09/19/96
		NOTICE OF ABANDONMENT		
This applica	tion is abandoned in	view of:		
1. 🗆 Appl	licant's failure to resp	ond to the Office letter, mailed		*
2. 🗶 Appl	licant's letter of expre	ss abandonment which is in compliance with	37 C.F.R. 1.138.	
	licant's failure to time od set in the Office le	ly file the response received	with	in the
mail	ing date of	the required issue fee within the statutory per of the Notice of Allowa	ince.	
		ceived on		
Π.	The issue fee has no	t been received in Allowed Files Branch as of		·
1	may petition the Com payment was unavoid	5 U.S.C. 151, and under the provisions of 37 ( missioner to accept the delayed payment of t dable. The petition must be accompanied by nitted, in the amount specified by 37 C.F.R. 1. e delay.	ne issue fee if the delay ir the issue fee, unless it ha	า เร
4		eceived the Notice of Allowance, a petition for holding of abandonment may be appropriate Q. 513.		Ce
draw	ings by	ly correct the drawings and/or submit new or a as received onas received on	substitute formal quired in the last Office ac	ction.
6. 🗆 The	reason(s) below.			
	•	SUPERVISORY PA GROUP	JONES TENT EXAMINER	

9/14/96

PTO-1432 (Rev. 4/93)

----

GeneDX 1010, pg. 236



Infre. patent application of:

Donna M. SHATTUCK-EIDENS et al.

Serial No. 08/409,305

Filed: 24 March 1995

Examiner: D. Rees

Group Art Unit: 1807

For: 17Q-LINKED BREAST AND OVAŘIAN CANCER SUSCEPTIBILITY GENE MUTATIONS

## **EXPRESS ABANDONMENT**

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

Sir:

1996

Applicants hereby expressly abandon the above-identified application, as of the date given hereinbelow.

Respectfully submitted

hnen

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. 24884-109347-04 Dated: 09 September 1996

## PETITION & UPPIC

0160 205

-år

- 1.et

103

PTO/SB/68 (11-85)

REQUEST FOR ACCESS OF ABANDONED APPLICATION UNDER 37 CFR 1.14 In re Application of Sh affuck-E:dews Application Number Fied 08/409305 3-24-95 Group Art Unit Externiner Assistant Commissioner for Patents Washington, DC 20231 I hereby request access underi37 CFR 1.14(a)(3)(v) to the application file record of the above- klerntified ABANDONED application, which is: (CHECK ONE) (A) referred to in United States Patent Number <u>5693 473</u> , column (B) referred to in United States Patent Number <u>5693 473</u> , column (C) an application that is open to public inspection as set forth in 37 CFR 1.11, Le Application No	der the Peperson's Reduction Act of 1905, no persons are required	Palent and 1 Is respond to a collection	Approved for use through 10/31/99. CMB DS Trademark Office; U.S. DEPARTMENT OF COM n of information unloss & displays a valid CMB control	MERCE
Sh attuck - E: devis         Application Number       Fied         08/409305       3-24-95         Group Art Unit       Exeminer         Paper No.       1/1         (A) referred to in United States Patent Number       5693.4773         (B) referred to in an application that is open to public inspection as set forth in 37 CFR 1.11, i.e         Application No.       0         (C) an application that claims the benefit of the filing date of an application that is open to public inspection, i.e., Application No.         (D) an application in which the applicant has filed an authorization to lay open the complete application to the public.         Please direct any correspondence concerning this request to the following address:	REQUEST FOR ACCESS OF ABANI	DONED APPLI	CATION UNDER 37 CFR 1.14	<b>(a)</b>
O8/ 409 305       3-24-95         Group Art Unit       Exeminer         Paper No.         Assistant Commissioner for Patents         Washington, DC 20231         I hereby request access under/37 CFR 1.14(a)(3)(iv) to the application file record of the above- identified ABANDONED application, which is: (CHECK ONE)		1	•	
O8/ 409 305       3-24-95         Group Art Unit       Exeminer         Paper No.         Assistant Commissioner for Patents         Washington, DC 20231         I hereby request access under/37 CFR 1.14(a)(3)(iv) to the application file record of the above- identified ABANDONED application, which is: (CHECK ONE)		Short	Hale F. L.	
O8/ 409 305       3-24-95         Group Art Unit       Exeminer         Paper No.         Assistant Commissioner for Patents         Washington, DC 20231         I hereby request access under/37 CFR 1.14(a)(3)(iv) to the application file record of the above- identified ABANDONED application, which is: (CHECK ONE)		Application Numb	ber Filed	
Assistant Commissioner for Patents Washington, DC 20231 I hereby request access under 37 CFR 1.14(a)(3)(iv) to the application file record of the above- kientified ABANDONED application, which is: (CHECK ONE)		08/		
Assistant Commissioner for Patents         Washington, DC 20231         I hereby request access under 37 CFR 1.14(a)(3)(iv) to the application file record of the above- identified ABANDONED application, which is: (CHECK ONE)        (A) referred to in United States Patent Number		Group Art Unit	Examiner	,
Assistant Commissioner for Patents         Washington, DC 20231         I hereby request access under 37 CFR 1.14(a)(3)(iv) to the application file record of the above- identified ABANDONED application, which is: (CHECK ONE)        (A) referred to in United States Patent Number				
Assistant Commissioner for Patents         Washington, DC 20231         I hereby request access under 37 CFR 1.14(a)(3)(iv) to the application file record of the above- identified ABANDONED application, which is: (CHECK ONE)        (A) referred to in United States Patent Number				
Washington, DC 20231         I hereby request access underi37 CFR 1.14(a)(3)(iv) to the application file record of the above- ktertified ABANDONED application, which is: (CHECK ONE)        (A) referred to in United States Patent Number	Assistant Commissioner for Patents		Paper No	—
Identified ABANDONED application, which is: (CHECK ONE)        (A) referred to in United States Patent Number				
dentified ABANDONED application, which is: (CHECK ONE)				
Identified ABANDONED application, which is: (CHECK ONE)        (A) referred to in United States Patent NumberStates Patent Number,        (B) referred to in an application that is open to public Inspection as set forth in 37 CFR 1.11, i.e.        (B) referred to in an application that is open to public Inspection as set forth in 37 CFR 1.11, i.e.        (B) referred to in an application that is open to public Inspection as set forth in 37 CFR 1.11, i.e.        (C) an application that claims the benefit of the filing date of an application that is open to public inspection, i.e., Application No, filed, or page, or page	hereby mayor screes under 37 CEP 4 14/	a)(?)(iid to the an	nication file means of the shown	
(B) referred to in an application that is open to public inspection as set forth in 37 CFR 1.11, i.e         Application No	kientified ABANDONED application, which is	: (CHECK ONE)	becanou the lecold of the above-	
(B) referred to in an application that is open to public inspection as set forth in 37 CFR 1.11, i.e Application No	(A) referred to in United States Patent N	umber <u>569</u>	<u>13473</u> , column	.,
Application No		•		
(C) an application that claims the benefit of the filing date of an application that is open to public inspection, i.e., Application No, filed, or, filed, filed, or, filed, filed, or, filed, f	Application No.			
inspection, i.e., Application No, filed, of (D) an application in which the applicant has filed an authorization to lay open the complete application to the public. Please direct any correspondence concerning this request to the following address: 	paper number,			
(D) an application in which the applicant has filed an authorization to lay open the complete application to the public.  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning this request to the following address:  Please direct any correspondence concerning the following address:  Please d				
application to the public. Please direct any correspondence concerning this request to the following address:				4
Please direct any correspondence concerning this request to the following address:		has filed an autho	orization to lay open the complete	
Culture     12-K-MACESSED       Signature     Date       Calvin Perry     5       Typed or printed name     FOR PTO USE CNILY       Approved by:     (initials)		a this manast to t		
Signature     190       Calvin Perry     FOR PTO USE ONLY       Typed or printed name     Approved by:       (initials)	riease direct any correspondence concernan	ig one request to t	the tollowing address:	
Signature     190       Calvin Perry     FOR PTO USE ONLY       Typed or printed name     Approved by:       (initials)				-
Signature     190       Calvin Perry     FOR PTO USE ONLY       Typed or printed name     Approved by:       (initials)				. ·
Signature     Signature       Calvin Perry     190       Typed or printed name     FOR PTO USE ONLY       Approved by:     (initials)				
Signature     Signature       Calvin Perry     190       Typed or printed name     FOR PTO USE ONLY       Approved by:     (initials)			13-	-
Signature     190       Calvin Perry     FOR PTO USE ONLY       Typed or printed name     Approved by:       (initials)	_ Calvi Kems		12-8-9400FF0-	
Approved by:	Signature		Date DED - OUED	RU
Approved by:	_ Calvin Kerry	_	× 1907	$\neq$
(initials)	Typed or printed name	· .	FOR PTO USE ONLY	
l Unit:			Unit:	
	en Hour Statement: This form is estimated to take 0.2 h	ours to complete. Tim		
en Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the in- . Any commants on the amount of time you are required to complete this form should be sent to the Chief Information Officer Trademark Office, Weshington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO stant Commissioner for Patents, Weshington, DC 20231. GeneDX 1010, pg.	Any comments on the amount of time you are required Trademark Office, Washington, DC 20231. DO NOT SEI stant Commissioner for Patients. Weshington, DC 20231	to complete this form NO FEES OR COMPLI	should be sain to the Chief Information Office ETED FORMS TO THIS ADDRESS, SEND TO COMPUTE 1010	<b>r, Pate</b> D: 229

er the Papaneonic Reduction A. (SDS, no paraons are rec	A and for use through 10/31/99. OMB 0651 Palent and Trade. 2 Office; U.S. DEPARTMENT OF COMM plind is respend to a collection of information unloss it displays a valid CMB control m
REQUEST FOR ACCESS OF ABA	NDONED APPLICATION UNDER 37 CFR 1.14
	In re Application of
	Shattuck - Eidens et.a
	Application Number Filed
	409,305 3/24/95
· · · ·	Group Art Unit Examiner 1807 Rel S
	Paper No. 12
Assistant Commissioner for Patents	
Washington, DC 20231	
(B) referred to in an application that i	It Number <u>5,693,473</u> , column, s open to public inspection as set forth in 37 CFR 1.11, i.e. , filed on page of
paper number	
(C) an application that claims the ber	nefit of the filing date of an application that is open to public
inspection, I.e., Application No	, filed, or
(D) an application in which the applic application to the public.	ant has filed an authorization to lay open the complete
Please direct any correspondence conce	ming this request to the following address:
S C Shuff	12/23/97
<u>S.</u> C. Shroff	Date
Signature SARLA_C.SHROFF	
Typed or printed name	FOR PTOLEGANTED DV
Abor of builde lighte	Approved by
	((Ingens)
	Unit: <b>CIAA</b>

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 sant to the Chief Information Officer, 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.

. ... .

restruction Act of 1925, nor parsons are required to respond to a collection of intermetion unwest it deputyee while CMB control number. REQUEST F RACCESS OF ABANDONED AFTLICATION UNDER 37 CFR 1.14(a) In re Application of D. SHATTUCK - GEDENS Application Number Fied 09 20 Group Art Unit Examines Paper No. Assistant Commissioner for Patents Washington, DC 20231 I hereby request access under 37 CFR 1.14(a)(3)(iv) to the application file record of the above-Identified ABANDONED application, which is: (CHECK ONE) , column (B) referred to in an application that is open to public inspection as set forth in 37 CFR 1.11, i.e., Application No. _ filed __ _____, on page _____ of paper number (C) an application that claims the benefit of the filing date of an application that is open to public inspection, i.e., Application No. _, filed ____ . 01 _ (D) an application in which the applicant has filed an authorization to lay open the complete application to the public. Please direct any correspondence concerning this request to the following address: -20 Date Signature BERGERT DUTTON - STEVE TROFFICI FOR PTO USE ONLY Typed or printed name Appria Diality Unit:

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 reduced to complete this form should be sent to the chief liferination officer, Patent and Trademats Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS Address. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

U. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a	Approved for use through 9/30/2007. OMB 0651-00 S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMER collection of information unless it displays a valid OMB control numb
EQUEST FOR ACCESS TO AN ABANDONED AP	PLICATION UNDER 37 CFR 1.14
In re Application of In re Application of Application Number OCrystal Drive Ington, VA 22202-3514 Pephone: (703) 308-2733 Application Number OS Constant Price Constant Price Constant OS	ued Prosecution Application (CPA) (37 CFR 1.53(d
United States Patent Application Publication No.	, page, line,
United States Patent Number 569347, Bolumn	, line, or
WIPO Pub. No, page, line	
<ul> <li>For published applications that are still pending, a member of the file contents; the pending application as originally filed; c application.</li> <li>For unpublished applications that are still pending: <ul> <li>(1) If the benefit of the pending application is claimed under 35 application that has:</li> <li>(a) issued as a U.S. patent, or</li> <li>(b) publication publication, or an international patent application is claimed or application is incorporated by reference or otherwise registration, a U.S. patent application, or an internation, or an internation, or an internation or an internation or an international patent application is incorporated by reference or otherwise registration, a U.S. patent application publication, or an internation, a U.S. patent application publication, or an internation, or an internation, or an internation, a U.S. patent application publication, or an internation, or an internation, and the public originally filed.</li> </ul></li></ul>	U.S.C. 119(e), 120, 121, or 365 in another ished as a statutory invention registration, a U.S. plication publication in accordance with PCT the file contents; the pending application as oplication. <u>identified</u> in a U.S. patent, a statutory invention mational patent application publication in
Azer Copy	10/30/07
Azer Cspy Azer Cspy	10/30/07 Date
Azie Coffy Azie b Esfery Typed or printed name	TO 3007 Date FOR PTO USE ONLY
A zie b Testery Typed or printed name	FOR PTO USE ONLY OCT 3 0 2007
Azieb Testery	FOR PTO USE ONLY OCT 3 0 2007 Approved by: (initials)
A zie b Testery Typed or printed name	FOR PTO USE ONLY OCT 3 0 2007

U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES FORMS TO THIS ADDRESS. BRING TO: File Information Unit, Room 2E04, 2900 Crystal Drive, Artington, Virginia.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

08 409305

PATENT APPLICATION SERIAL NO.

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

PTO-1556 (5/87)

ľ	PATENT			DETERMINAT ber 1, 1995	TION RECO		pplication of		ket Numbe 5305	
		and the second s	S FILED -		lumn 2)	SMALL	ENTITY	OR		
FOF	ł		ER FILED	NUMBER		RATE	FEE		RATE	FEE
BAS	IC FEE	Sign Sy		RE: 12.3			375.00	OR		750.0
гот	AL CLAIMS		18 minu	s 20 = *		x\$11=		OR	x\$22=	
NDE	PENDENT CL	AIMS	2 min	us 3 = *		x39=		OR	x78=	
MUL	TIPLE DEPEND	DENT CLAIM PRE	SENT			+125=		OR	+250=	
. 11 11	he difference in co	blumn 1 is less than	zero, enter "0"	in column 2		TOTAL		OR	TOTAL	750
ł		CLAIMS AS (Column 1)	AMENDE	<b>) - PART II</b> (Column 2)	(Column 3)	SMALL	ENTITY	OR		R THAN ENTITY
AMENDMENT A		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE	ADDI- TIONAL FEE		RATE	ADDI- TIONA FEE
MON	Total	•	Minus	**	s	x\$11=	1000	OR	x\$22=	
ME	Independent		Minus	***	=	x39=	1.521	OR	x78=	
∢	FIRST PRES	SENTATION OF	MULTIPLE	DEPENDENT CL	AIM	+125=	2	OR	+250=	
		(Column 1)		(Column 2)	(Column 3)	TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	
ENT B		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE	ADDI- TIONAL FEE		RATE	ADDI- TIONA FEE
NDMEN	Total	•	Minus	**	=	x\$11=		OR	x\$22=	
A MA	Independent		Minus	***	=	x39=		OR	x78=	
∢	FIRST PRE	SENTATION OF	MULTIPLE	DEPENDENT CL	AIM	+125=		OR	+250=	1
		(Column 1)		(Column 2)	(Column 3)	TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	
ENTC		CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE	ADDI- TIONAL FEE		RATE	ADDI- TIONA FEE
AMENDMENT	Total	•	Minus	**	=	x\$11=		OR	x\$22=	
MEN	Independent	•	Minus	***	=	x39=		OR	x78=	
4	FIRST PRE	SENTATION OF	MULTIPLE	DEPENDENT CL	AIM	+125=		OR	+250=	1
** !!	the entry in colu the "Highest Nur	mn 1 is less than the	ne entry in colu	umn 2, write "0" in col S SPACE is less than	lumn 3.	TOTAL ADDIT. FEE	(	OR	TOTAL ADDIT. FEE	

Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE



1

 $\mathbb{R}^{n}$ 

POSITION	ID NO.	DATE
CLASSIFIER	21	4/7/95
EXAMINER	405	8-8-95
TYPIST	17	3-4-91
VERIFIER	mes	3-61
CORPS CORR.		t
SPEC. HAND	412	2-27-96
FILE MAINT.		
DRAFTING		

## **INDEX OF CLAIMS**

Claim       Date         Image: Im				 		•								
Image: Stress of the second	Clair	n 📃		 	Dat	е								
SYMBOLS           3	Final	E 1 N	2/1/1											·
2         3         4         5         6         6         6         6         6         6         6         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7         7		<del>.    .</del>	74	 					-1					
3		<b>A</b>  -;		 	· · ·									
4				 										
5				 										
6				 										
7				 										
8				 			L							
9														
10			$\downarrow$	 										
11		_	11	 										
12			11											
13				 										
14			Ш	 										
14       SYMBULS         16       SYMBULS         21       SYMBULS         22       SYMBULS         23       SYMBULS         24       Standard         25       Standard         26       Standard         27       Standard         28       Standard         29       Standard         30       Standard         31       Standard         32       Standard         33       Standard         34       Standard         35       Standard         36       Standard         37       Standard         38       Standard         40       Standard         44       Standard         44       Standard         44       Standard         44       Standard         48       Standard	1	3		 										
16														
17			$\parallel$											
18       19         20       11         21       11         22       11         23       11         24       11         25       11         26       11         27       11         28       11         29       11         30       11         32       11         33       11         34       11         35       11         36       11         37       11         38       11         40       11         42       11         43       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11         44       11			11											
19														
20       21       22         21       22       23         22       23       24         25       26       27         28       29       28         30       29       20         31       23       20         32       23       20         33       29       20         30       29       20         33       29       20         33       29       20         33       20       20         33       20       20         33       20       20         33       20       20         34       20       20         35       20       20         36       20       20         37       20       20         40       20       20         41       20       20         42       20       20         43       20       20         44       20       20         44       20       20         44       20       20         48       20       20		8	11											
21	1	9												
22       23       24         24       25       26         26       27       28         29       29       28         30       29       28         31       29       28         32       31       28         33       29       20         31       29       20         32       31       20         33       20       20         34       20       20         35       20       20         36       20       20         38       20       20         40       20       20         41       20       20         43       20       20         44       20       20         44       20       20         44       20       20         44       20       20         44       20       20         44       20       20         44       20       20         44       20       20         44       20       20         48       20       20	2	0		 								÷		
23       24         24       25         26       27         28       29         30       31         31       32         32       33         33       34         35       35         36       37         38       39         40       41         42       43         44       44         45       46         47       48         49       49	2	1	-											
24	2	2	1											
25	2	3												
26       27       28         29       29       29         30       30       30         31       31       31         32       32       31         33       32       31         33       32       31         33       33       32         34       34       35         35       36       36         37       38       39         40       33       34         41       32       33         44       44       44         45       46       47         48       49       49	2	4	Τ											
27       28       30       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       31       32       32       31       31       31       31       32       32       32       32       32       32       31       31       31       32       33       31       31       32       33       31       31       31       32       33       32       33       33       33       32       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       33       34 <td< th=""><th>2</th><th>5</th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th><th></th></td<>	2	5												
28	2	26												
29	2	27												
30       SYMBOLS         31	2	28				•								
30				 			<u> </u>							
31     Allowed       32     Allowed       33     Allowed       33     Allowed       33     Allowed       33     Allowed       33     Allowed       34     Allowed       35     Allowed       36     Allowed       37     Allowed       38     Allowed       39     Allowed       40     Allowed       41     Allowed       42     Allowed       43     Allowed       44     Allowed       45     Allowed       46     Allowed       48     Allowed       49     Allowed				 			İ				,	31		Rejected
34							ļ	ļ		) -	-			Allowed
34				 			L			-+	(IN)	ougn nu	moeral)	. Restricted
35     0       36     0       37     0       38     0       39     0       40     0       41     0       42     0       43     0       44     0       45     0       46     0       47     0       48     0			_	 			Í				l		•••••	Non-elected
35     0       36     0       37     0       38     0       39     0       40     0       41     0       42     0       43     0       44     0       45     0       46     0       47     0       48     0	3	4		 			ļ	ļ			۰۰۰۰۰ (			Appeal
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$				 			<b> </b>	ļ	ļ		)	•••••		. Ubjected
38				 			ļ							
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			$\downarrow$	 			<u> </u>							
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			$\rightarrow$	 			<b> </b>							
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$			_	 		L				ł				
42				 						ļ				
43				 				L	L					
44     45       45     46       47     48       49     49							ļ			ł				
45				 				<b> </b>						
46			_1				ļ	<b> </b>	ļ	]				
47				 			ļ							
48 49 49				 			ļ	<u> </u>	ļ	1				
49				 		ļ		ļ		ł				
				 	· ·	L	ļ	<u> </u>						
50 50								ļ	L	1				
	5	50		 		[·		L		].				

								·	
CI	aim				Dat	e			
_	Original								
Final	ig								
ш	ō								
	51								
	52								
	53								
	54								
	55								
	56								
_	57								
	58								
	59								
	60								
	61								
	62								
	63							h	
	64								
	65					ļ		l	
	66								
	67								
	69								
	70								
	71 '				2				
	72								
	73					·			
	74								
·	75				ļ		<b> </b>		
	76				L			I	
	77					•			
	78								
	79			[		<u> </u>		ļ	
	80						<u> </u>	<b> </b>	
			<del> </del>	<u>}</u>					ł
	81						ļ	ļ	<u> </u>
	82	<b> </b>			<u> </u>			Ļ	<b> </b>
	83			L		<u> </u>			<u> </u>
	84				L				
	85								<u> </u>
	86	1	1					1	1
		<u> </u>		<u> </u>	·		<u> </u>		<u> </u>
	87	· · ·					<del> </del>	<u> </u>	<u> </u>
	88		<u> </u>		<b> </b>		<b> </b>	ļ	
	89	·		ļ			L	L	L
	90								1
	91	<u> </u>	[			<u> </u>			
<u> </u>		<u> </u>	-		<u> </u>	<u> </u>	<u>├</u> ──	┨────	┼───
	92		ļ		<del> </del>	<u> </u>	<b> </b>		ļ
	93			<b> </b>		ļ	L	L	<b> </b>
	94								
	95				Ι		[		
	96	1		1	<u> </u>	1	ţ	1	<u> </u>
	97	t			t	<u> </u>	<u> </u>	<u> </u>	
				┣──	┣──	<u>}                                    </u>	<b> </b>	┨───	
	98		ļ	L		ļ	I		ļ
	99								
	100	1	1	1	1	1	1	<u> </u>	

GeneDX 1010, pg. 245

## (LEFT INSIDE)

		SEAR	CHED	
	Class	Sub.	Date	Exmr.
	936 435		6/19/96	
29			Y	

INTERFERENCE SEARCHED						
Class	Sub.	Date	Exmr.			

APS, SCISEARCH, MEDLINE, EMBASE, CANCERLIT, BIOGIS BIOBUSINESS, LIN, GENDOWK, CAPLUS, NLDB, PROMT, LIFESCI BIOBUSINESS, TOX LIT DEERES PNI, TOX LINE, EMBAL CIN,	Date	Exmr.
MEDLINE, EMBASE, CANCERLIT, BIOSIS BIOBUSINESS, LIN, GENDOWK, CAPLUS, NLDB, PROMT, LIPESCI BIOBUSINESS, TOX LIT DEENES PNI, TOX LINE,	6 19 12	DP
DISSABS, JICST- EPLUS, NTIS, BIOTECHABS, BIOTECHABS, PATOSNO; NPIDS, WPINDEX, CABA, PATOSER, CEABA, CEN; CONFSCI, DRUGNLIAIDSLINE, CHACS, PRUGU, MCOLSE GEORICHED VELOTED CHACS, PRUGU,		AP

(RIGHT OUTSIDE)

GeneDX 1010, pg. 246