HUMAN MUTATION 22:121–128 (2003)

RESEARCH ARTICLE

Mutational Analysis of the *BRCA1*-Interacting Genes *ZNF350/ZBRK1* and *BRIP1/BACH1* Among *BRCA1 and BRCA2*-Negative Probands From Breast-Ovarian Cancer Families and Among Early-Onset Breast Cancer Cases and Reference Individuals

Joni L. Rutter,[1] Amelia M. Smith,[1] Michael R. Dávila,[1] Alice J. Sigurdson,[2] Ruthann M. Giusti,[3] Marbin A. Pineda,[1] Michele M. Doody,[2] Margaret A. Tucker,[4] Mark H. Greene,[3] Jinghui Zhang,[1] and Jeffery P. Struewing[1]*

[1]Laboratory of Population Genetics, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland; [2]Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics; National Cancer Institute, Bethesda, Maryland; [3]Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics; National Cancer Institute, Bethesda, Maryland; [4]Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics; National Cancer Institute, Bethesda, Maryland

Communicated by Mark H. Paalman

Two potential breast cancer susceptibility genes, encoding the BRCA1-interacting proteins ZNF350 (or ZBRK1) and BRIP1 (or BACH1), have been identified in yeast two-hybrid screens. We sequenced these genes in probands from 21 families with potentially inherited breast/ovarian cancer, all of which were negative for *BRCA1/BRCA2* mutations. Families had at least one case of male breast cancer, two cases of ovarian cancer, or three or more cases of breast and ovarian cancer. In addition, 58 early-onset (before age 35) breast cancer cases and 30 reference individuals were analyzed. Of 17 variants detected in *ZBRK1*, a missense mutation Val524Ile was identified in the proband of one high-risk family, but no other family members were available for testing. Of 25 variants identified in *BRIP1*, in addition to four common silent or missense mutations, we identified Gln540Leu, a non-conservative amino acid change, in a single familial proband with inflammatory breast cancer, but this mutation was not present in her three relatives with breast cancer. Haplotype analysis suggests that all *ZBRK1* SNPs fall within a single block with two SNPs capturing 92% of the haplotype diversity, while the *BRIP1* SNPs fall in two blocks, with five SNPs capturing 89% of the haplotype diversity. Based on sequencing of *ZBRK1* and *BRIP1* in 21 *BRCA1/2*-negative probands from inherited breast/ovarian cancer families, it appears unlikely that mutations in these genes account for a significant fraction of inherited breast cancer. Further analysis in unselected cases will be required to know whether the identified variants play a role in genetic predisposition to breast cancer in the general population. Hum Mutat 22:121–128, 2003. Published 2003 Wiley-Liss, Inc.[†]

KEY WORDS: cancer; breast cancer; ovarian cancer; DHPLC; haplotype; linkage disequilibrium; ZNF350; ZBRK1; BRIP1; BACH1; BRCA1; BRCA2; SNP

DATABASES:
BRIP1 – OMIM: 605882; GenBank: AF360549, NM_032043.1 (mRNA), NT_010783.13 (Chr 17 genomic contig)
ZNF350 – OMIM: 605422; GenBank: AF295096, NM_021632.1 (mRNA), AC011460.3 (genomic)
http://lpg.nci.nih.gov/LPG/struewing/pubs (Struewing Lab – additional data available)

INTRODUCTION

Among clearly inherited breast and ovarian cancer families (those with five or more early-onset cases of breast cancer, ovarian cancer, or male breast cancer), approximately two-thirds have germline mutations in the *BRCA1* (MIM# 113705) or *BRCA2* (MIM# 600185)

Received 7 January 2003; accepted revised manuscript 10 March 2003.

*Correspondence to: Jeffery P Struewing, NIH, LPG, 41 Library Dr., Room D702, Bethesda, MD 20892-5060. E-mail: Js140a@nih.gov

DOI 10.1002/humu.10238
Published online in Wiley InterScience (www.interscience.wiley.com).

genes that are detectable using PCR-based techniques [Szabo and King, 1997]. While some additional families may have abnormalities in *BRCA1* or *BRCA2*, such as large genomic rearrangements or regulatory mutations, a sizeable fraction appear to be unexplained by mutations in these two genes. Linkage analyses of large collections of families without *BRCA1/2* mutations are ongoing, but it does not appear that mutations in a single gene will account for a large proportion of the remaining families.

The *BRCA1* and *BRCA2* genes encode large, multifunctional proteins that do not share significant homology to other well-characterized proteins or to each other [Venkitaraman, 2002]. Biochemical studies have identified a large number of interacting proteins and upstream and downstream targets, many relating to DNA damage repair or transcriptional regulation. Two novel proteins that interact with BRCA1 were recently identified in yeast two-hybrid assays [Zheng et al., 2000; Cantor et al., 2001]. We studied whether mutations in the genes *ZNF350* (originally termed *ZBRK1*; MIM# 605422) and *BRIP1* (originally termed *BACH1*; MIM# 605882) that encode these BRCA1-interacting proteins might account for some proportion of *BRCA1/2*-negative probands from potentially inherited breast/ovarian cancer families. In addition, we screened for variants in an anonymized series of early-onset breast cancer cases and cancer-free reference individuals to identify variants that could be studied as lower-penetrance susceptibility alleles in other studies. We also studied linkage disequilibrium and haplotypes derived from the single nucleotide polymorphism (SNP) discovery and genotyping results emanating from this work.

MATERIALS AND METHODS

BRCA1/BRCA2-Negative Probands

Within the NCI's Family Cancer Registry, after excluding the 31 families with identified *BRCA1* or *BRCA2* mutations, there were 21 probands with breast or ovarian cancer who had no detectable deleterious mutations upon full sequencing of the *BRCA1* and *BRCA2* genes. The *BRCA1/2*-negative families had slightly fewer ovarian cancers (average 1.0 per family) compared to mutation positive families (2.8 per family), but similar numbers of breast cancer cases (3.1 vs. 3.3) [Struewing et al., 1995]. Each *BRCA1/2*-negative family in the present study had at least one case of male breast cancer, two cases of ovarian cancer, or three or more cases of breast and ovarian cancer. All families were of European ancestry.

Early-Onset Breast Cancer Cases and Cancer-Free Reference Individuals

Among all 61 subjects diagnosed with breast cancer before age 35 from a national cohort of radiologic technologists [Mohan et al., 2002], limited *BRCA1* mutation testing was performed on samples rendered anonymous [Struewing et al., 1996] and the 58 negative cases were included in this analysis, having obtained a waiver of the requirement for full IRB review. The samples are linked only to limited epidemiologic information, and the ethnicity of individual subjects is not known, but most are of European ancestry. In addition, samples from 30 unrelated cancer-free individuals of European ancestry (subjects 01 and 02 from families 1328, 1331, 1332, 1333, 1341, 1344, 1346, 1347, 1349, 1362, 1400, 1408, 1413, 1416, 1423 from the *Centre d'Etude du Polymorphisme* (CEPH) family collection) were obtained from the Coriell Repository. Analysis of these 88 subjects was performed blinded as to case or reference status.

Sequence Analysis

The intron–exon boundaries of *ZBRK1* and *BRIP1* were determined by aligning GenBank mRNA records (AF295096 and AF360549, respectively) with genomic sequence records (NT_010757 and NT_011091, respectively). *ZBRK1* spans approximately 10 kb and is composed of four exons, while *BRIP1* spans approximately 180 kb and is composed of 20 exons. PCR amplicons were designed to amplify the entire mRNA encoding portions and flanking intronic sequence from genomic DNA, using nine primer pairs ranging in size from 208 bp to 428 bp for *ZBRK1* and 23 primer pairs ranging in size from 250 bp to 549 bp for *BRIP1*.

The 21 *BRCA1/2*-negative probands were analyzed by sequencing all PCR products with both the forward and reverse primers used for amplification. The sequencing reaction included two microliters of PCR product as template in an eight-microliter reaction using 20 pmoles of primer and Big Dye Terminator Ready Reaction Mix (Applied Biosystems, Foster City, CA) at 0.8 × concentration. The reactions were cleaned up using Microcon 96-well spin plates, dried down, resuspended in 10 microliters of deionized formamide and analyzed on an ABI 3100 sequencer (Applied Biosystems). Two reviewers inspected all sequence tracings, with the forward and reverse reads aligned using Vector NTI (Informax, Inc., North Bethesda, MD).

Analysis of the early-onset breast cancer cases and CEPH reference individuals included both denaturing high performance liquid chromatography (DHPLC) and sequencing. DHPLC analysis was performed on a WAVE System (Transgenomic, Inc., Omaha, NE) using the same PCR primers as for sequencing (gradients and DHPLC conditions are available at http://lpg.nci.nih.gov/LPG/struewing/pubs). For amplicons in which there were no polymorphisms, only DHPLC analysis was performed. For amplicons with one or more common polymorphisms, both DHPLC and sequencing in one direction were usually performed, because distinguishing among the several possible genotype combinations in the DHPLC chromatograms alone was not always possible. Variants observed in the DHPLC chromatograms were sequenced as above to determine the exact nucleotide change. STAR software (Varian, Inc., Walnut Creek, CA) was also used to group the DHPLC chromatograms into clusters based on their patterns as an aid in identifying samples requiring sequencing, or to determine a subject's genotype for common polymorphisms.

Further Accession Numbers and Supplementary Data

Genomic records used were AC011460.3 and NM_010783.13 and mRNA records used were NM_021632.1 and NM_032043.1 for *ZBRK1* and *BRIP1*, respectively. All variants were deposited into dbSNP with accession numbers ss6905442 to ss6905458 for *ZBRK1* and ss6905459 to ss6905483 for *BRIP1*. Primer sequences, PCR and DHPLC thermal cycling conditions, and the genetic data for each subject are available at http://lpg.nci.nih.gov/LPG/struewing/pubs.

Statistical Analysis

No formal statistical comparisons of variant allele frequencies between the early-onset breast cancer cases and reference individuals were made because the base populations for the two groups are likely to be dissimilar. Nucleotide diversity was calculated separately for the amino acid coding nucleotides (n = 1599 for *ZBRK1* and n = 3750 for *BRIP1*) and non-coding nucleotides (n = 1084 for *ZBRK1* and n = 5115 for *BRIP1*) that were analyzed [Hartl and Clark, 1997]. We used the publicly

available computer algorithms PHASE and SNPHAP to predict haplotypes, as implemented in the HapScope suite of programs [Johnson et al., 2001; Stephens et al., 2001; Zhang et al., 2002]. Haplotypes were estimated after excluding variants with allele frequencies below 1% (and before block boundaries were estimated), and the minimum set of variants required to distinguish all haplotypes with a frequency above 2% was calculated using a greedy partition algorithm as implemented in HapScope [Zhang et al., 2002]. The normalized measure of linkage disequilibrium, D′, and exact *P* values for allelic association, were calculated for all pair-wise combinations of markers using Arlequin Version 2.001 software (http://anthro.unige.ch/arlequin).

RESULTS

A total of 12 nucleotide variants were identified in *ZBRK1* and 11 variants were identified in *BRIP1* among the 21 *BRCA1/2*-negative probands (Tables 1 and 2 and Fig. 1). The *ZBRK1* mutations Q393Q (g.1362A>G) and V524I and a variant in the 3′UTR (c.1845C>T) were identified in one proband each, but no other affected family members were available to determine whether these variants segregated with cancer in the family. Neither the silent nor the missense mutations were observed in the 58 early-onset breast cancer cases or 30 reference individuals, but the 3′UTR variant was present in 7% and 8% of these subjects, respectively. The other variants observed in the *BRCA1/2*-negative probands had allele frequencies of 10% or greater, were observed in the 88 other individuals analyzed, and are unlikely to be related to the cancer clustering in these families (Table 1). A missense mutation Q540L and an intronic variant IVS14+26delT were identified in *BRIP1* in one proband each; neither was observed in the early-onset breast cancer cases nor in reference individuals. Another intronic variant, IVS5–31G>C, was identified in two probands but in many of the other 88 individuals. The other six variants identified in the high-risk family probands were common, and had allele frequencies of 25% or greater in the reference individuals (Table 2).

The glutamine to leucine change at residue 540 (Q540L) in *BRIP1* is a non-conservative change that does not fall within any of the helicase domains of this protein. Samples from additional family members were available for analysis, but neither the proband's daughter nor her two nieces who had breast cancer carried this mutation. No other family members were available for analysis of the *BRIP1* IVS14+26delT variant, nor for the one family in which the proband had the IVS5–31G>C. In the second family with the intron 5 variant, all three relatives with breast or ovarian cancer also carried this mutation. One other female over age 50 without cancer from this family was available for analysis, and she also carried the IVS5–31G>C.

Analysis of the early-onset breast cancer cases and the reference individuals revealed five additional *ZBRK1* variants and 14 additional *BRIP1* variants that were not observed in the *BRCA1/2*-negative probands (Tables 1 and 2). Of seven additional *BRIP1* missense mutations identified, four were identified in the 30 reference individuals and not in the early-onset breast cancer cases. Four breast cancer cases carried both the T385T(g.1338A>C) and S472P variants in *ZBRK1*, while none of the reference individuals did, but the frequencies of most variants were similar between the

TABLE 1. Results of Mutational Analysis of ZNF350/*ZBRK1*

Number[b]	Protein variant (aa codon)	DNA variant[c]	Segment	Genomic location[d]	Allele frequency (%)[a]		
					BRCA1/2 Negative Probands[e]	Breast Ca cases dx <35[f]	Reference group[f]
1	(D35)	288T>C	Exon 2	29248	7/42 (0.17)	26/116 (0.22)	16/60 (0.27)
2	L66P[g]	380T>C	Exon 3	29671	7/42 (0.17)	14/116 (0.12)	11/60 (0.18)
3		IVS3+18A>G[g]	Intron 3	29730	6/42 (0.14)	25/116 (0.22)	16/60 (0.27)
4		IVS3+46A>T[g]	Intron 3	29758	7/42 (0.17)	14/116 (0.12)	11/60 (02.18)
5		IVS3+62G>A[g]	Intron 3	29774	6/42 (0.14)	27/116 (0.23)	16/60 (0.27)
6	(C236)	891C>T	Exon 4	32545	6/42 (0.14)	14/116 (0.12)	10/60 (02.17)
7	(P373)[g]	1302C>A	Exon 4	32956	4/42 (0.10)	14/116 (0.12)	9/60 (0.15)
8	(T385)	1338A>C	Exon 4	32992	0/42	4/116 (0.03)	0/60
9	(Q393)[h]	1362A>G	Exon 4	33016	1/42 (0.02)	0/116	0/60
10	S472P	1597T>C	Exon 4	33251	0/42	4/116 (0.03)	0/60
11	(S476)	1611C>T	Exon 4	33265	0/42	0/116	1/60 (0.02)
12	R501S[g]	1686A>T	Exon 4	33340	4/42 (0.10)	14/116 (0.12)	9/60 (0.15)
13	V524I[h]	1753G>A	Exon 4	33407	1/42 (0.02)	0/114[i]	0/60
14		1845C>T[g] (73rd nt of 3′UTR)	3′UTR	33509	1/42 (0.02)	8/116 (0.07)	5/60 (0.08)
15		1947T>A (165th nt of 3′UTR)	3′UTR	33601	0/42	0/116	1/60 (0.02)
16		1955T>A (173rd nt of 3′UTR)	3′UTR	33609	0/42	0/116	1/60 (0.02)
17		2067C>T[g] (285th nt of 3′UTR)	3′UTR	33721	4/42 (0.10)	14/116 (0.12)	9/60 (0.15)

[a]Frequency of the second allele listed.
[b]SNP number for cross-referencing to Figure 2.
[c]Numbering based on RefSeq NM_021632.1 (AF295096), with most common allele listed first.
[d]Nucleotide position within AC011460.3.
[e]Based on complete sequencing.
[f]Breast cancer cases diagnosed before age 35 from radiologic technologist cohort; reference group consists of 30 CEPH parents.
[g]In dbSNP as of Nov. 1, 2002.
[h]Within *BRCA1/BRCA2*-negative proband families, no other affected subjects were available for analysis.
[i]One subject was untyped for this SNP.

TABLE 2. Results of Mutational Analysis of *BRIP1/BACH1*

Number[b]	Protein variant (aa codon)	DNA variant[c]	Segment	Genomic location[d]	Allele frequency (%)[a]		
					BRCA1/2 Negative probands[e]	Breast Ca cases dx ≤35[f]	Reference group[f]
1		−64G>A	–	15233769	14/42 (0.33)	64/116 (0.55)	24/60 (0.40)
2		12C>T	Exon 1	15233726	0/42	1/116 (0.01)	0/60
3		IVS1+12G>A	Intron 1	15233615	12/42 (0.29)	38/116 (0.33)	16/60 (0.27)
4		IVS1+22G>C	Intron 1	15233605	0/42	1/116 (0.01)	0/60
5		IVS2+72T>G	Intron 2	15231718	0/42	1/116 (0.01)	0/60
6		IVS4-28G>A	Intron 4	15219627	0/42	0/116	2/60 (0.03)
7		IVS5-31C>G	Intron 5	15217594	2/42 (0.05)	19/116 (0.16)	5/60 (0.08)
8	R173C	658C>T	Exon 6	15217554	0/42	0/116	1/60 (0.02)
9	V193I	718G>A	Exon 6	15217494	0/42	1/116 (0.01)	0/60
10	L195P	725T>C	Exon 6	15217487	0/42	1/116 (0.01)	0/60
11	R419W	1396C>T	Exon 9	15169567	0/42	0/116	1/60 (0.02)
12	F531V	1732T>G	Exon 11	15154650	0/42	0/116	1/60 (0.02)
13	Q540L[g]	1760A>T	Exon 11	15154622	1/42 (0.02)	0/116	0/60
14		IVS12-47C>G	Intron 12	15150791	14/40[i] (0.35)	36/116 (0.31)	10/58[i] (0.17)
15		IVS14+7G>A	Intron 14	15146737	0/42	0/116	1/60 (0.02)
16		IVS14+26delTh[h]	Intron 14	15146718	1/42 (0.02)	0/116	0/60
17		IVS16-139C>T	Intron 16	15086545	0/42	1/116 (0.01)	0/60
18	C832Y	2636G>A	Exon 18	15063853	0/42	0/116	1/60 (0.02)
19	(E879)	2778G>A	Exon 19	15056447	16/42 (0.38)	27/114i (0.24)	22/56[i] (0.39)
20	P919S	2896C>T	Exon 19	15056329	17/42 (0.40)	39/116 (0.34)	24/60 (0.40)
21	V935G	2945T>G	Exon 19	15056280	0/42	1/116 (0.01)	0/60
22		IVS19+83T>A	Intron 19	15056096	10/42(0.24)	41/116 (0.35)	27/60 (0.45)
23	(Y1137)	3552C>T	Exon 20	15053978	17/42 (0.40)	38/116 (0.33)	24/60 (0.40)
24		3968A>C (77th nt of 3′UTR)	3′UTR	15053562	0/42	0/116	1/60 (0.02)
25		4049C>T (158th nt of 3′UTR)	3′UTR	15053481	17/42 (0.40)	38/116 (0.33)	24/60 (0.40)

[a]Frequency of the second allele listed.
[b]SNP number for cross-referencing to Figure 2.
[c]Numbering based on RefSeq NM_032043.1 (AF360549), with most common allele listed first.
[d]Nucleotide position within NT_010783.13.
[e]Based on complete sequencing.
[f]Breast cancer cases diagnosed before age 35 from radiologic technologists cohort;reference group consists of 30 CEPH parents.
[g]Not present in any of proband's three affected relatives.
[h]Within *BRCA1/2*-negative proband families, no other affected subjects were available for analysis.
[i]One or two subjects were untyped for this SNP.

early-onset breast cancer cases and reference individuals (no formal statistical comparisons were made).

Our extensive genetic analysis of these two genes in 218 chromosomes allowed us to examine diversity and linkage disequilibrium measures. The calculated amino acid encoding nucleotide diversity was approximately twice as high for *ZBRK1* (8.6×10^{-4}) compared with *BRIP1* (4.4×10^{-4}), while for non-coding regions the values were 1.2×10^{3} and 4.9×10^{-4}, respectively. The SNPHAP and PHASE programs resulted in nearly identical *ZBRK1* haplotype predictions for each of the 109 subjects, with only one difference when excluding variants at 1% allele frequency or below, and no differences when excluding variants at 10% frequency or lower. For *BRIP1*, in which there were no variants with allele frequencies between 1 and 10%, the predicted haplotypes differed for six subjects when variants below 1% allele frequency were excluded.

There were only three *ZBRK1* haplotypes with estimated frequencies of 2% or greater, with two variants, namely D35D (or alternatively IVS3+18A>G or IVS3+62G>A) and L66P (or alternatively IVS3+46A>T) capturing all of this variability. (Fig. 2) *BRIP1* showed much more haplotype diversity, with nine haplotypes at 2% frequency or greater, and no single haplotype had an estimated frequency above 25% (Fig. 2). Thus, five variants would need to be analyzed to capture the common haplotype diversity in this gene. By graphical inspection of the pair-wise D′ and *P* values for the *ZBRK1* variants that span approximately 10 kilobases, they appear to fall within a single haplotype block, while there appear to be two haplotype blocks across *BRIP1*, which spans approximately 180 kilobases (Figs. 2 and 3).

DISCUSSION

Approximately one-third of families with inherited forms of breast and ovarian cancer do not segregate mutations in *BRCA1* or *BRCA2*, two genes identified through linkage analysis in multiple-case families. Additional linkage studies in families that are negative for *BRCA1/2* suggest that mutations in no single gene account for a large proportion of families. If mutations in several genes acting in concert account for small subsets of these *BRCA1/2*-negative families, identifying such genes by linkage analyses and positional cloning will be difficult. An alternative approach is to analyze candidate genes to search for mutations that alter the normal biological function of the gene or its product and that segregate with cancer in a family. Some breast cancer

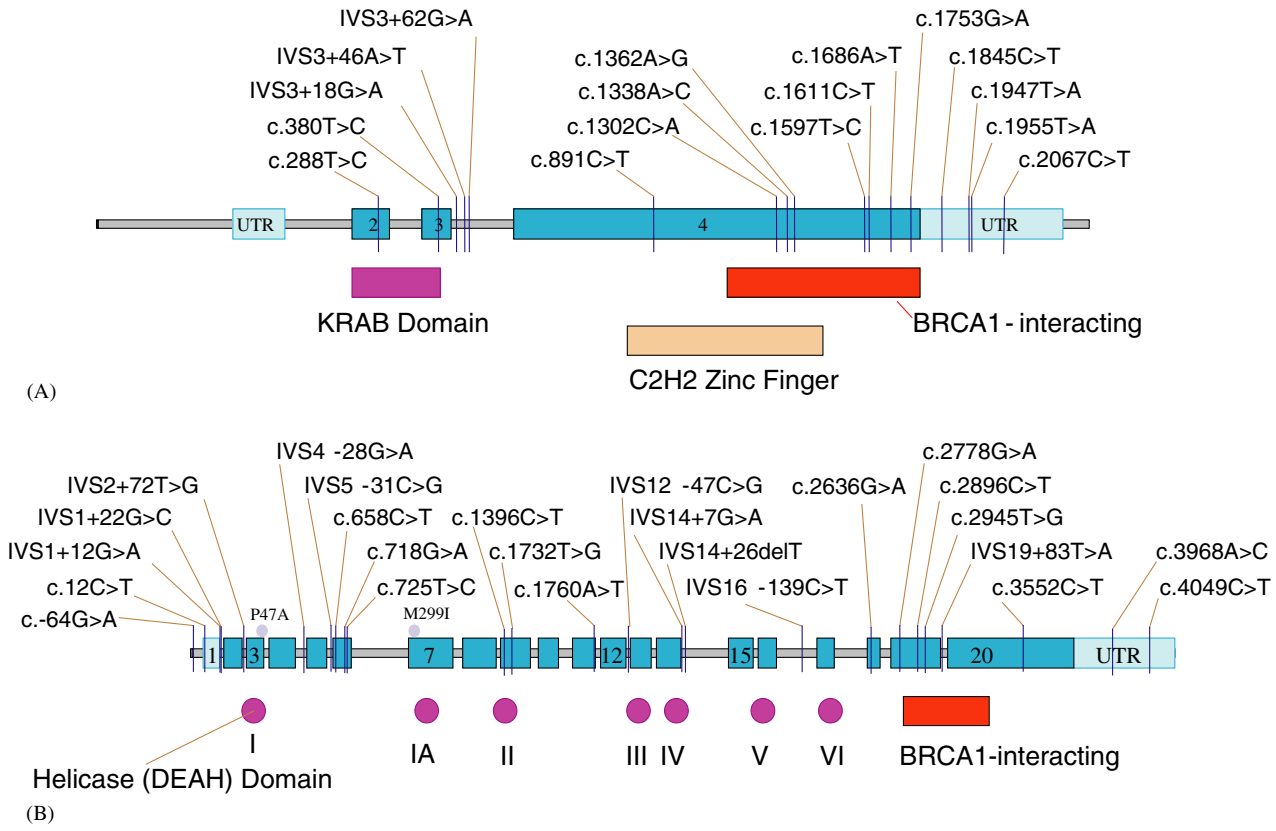

FIGURE 1. **Schematic diagram showing the genomic structure and position of identified genetic variants. Introns are shown as a thin horizontal line and are not drawn to scale; exons are shown as thick horizontal lines and are drawn roughly to scale, with selected ones numbered. A:** ***ZNF350/ZBRK1*** **gene showing four exons with 17 genetic variants. The KRAB, C2H2 zinc finger, and BRCA1-interacting domains are indicated by rectangles below the gene. B:** ***BRIP1/BACH1*** **gene showing 20 exons with 25 genetic variants. The DEAH helicase domains are shown as circles and the BRCA1-interacting domain as a rectangle below the gene. The two mutations identified previously in early-onset breast cancer cases [Cantor et al., 2001] are indicated with small circles above the gene. See Table 1 (*ZNF350*) and Table 2 (*BRIP1*) for the protein-based names of the exonic mutations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]**

families have been attributed to such mutations [Vahteristo et al., 2001; Chenevix-Trench et al., 2002].

The biological basis of the cancer predisposition in *BRCA1* and *BRCA2* mutation carriers has not been clearly elucidated. BRCA1 and BRCA2 are large, multifunctional proteins, involved in DNA damage sensing and repair, transcriptional regulation, and transcription-coupled repair, but whether these or other functions account for their carcinogenic effect are not known. They interact with many different proteins, including two that were identified through yeast two-hybrid analyses. The *ZBRK1* gene is a novel gene identified using an N-terminal fragment of *BRCA1* as bait [Zheng et al., 2000]. This gene encodes a protein with a KRAB domain and eight zinc fingers involved in repression of *GADD45* and, potentially, other targets. No mutational analysis has been published, but there are four SNPs in dbSNP that map to the mRNA. The *BRIP1* gene, originally termed *BACH1*, was identified using a C-terminal fragment of *BRCA1* as bait [Cantor et al., 2001]. *BRIP1* contains six DEAH-helicase homology regions and interacts with the BRCT domains of *BRCA1*. Mutations within the helicase domains of *BRIP1* interfered with double-strand DNA break repair in a BRCA1-binding dependent manner, and two germline missense mutations in *BRIP1* were identified among 65 breast cancer cases [Cantor et al., 2001]. One of these, P47A, identified in a patient with a strong family history of breast and ovarian cancer who was negative for *BRCA1/BRCA2* mutations, affected a conserved residue in the helicase domain and produced an unstable protein. No other family members from the two missense-carrying cases were available for analysis, but this finding does raise the possibility that mutations in *BRIP1* may account for some (presumably small) fraction of inherited breast cancer cases.

We screened the *ZBRK1* and *BRIP1* genes in a series of individuals at high risk of genetic forms of breast cancer, including 21 probands from families with multiple cases of breast and/or ovarian cancer who were negative for *BRCA1/2* mutations, and 58 individuals who were diagnosed with breast cancer before age 35 [Struewing et al., 1995, 1996; Mohan et al., 2002]. We did not identify any mutations that were clearly related to breast cancer. For *BRIP1* we did not observe the two previously reported mutations P47A and M299I, but we did identify

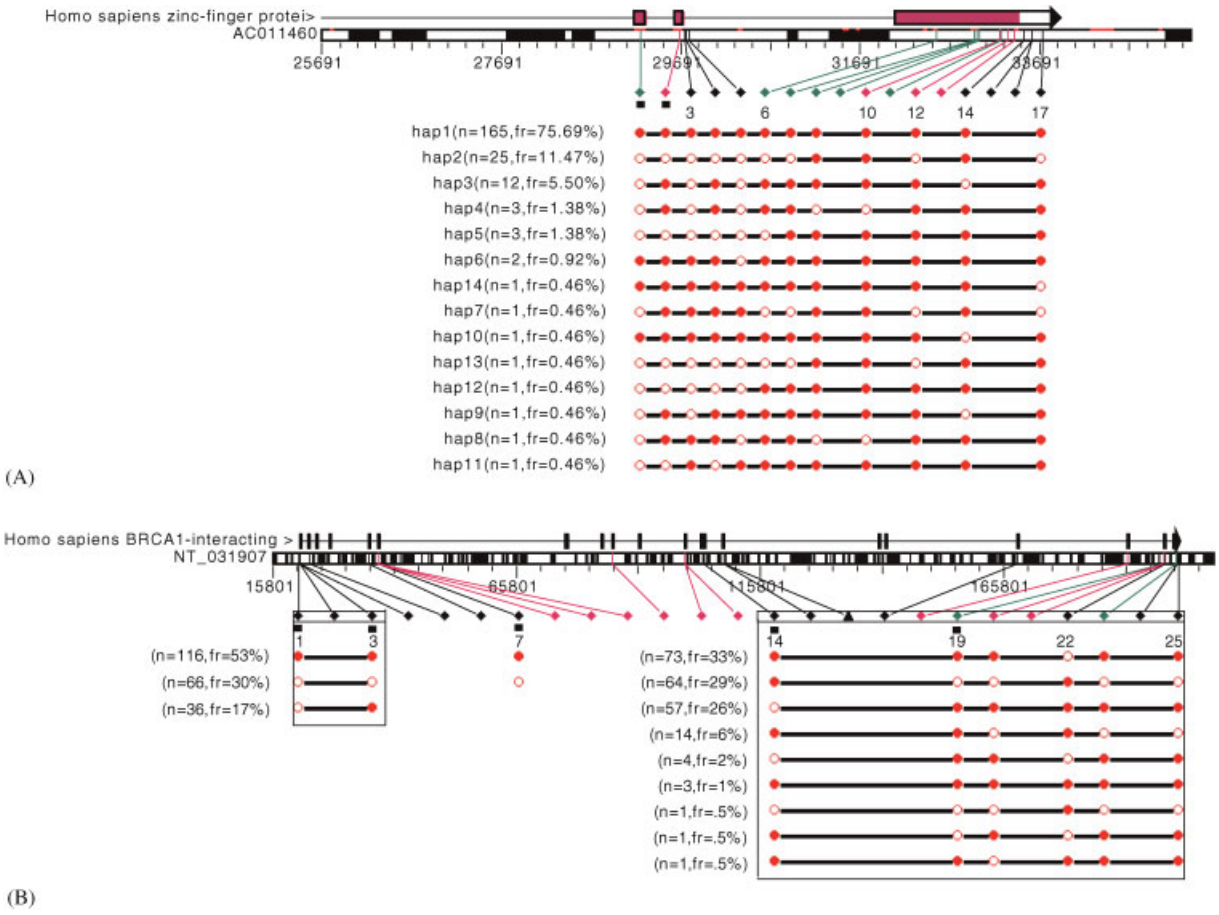

FIGURE 2. **Schematic representation of calculated haplotypes as displayed by HapScope, with minor modifications. The exons and introns are drawn to scale (different for each gene) at the top of the diagram as horizontal lines, with nucleotide positions of the genomic record indicated by the ruler. Repetitive elements are shown as blackened areas in the genomic record bar. Symbols (diamonds = nucleotide substitutions, triangles = insertion/deletion) below the ruler represent the SNPs and their relative location indicated by lines connecting to the ruler. Small squares below a SNP symbol indicate a minimum set required to identify all haplotypes with a frequency greater than 2%. Numbers immediately below some SNP symbols correspond to the variant numbers from Tables 1 and 2, for cross-reference). Haplotypes are shown in the lower half of the figure, with each SNP used to calculate the haplotypes represented by a circle; darkened circles represent the common allele; open circles represent the variant allele. The number of occurrences and percent frequency are shown to the left of each haplotype. A: *ZNF350/ZBRK1* showing the 12 SNPs with allele frequencies greater than 1% used to estimate haplotypes. D35D and P373P are indicated with small squares as the minimum SNPs required to identify the three haplotypes with a frequency greater than 2%. B: *BRIP1/BACH1* showing the nine SNPs with allele frequencies greater than 1% used to estimate haplotypes, and the two estimated haplotype blocks. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]**

a non-conservative missense mutation, Q540L, in a *BRCA1/2*-negative proband with breast cancer. Analysis of seven additional relatives, however, suggested that this mutation was not the underlying cause of the breast cancer in this family, as none of the three other relatives with breast cancer carried this mutation. The conservative *ZBRK1* missense mutation V524I was identified in a *BRCA1/2*-negative proband and in none of the early onset breast cancer cases or reference individuals, but no other family members were available for analysis and it is unclear whether this is related to the cancer occurrence. There were only two other *ZBRK1* or *BRIP1* variants observed exclusively in the *BRCA1/2*-negative probands; one was a silent mutation and the other an intronic variant that appear unlikely to be causative mutations. If we assume none of the 21 probands carries a disease-related mutation, the upper 95% confidence limit on this proportion (0/21) would be 13% [Hanley and Lippman-Hand, 1983].

We identified many sequence variants in the two genes, including two silent mutations in *ZBRK1* and two silent and one missense mutation in *BRIP1* that had allele frequencies of at least 10% but are not in dbSNP (as of September, 2002). The observed nucleotide diversity was about twice as high for *ZBRK1* as *BRIP1*, but in comparison with 15 other genes analyzed in 93 subjects of worldwide distribution, only four genes had higher amino-acid coding diversity than *BRIP1*, suggesting that neither gene is under strong functional constraint [Thorstenson et al., 2001]. Our linkage disequilibrium and haplotype analyses appear to be consistent with recent evidence suggesting that SNPs fall within blocks of extended disequilibrium [Rioux et al., 2001; Gabriel et al., 2002]. All *ZBRK1* variants were in

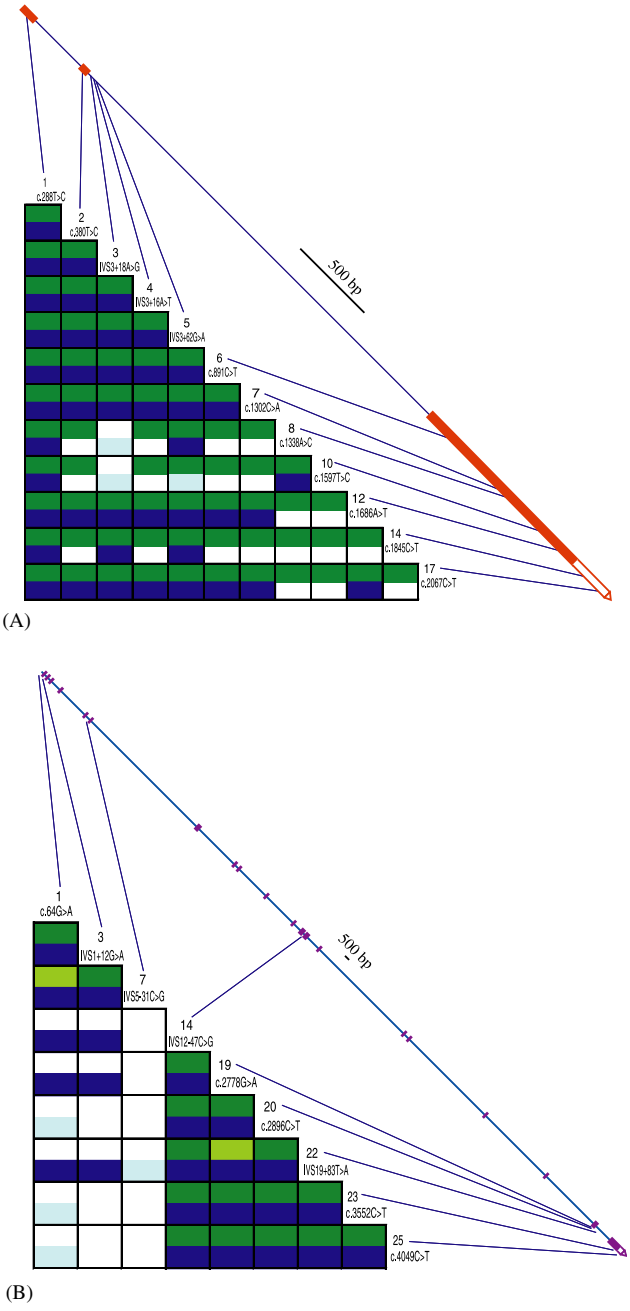

FIGURE 3. **Graphical representation of measures of linkage disequilibrium between SNPs with allele frequencies greater than 1%. The relative location of SNPs are indicated by lines connecting to the linear representation of the genomic structure of the gene, which are drawn to a different scale for each gene; a 500-bp line above the gene line indicates the scale. For each pair-wise comparison, the upper half of the square indicates the magnitude of D′, with dark shading representing values >0.9, light shading representing values between 0.8 and 0.9, and open rectangles representing values <0.8; the lower half represents the exact *P* value for the 2 × 2 association between SNPs, with dark shading representing values <0.001, light shading representing values between 0.05 and 0.001, and open rectangles representing values >0.05. A: *ZNF350/ZBRK1*. B: *BRIP1/BACH1*. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]**

strong linkage disequilibrium, and analyzing the D35D and L66P variants would capture the common haplotype diversity across this gene, which appears to fall within a single block. While less diverse at the nucleotide level, *BRIP1* showed considerably more haplotype diversity, with the variants in *BRIP1* falling within two haplotype blocks.

Our SNP discovery effort in the early-onset breast cancer cases and reference individuals identified several common variants that may warrant further study in breast cancer or other cancer case-control studies. Within *ZBRK1*, the two non-conservative missense mutations L66P and R501S are common (allele frequencies of 10% or greater). Another non-conservative mutation, S472P, was present in 7% of early-onset breast cancer cases but none of the reference individuals. Two *BRIP1* variants, the non-conservative missense mutation (P919S) that is very common (allele frequency >30%), and a newly described, common variant 64 nucleotides upstream from the transcription start site (c.–64G>A) that might affect gene regulation, may warrant further investigation as susceptibility alleles.

In summary, we performed mutational analysis of the genes *ZBRK1* and *BRIP1*, encoding two recently identified BRCA1-interacting proteins, as possible explanations for inherited breast and ovarian cancer clustering in families that are negative for germline mutations in *BRCA1* and *BRCA2*. We did not identify any mutations that appear to be causative in these families and it appears unlikely that mutations in these genes account for a large fraction of inherited forms of breast cancer. Additional analysis in early-onset breast cancer cases and reference individuals identified several newly described variants that make attractive candidates as susceptibility alleles for breast and other cancers.

REFERENCES

Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA, Livingston DM. 2001. Bach1, a novel helicase-like protein, interacts directly with brca1 and contributes to its DNA repair function. Cell 105:149–160.

Chenevix-Trench G, Spurdle AB, Gatei M, Kelly H, Marsh A, Chen X, Donn K, Cummings M, Nyholt D, Jenkins MA, Scott C, Pupo GM, Dork T, Bendix R, Kirk J, Tucker K, McCredie MR, Hopper JL, Sambrook J, Mann GJ, Khanna KK. 2002. Dominant negative ATM mutations in breast cancer families. J Natl Cancer Inst 94:205–215.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. 2002. The structure of haplotype blocks in the human genome. Science 296:2225–2229.

Hanley JA, Lippman-Hand A. 1983. If nothing goes wrong, is everything all right? Interpreting zero numerators. JAMA 249:1743–1745.

Hartl DL, Clark AG. 1997. Principles of population genetics. Sunderland, MA: Sinauer Associates, Inc. p. 345–349.

Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG, Todd

JA. 2001. Haplotype tagging for the identification of common disease genes. Nat Genet 29:233–237.

Mohan AK, Hauptmann M, Linet MS, Ron E, Lubin JH, Freedman DM, Alexander BH, Boice JD Jr, Doody MM, Matanoski GM. 2002. Breast cancer mortality among female radiologic technologists in the United States. J Natl Cancer Inst 94:943–948.

Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z, Delmonte T, Kocher K, Miller K, Guschwan S, Kulbokas EJ, O'Leary S, Winchester E, Dewar K, Green T, Stone V, Chow C, Cohen A, Langelier D, Lapointe G, Gaudet D, Faith J, Branco N, Bull SB, McLeod RS, Griffiths AM, Bitton A, Greenberg GR, Lander ES, Siminovitch KA, Hudson TJ. 2001. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet 29:223–228.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989.

Struewing JP, Brody LC, Erdos MR, Kase RG, Giambarresi TR, Smith SA, Collins FS, Tucker MA. 1995. Detection of eight *BRCA1* mutations in 10 breast/ovarian cancer families, including 1 family with male breast cancer. Am J Hum Genet 57:1–7.

Struewing JP, Tarone RE, Brody LC, Li FP, Boice JD Jr. 1996. *BRCA1* mutations in young women with breast cancer. Lancet 347:1493.

Szabo CI, King MC. 1997. Population genetics of BRCA1 and BRCA2. Am J Hum Genet 60:1013–1020.

Thorstenson YR, Shen P, Tusher VG, Wayne TL, Davis RW, Chu G, Oefner PJ. 2001. Global analysis of ATM polymorphism reveals significant functional constraint. Am J Hum Genet 69:396–412.

Vahteristo P, Tamminen A, Karvinen P, Eerola H, Eklund C, Aaltonen LA, Blomqvist C, Aittomaki K, Nevanlinna H. 2001. p53, CHK2, and CHK1 genes in Finnish families with Li-Fraumeni syndrome: Further evidence of CHK2 in inherited cancer predisposition. Cancer Research 61:5718–5722.

Venkitaraman AR. 2002. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 108:171–182.

Zhang J, Rowe WL, Struewing JP, Buetow KH. 2002. HapScope: a software system for automated and visual analysis of functionally annotated haplotypes. Nucleic Acids Res 30:5213–5221.

Zheng L, Pan H, Li S, Flesken-Nikitin A, Chen LP, Boyer GT, Lee HW. 2000. Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1. Mol Cell 6:757–768.