

## Clinical Classification of *BRCA1* and *BRCA2* DNA Sequence Variants: The Value of Cytokeratin Profiles and Evolutionary Analysis—A Report From the kConFab Investigators

Amanda B. Spurdle, Sunil R. Lakhani, Sue Healey, Suzanne Parry, Leonard M. Da Silva, Ross Brinkworth, John L. Hopper, Melissa A. Brown, Davit Babikyan, Georgia Chenevix-Trench, Sean V. Tavtigian, and David E. Goldgar

### ABSTRACT

#### Purpose

Rare missense substitutions and in-frame deletions of *BRCA1* and *BRCA2* genes present a challenge for genetic counseling of individuals carrying such unclassified variants. We assessed the value of tumor immunohistochemical markers in conjunction with genetic and evolutionary approaches for investigating the clinical significance of unclassified variants.

#### Patients and Methods

We studied 10 *BRCA1* and 12 *BRCA2* variants identified in Australian families with breast cancer. Analyses assumed a prior probability based on revised cross-species sequence alignment methods assessing amino acid evolutionary conservation and position, combined with likelihoods from data on co-occurrence with pathogenic mutations in the same gene, segregation analysis, and immunohistochemistry. We specifically explored the value of estrogen receptor, cytokeratin 5/6, and cytokeratin 14 as tumor markers of *BRCA1* mutation status.

#### Results

Posterior probabilities classified 72% of variants. *BRCA1* variants IVS18+1 G>T (del exon 18) and 5632 T >A (V1838E) were classified as pathogenic, with >99% posterior probability of being deleterious, and tumor histopathology was particularly important for their classification. *BRCA2* variant classification was improved over previous studies, largely by incorporating the prior probability of pathogenicity based on amino acid cross-species sequence alignments.

#### Conclusion

Variant classification was considerably improved by analysis of estrogen receptor, cytokeratin 5/6, and cytokeratin 14 tumor expression, and use of updated methods estimating the clinical relevance of amino acid evolutionary conservation and position. These methodologies may assist genetic counseling of individuals with unclassified sequence variants.

*J Clin Oncol* 26:1657-1663. © 2008 by American Society of Clinical Oncology

### INTRODUCTION

Screening of the breast cancer susceptibility genes *BRCA1* and *BRCA2* identifies numerous nucleotide sequence changes of varying clinical significance. The effect of rare changes predicted to cause missense substitutions or in-frame exon deletions is often not clear, and presents a challenge in the clinical setting. The scale of the problem is considerable, with approximately 30% of *BRCA1* and 60% of *BRCA2* entries in the Breast Cancer Information Core database for *BRCA1* and *BRCA2* variation<sup>1</sup> described as unclassified variants (UVs).

An integrated approach to classification of UVs in *BRCA1* and *BRCA2* into high-risk mutations and neutral variants was developed to define a reliable

protocol for prediction of the clinical significance of UVs.<sup>2</sup> This multifactorial likelihood model used data on co-occurrence of the UV with pathogenic mutations in the same gene, segregation in families, and amino acid physicochemical properties and evolutionary conservation. The model was used to estimate the odds of causality, a ratio of the likelihood of the observed data under the hypothesis of causality to that under the hypothesis of neutrality. Because the model cannot distinguish between variants that are truly benign and those that might be associated with modest risk, neutral variants are sometimes alternatively termed to be of low/little clinical significance (neutral/LCS). We recently revised the model to take into account relevant features of *BRCA1*- and *BRCA2*-associated tumors,

From the Queensland Institute of Medical Research; School of Medicine, and School of Molecular and Microbial Sciences, University of Queensland, Brisbane; Centre for Genetic Epidemiology, University of Melbourne and Peter MacCallum Cancer Centre, Melbourne, Australia; International Agency for Research on Cancer, Lyon, France; and the Department of Dermatology, University of Utah, Salt Lake City, UT. Submitted June 26, 2007; accepted December 5, 2007.

The Kathleen Cunningham Foundation Consortium for Research into Familial Breast Cancer (kConFab) is supported by grants from the National Breast Cancer Foundation, the National Health and Medical Research Council (NHMRC) and by the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. The kConFab clinical follow-up study was funded by NHMRC Grants No. 145684 and 288704; and by a grant from the Susan G. Komen Breast Cancer Foundation, and the NHMRC. J.L.H. and G.C.T. are NHMRC senior principal research fellows; A.B.S. is funded by an NHMRC career development award; L.Da.S. is supported by a fellowship from the Ludwig Institute for Cancer Research; D.B. is a recipient of a postdoctoral fellowship from the International Agency for Research on Cancer; and S.V.T. and D.E.G. were supported in part by the INHERIT BRCA program from the Canadian Institute for Health Research, and a subaward agreement from the Mayo Clinic, Rochester, MN.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Amanda B. Spurdle, PhD, Queensland Institute of Medical Research, c/o Royal Brisbane Hospital Post Office, Herston, Queensland 4029, Australia; e-mail: Amanda.Spurdle@qimr.edu.au.

© 2008 by American Society of Clinical Oncology

0732-183X/08/2610-1657/\$20.00

DOI: 10.1200/JCO.2007.13.2779

including the characteristic histopathology associated with pathogenic mutations in *BRCA1*. Use of the revised model classified 56% of 25 unselected UVs (seven of 10 *BRCA1* and seven of 15 *BRCA2*) from Australian families with breast cancer.<sup>3</sup> Fewer *BRCA2* UVs were classified, largely because *BRCA2* evolutionary predictions were less robust (fewer cross-species sequences were available at that time, relative to *BRCA1*), and because *BRCA2* mutations are not associated with very distinctive pathologic features. Overall, these findings suggested that further extension of the model may improve classification rates, and that classification of *BRCA2* UVs particularly would be improved with more robust evolutionary predictions.

Several immunohistochemical (IHC) markers, including tumor expression of estrogen receptor (ER), cytokeratin 5/6 (CK5/6), cytokeratin 14 (CK14), and cytokeratin 17 (CK17) have been shown to be indicators of *BRCA1* mutation status.<sup>4-7</sup> We undertook a study to investigate the clinical significance of a new panel of *BRCA1* and *BRCA2* UVs identified in Australian families with breast cancer, specifically exploring the value of a subset of these additional IHC markers as histopathologic classifiers of *BRCA1* tumors, and considering the prior probability of pathogenicity based on improved evolutionary conservation analysis methods.

## PATIENTS AND METHODS

### Subjects and Laboratory Methodology

Pedigrees with UVs in *BRCA1* and *BRCA2* were ascertained by the Kathleen Cuninghame Foundation Consortium for Research into Familial

Breast Cancer (kConFab).<sup>8</sup> Participants provided written informed consent, and relevant ethics committees approved the study. A total of 276 individuals (proband and family members) were available for genetic screening. Cancer-free controls ( $n = 180$ ) from the Australian Breast Cancer Family Study have been described previously.<sup>3</sup> Denaturing high performance liquid chromatography and sequencing were as described previously,<sup>3</sup> using primers detailed in Table 1. IHC was performed using standard methods.<sup>6,9</sup> The study included 23 families with 22 different UVs (Tables 2 and 3).

### Protein Modeling

Molecular modeling was carried out on a Silicon Graphics Inc workstation using the Insight II software package (Accelrys, San Diego, CA). Modeling was carried out on the crystal structure of the ring domain (1JM7.pdb),<sup>10</sup> the *BRCA1*-COOH terminal (BRCT) repeat region of *BRCA1* (1JNX.pdb),<sup>11</sup> and the *BRCA2* crystal structure published by Shin et al.<sup>12</sup>

### Prior Probability of Pathogenicity From Amino Acid Conservation and Location of the Mutation in Specific Known Functional Domains

Missense substitutions and in-frame deletions were classified according to their location within one of two recognized functional domains of the proteins, the C-terminus region containing the *BRCA1* BRCT repeats, defined loosely as amino acids 1396 to 1862, and the *BRCA2* DNA-binding domain (DBD; amino acids 2500 to 3098). Variants were further categorized according to whether the wild-type residue involved in the substitution/deletion was evolutionarily conserved through to the pufferfish *Tetraodon*, using multiple sequence alignments. Heterogeneity analysis of 1,433 variants in the Myriad Genetics Laboratory database was used to estimate the proportion of deleterious variants in three classifications<sup>13</sup>: invariant position in BRCT/DBD domain, proportion = 0.73; variable

Table 1. DHPLC Primer Sequences and Conditions

Variant	Sequence		DHPLC Melt Temperature Used (°C)
	Forward (5'>3')	Reverse (5'>3')	
BRCA1 655 A>G (Y179C)	GGAAACCAGTCTCAGTGTCCA	CACTTCCCAAAGCTGCCTAC	58
BRCA1 1575 T>C (F486L)	TGGGAAAACCTATCGGAAGA	CCGTTTGGTTAGTCCCTGA	57
BRCA1 1767 A>C (N550H)	CCTACATCAGGCCTTCATCC	AGGTTCCAGCTTTCGTTTTGAA	56
BRCA1 2596 C>A (T826K)	GGCACTCAGGAAAGTATCTCG	AATGACTGGCGCTTTGAAAC	58
BRCA1 2878 T>C (V920A)	ACATTCTCTGCCACTCTGG	TTTCGTTGCCTCTGAACTGA	57
BRCA1 3415 C>T (P1099L)	CCAGTGATGAAAACATTC AAGC	GCATGACTACTTCCCATAGGC	55
BRCA1 3446 A>C (K1109N)			
BRCA1 3827 T>G (N1236K)	ATTTGGCTCAGGGTTACCG	TGCCTTTGCCAATATTACCTG	58
BRCA1 IVS 18 + 1 G>T (del exon 18)	GAGGCTCTTAGCTTCTTAGGAC	CCAGCATCACCAGCTTATCT	56
BRCA1 5632 T>A (V1838E)	AGGACCTGGAGTCGATTG	GCAGTCAGTAGTGGCTGTGG	61
BRCA2 281 G>A (R18H)	CGTAGGTA AAAATGCCTATTGGA	TGTGGTTAACCTGCAAACGA	56
BRCA2 1766 A>G (K513R)	TCTTGCAGTAAAGCAGGCAAT	AGGAGTCCTCCTCTGTGAGC	56
BRCA2 2048 A>C (K607T)	TTTGCTCACAGAAGGAGGACT	TGCTCAAAGCTGGGCTGAAC	56
BRCA2 3031 G>A (D935N)	TAAGTGTCAATCCAGACTCTG	CTATATTCAAAGGAGATGTCCG	55
BRCA2 3743 C>T (S1172L)	GAAGAGTACATTTGAAGTGCCTGA	CCCTAAACCCCACTTCATT	58
BRCA2 4391 C>A (T1388N)	GAATTTGATGGCAGTGATTC A A	TTTTGCTCCGTTTTAGTAGCA	54
BRCA2 5506 T>G (S1760A)	GACAAAAATCATCTCTCCGAAAA	CAGTTTGTGGGTATGCATTTG	55
BRCA2 7643 A>C (K2472T)	CATTGATGGACATGGCTCTG	GGGAAAACCATCAGGACATT	55
BRCA2 7772 C>T (T2515I)	CAAGTCTTCAGAATGCCAGAGA	CACTCTGCATAAAAAGCCATCAG	60
BRCA2 8377 G>T (A2717S)	TGCAAAAACACTTGTCTCTGTG	CAGTACATCTAAGAAATTGAGCATCC	58
BRCA2 8801 A>G (Q2858R)	TGCCTGGCCTGATACAATTA	TGTCCTTGTGTCTATTCTTTG	57
BRCA2 9078 G>T (K2950N)	TGGTCACAGGGTATTTTCAGTG	GCTTACAATACGCAACTTCCA	58
BRCA2 9133 G>A (V2969M)			
BRCA2 9345 G>A (del exon 23)*	CATACAGTTAGCAGCGACAAAAA	CCTCAGAACAAAGATGGCTGA	56

Abbreviation: DHPLC, denaturing high performance liquid chromatography.

\*Also known as P3039P.

Evaluation of *BRCA1/2* Unclassified Sequence Variants

**Table 2.** *BRCA1* Unclassified Variants—Likelihood Scores and Posterior Probability of Pathogenicity

<i>BRCA1</i> Unclassified Variant	Conserved to Tetraodon and in BRCT Domain	Sequence Alignment Prior Probability	Myriad			ER	CK 5/6	CK 14	Overall Grade
			Frequency	No. of Different Deleterious Mutations Estimated in Trans					
655 A>G (Y179C)*	Outside	.02	104	3	Positive	Negative	Negative	2	
1575 T>C (F486L)*	Outside	.02	106	3	Positive	Negative	Negative	2	
1767 A>C (N550H)*	Outside	.02	103	3	Positive	Negative	Negative	2	
2596 C>A (T826K)	Outside	.02	67	3	NA	NA	NA	0	
2878 T>C (V920A)†	Outside	.02	1	0	NA	NA	NA	0	
3415 C>T (P1099L)	Outside	.02	42	2	NA	NA	NA	0	
					Positive	Negative	Negative	1	
					Positive	Negative	Negative	1	
3446 A>C (K1109N)‡	Outside	.02	11	0	Negative	NA	NA	3	
3827 T>G (N1236K)	Outside	.02	68	0	NA	NA	NA	0	
5632 T>A (V1838E)	Variable, inside	.08	2	0	Negative	NA	Positive	3	
IVS 18 + 1 G>T	Invariant, inside	.73	4	0	Negative	Positive	Positive	3	
					Negative	Positive	Negative	3	
del exon 18					Negative	Positive	Positive	3	
					Negative	NA	Positive	3	
					Negative	Positive	Negative	3	
					Negative	Positive	Negative	3	

<i>BRCA1</i> Unclassified Variant	LR		Bayes Odds for the Variant	Odds for Causality	Intermediate Classification	Posterior Probability of a Variant Being Deleterious	Classification With Sequence Alignment
	Co-Occurrence	LR Pathology					
655 A>G (Y179C)*	8.49093E-07	0.143	0.0007	8.49942E-11	Neutral/LCS	1.73458E-12	Neutral/LCS
1575 T>C (F486L)*	9.1225E-07	0.143	0.0007	9.13162E-11	Neutral/LCS	1.8636E-12	Neutral/LCS
1767 A>C (N550H)*	0.0000008	0.143	0.0007	8.200E-11	Neutral/LCS	1.673E-12	Neutral/LCS
2596 C>A (T826K)	0.0000002	1	0.4985	0.0000001	Neutral/LCS	2.291E-09	Neutral/LCS
2878 T>C (V920A)†	1.0365238	1	0.9811	1.0169335	Unclassified	0.0203318	Unclassified
3415 C>T (P1099L)	0.0000336	0.020449	0.00222216	1.528E-09	Neutral/LCS	3.119E-11	Neutral/LCS
3446 A>C (K1109N)‡	1.4837894	2.95	0.0078	0.0341420	Unclassified	0.0006963	Neutral/LCS
3827 T>G (N1236K)	11.465478	1	0.2975	3.4109796	Unclassified	0.0650814	Unclassified
5632 T>A (V1838E)	1.0743817	626.912	5.9	3973.9023	Pathogenic	0.9971145	Pathogenic
IVS 18 + 1 G>T (del exon 18)	1.1542960	7561.5232	12.87	112332.40	Pathogenic	0.9999967	Pathogenic

NOTE. The Posterior Probability is calculated from the prior probability (based on sequence data) and the Odds for Causality using Bayes rule. Example calculation of posterior probability, for variant *BRCA1* V1838E: the Prior Probability for V1838E is .08 since variant is inside the BRCT domain but is not conserved (see Patients and Methods). The Odds for Causality is 3973.9023, calculated as the product of the individual statistically independent components (LR Co-occurrence [1.0743817] × LR Pathology [626.912] × LR Segregation [5.9]). Note the Intermediate Classification is based on the Odds for Causality, where a variant with odds greater than 1,000:1 is considered pathogenic, and a variant with odds less than 1:100 is considered neutral/LCS (after Goldgar et al<sup>2</sup>). The Posterior Probability for V1838E is 0.9971145 = Posterior Odds/(Posterior Odds+1), where the Posterior Odds = Prior Probability (.08) × Odds for causality (3973.9023) × (1/1-prior probability).

Abbreviations: BRCT, *BRCA1* C-terminus domain; ER, estrogen receptor status; CK5/6, cytokeratin 5/6 status; CK14, cytokeratin 14 status; LR, likelihood ratio; LCS, low clinical significance; NA, not available.

\*These three variants occur in cis in a single family. All tumor information is thus derived from a single tumor. A breast-affected individual from the same family has subsequently been reported to carry the pathogenic mutation *BRCA1* 546 G>T (E143X).

†A pathogenic mutation *BRCA1* 5622 C>T (R1853X) was subsequently identified in an affected individual in a distant branch of the family.

‡The youngest affected member of the family was also subsequently found to carry *BRCA1* 546 G>T (E143X) in addition to the unclassified variant.

position in BRCT/DBD domain, proportion = 0.08; and position outside of BRCT/DBD domain whether invariant or variable, proportion = 0.02. The values were then used as prior probabilities of being deleterious for classification of the studied variants.

**Co-Occurrence With Pathogenic Mutations**

We queried the Myriad Genetic Laboratories database of approximately 90,000 full-sequence tests to determine the number of times a UV was observed, and the number of different deleterious mutations observed to co-occur with each variant, as a measure of the number of times the UV is seen in trans with a deleterious mutation. Phase of the variant and mutation was established for a subset of individuals. Observations for variants were excluded if in cis with a mutation, included if in trans with a mutation, and assumed to be in trans with at least *n*-1 observations for *n* observations with different deleterious mutations of unknown phase.<sup>14</sup>

**Histopathology and Pedigree Causality Analysis**

Available invasive tumor sections were analyzed by two pathologists (S.L., L.Da.S.) blind to UV status, for parameters recognized to be associated with *BRCA1* or *BRCA2* mutation status.<sup>6,15,16</sup> IHC scoring was performed by a single pathologist (S.L.) as described previously.<sup>6</sup>

**Pedigree Causality Analysis, Derivation of Probabilities, Multifactorial Likelihood Scoring**

Bayes factor analysis was as described previously.<sup>3,13,17</sup> Probabilities were derived for each of the following components, under the assumption that each factor was independent. For the co-occurrence component, we estimated the likelihood that a UV was causal as described previously.<sup>2</sup> The Bayes factor was included directly as a likelihood ratio (LR) score for the pedigree analysis component. For *BRCA1*, ER, CK5/6, and CK14 expression was used for calculating histopathologic LR scores, based on the previously reported

**Table 3.** *BRCA2* Unclassified Variants—Likelihood Scores and Posterior Probability of Pathogenicity

<i>BRCA2</i> Unclassified Variant	Conserved to Tetraodon and in DBD Domain	Sequence Alignment Prior Probability	Myriad			
			Frequency	No. of Different Deleterious Mutations Estimated in Trans	Tubule Formation (%)	LR Co-Occurrence
281 G>A (R18H)	Outside	.02	5	0	> 75	1.1328651
1766 A>G (K513R)*	Outside	.02	9	0	< 10	1.2517591
					< 10	
2048 A>C (K607T)†	Outside	.02	6	0	< 10	1.1614856
3031 G>A (D935N)	Outside	.02	205	5	> 75	0.0000133
					< 10	
					< 10	
3743 C>T (S1172L)	Outside	.02	62	1	< 10	4.6968833
5506 T>G (S1760A)	Outside	.02	1	0	> 75	1.0252638
7643 A>C (K2472T)	Outside	.02	2	0	10-75	1.0511659
7772 C>T (T2515I)	Variable, inside	.08	141	3	< 10	0.0018610
8377 G>T (A2717S)	Variable, inside	.08	257	4	NA	0.0012803
8801 A>G (Q2858R)	Variable, inside	.08	11	0	10-75	1.3158066
9078 G>T (K2950N)‡	Variable, inside	.08	196	2	10-75	0.1927832
9345 G>A (del exon 23)	Invariant, inside	.73	28	0	NA	2.0109347

  

<i>BRCA2</i> Unclassified Variant	LR Pathology	Bayes Odds for the Variant	Odds for Causality	Intermediate Classification	Posterior Probability of a Variant Being Deleterious	Classification With Sequence Alignment
281 G>A (R18H)	0.5	0.01990	0.0112720	Unclassified	.0002300	Neutral/LCS
1766 A>G (K513R)*	1.44	0.00400	0.0072101	Neutral/LCS	.0001471	Neutral/LCS
2048 A>C (K607T)†	1.2	0.18660	0.2600799	Unclassified	.0052797	Unclassified
3031 G>A (D935N)	0.72	2.70000	0.0000259	Neutral/LCS	.0000005	Neutral/LCS
3743 C>T (S1172L)	1.2	0.00786	0.0443145	Unclassified	.0009036	Neutral/LCS
5506 T>G (S1760A)	0.5	1.17000	0.5997793	Unclassified	.0120924	Unclassified
7643 A>C (K2472T)	0.5	0.00001	0.0000053	Neutral/LCS	.0000001	Neutral/LCS
7772 C>T (T2515I)	1.2	0.93200	0.0020814	Neutral/LCS	.0001810	Neutral/LCS
8377 G>T (A2717S)	1	0.12110	0.0001550	Neutral/LCS	.0000135	Neutral/LCS
8801 A>G (Q2858R)	0.5	0.93800	0.6171133	Unclassified	.0509291	Unclassified
9078 G>T (K2950N)‡	0.5	0.04385	0.0042263	Neutral/LCS	.0003674	Neutral/LCS
9345 G>A (del exon 23)	1	1.12000	2.2522468	Unclassified	.8589445	Unclassified

NOTE. The Posterior Probability is calculated from the prior probability (based on sequence data) and the Odds for Causality using Bayes rule. Example Calculation of posterior probability, for variant *BRCA2* R18H: the Prior Probability for R18H is .02 since variant is outside the DBD domain (see Patients and Methods). The Odds for Causality is 0.0112720, calculated as the product of the individual statistically independent components (LR Co-occurrence (1.1328651) × LR Pathology (0.5) × LR Segregation (0.01990)). Note the Intermediate Classification is based on the Odds for Causality, where a variant with odds greater than 1,000:1 is considered pathogenic, and a variant with odds less than 1:100 is considered neutral/LCS (after Goldgar et al<sup>2</sup>). The Posterior Probability for R18H is .0002300 = Posterior Odds/(Posterior Odds + 1), where the Posterior Odds = Prior Probability (.02) × Odds for causality (0.0112720) × (1/1 - prior probability).  
Abbreviations: DBD, *BRCA2* DNA-binding domain; LR, likelihood ratio; LCS, low clinical significance; NA, not available.  
\*Members of this family were found to carry the *BRCA1* 2080delA (Stop700) pathogenic mutation.  
†Some members of this family also carry *BRCA1* 5622 C>T (R1835X).  
‡Two families were included in the analysis. An additional family carrying *BRCA2* 9132 del C (STOP 2975) in cis with the unclassified variant was excluded from study.

prevalence of the combined immunotypes of these independent predictors of *BRCA1* mutation status in breast tumors.<sup>6</sup> The likelihoods for causality were: ER positive (irrespective of CK score) = 0.14:1; negative for all three markers = 0.87:1; ER negative, CK14 negative, CK5/6 positive = 5.6:1; ER negative, CK14 positive, CK5/6 negative = 2.6:1; ER negative, CK14 positive, CK5/6 positive = 27.4:1. Categories were collapsed for scoring the single ER-negative, CK14-positive tumor lacking CK5/6 data (LR, 8.8:1). The likelihood was calculated from ER expression and grade (LR, 2.95:1), as previously described,<sup>3</sup> for the single ER-negative grade 3 tumor lacking CK data. For *BRCA2*, tubule formation was used for LR estimates, based on the previously reported prevalence of this independent predictor of *BRCA2* mutation status.<sup>16</sup> Likelihoods were: tubule formation in less than 10% of tumor = 1.2:1; tubule formation  $\geq 10\%$  = 0.5:1, as used previously.<sup>3</sup>

The individual LRs were multiplied to calculate an overall multifactorial likelihood ratio assuming statistical independence of the sources of information. Bayes rule was then used to calculate a posterior probability that the variant was deleterious from the multifactorial LR, and the prior

probability determined from sequence alignment. Variants with a posterior probability  $\geq 0.99$  were classified as pathogenic, and those with probability  $\leq 0.001$  were classified as neutral/LCS. Example calculations are provided in Tables 2 and 3.

## RESULTS

We analyzed a total of 10 *BRCA1* UVs and 12 *BRCA2* UVs in this study (Tables 2 and 3). All UVs were predicted missense substitutions, except for two in-frame exon deletions, *BRCA1* IVS18+1 G>T (del exon 18) and *BRCA2* 9345 G>A (del exon 23). All UVs fell in regions reported to interact with at least one other protein thought to be involved in DNA repair. Protein modeling predictions were possible for only two variants: *BRCA1* 5632 T>A (V1838E) in the BRCT

domain, and *BRCA2* 5506 T>G (S1760A) in the BRAF35 interaction domain. The *BRCA1* V1838E substitution creates a charged hydrophilic group in a hydrophobic patch, and the glutamine side chain is predicted to clash with the Leu1790 side chain, and to a lesser extent with the Phe1761 side chain. This alteration might thus be expected to have serious effects on protein structure. The *BRCA2* S1760A substitution is between the *BRCA2* repeat 5 (1664 to 1698) and repeat 6 (1837 to 1871), and would not cause a major effect on protein structure, but may be important in that it would remove a potential *N*-linked glycosylation site. Screening of controls for all UVs under study identified a single UV, *BRCA1* 1767 A>C (N550H), in 1/180 controls.

Amino acid position and sequence alignment analysis was used to group missense substitutions into categories, as described in the Patients and Methods section. Overall, eight variants fell within the BRCT/DBD domains, only two of which were conserved to *Tetraodon* (Tables 2 and 3). Prior probabilities were assigned according to category, for inclusion in the multifactorial analysis.

All the UVs seen in these families have been identified at least once by Myriad Genetic Laboratories Inc. Five *BRCA1* and five *BRCA2* UVs were considered to co-occur in trans with at least one known deleterious mutation in the same gene, and phase was proven to be in trans at least once for all but one (*BRCA2* 3743 C>T [S1172L]). LRs based on co-occurrence alone were indicative of neutrality/LCS for five of five *BRCA1* UVs and three of five *BRCA2* UVs with co-occurrent mutations. Furthermore, a homozygote with *BRCA2* 3031 G>A (D935N) and no clinical features of Fanconi anemia has been identified in the Myriad data set, further suggesting that this variant is neutral/LCS. Segregation analysis provided firm evidence regarding causality for eight UVs, including the single family carrying the *BRCA1* 655 A>G (Y179C), 1575 T>C (F486L), and 1767 A>C (N550H) variants in cis.

There was at least one tumor sample available for pathology review and IHC screening for all but two *BRCA1* and one *BRCA2* UVs (Tables 2 and 3). The specific *BRCA1* histopathologic features included in LR estimates provided evidence that *BRCA1* 5632 T>A V1838E and IVS18+1 G>T (del exon 18) were likely pathogenic. All these tumors were ER negative and high grade, all five with available CK5/6 data expressed this basal marker of *BRCA1* mutation status, and CK14 was expressed in four of seven tumors. No other tumors from carriers of *BRCA1* UVs displayed cytokeratin expression consistent with a *BRCA1* mutation. Of the 14 *BRCA2* tumors screened, one of two tumors for K513R had histopathologic features of *BRCA1* tumors (high grade, ER negativity, focal positivity for CK14). This result was consistent with the fact that this individual (but not the other relative screened) was found to carry a *BRCA1* pathogenic mutation during the course of this study.

The combined odds of causality for each variant were derived as described in the Patients and Methods section, including LRs for pathology, co-occurrence, and segregation data, to generate an intermediate classification. Of the 10 *BRCA1* UVs studied, five were classified as neutral/LCS and two as pathogenic. The overall classification rate was less for the 12 *BRCA2* UVs, with six classified as neutral/LCS. After combination of the multifactorial odds with prior probabilities based on sequence analysis, two *BRCA1* variants (2878 T>C [V920A]; 3827 T>G [N1236K]), and four *BRCA2* variants (2048 A>C

[K607T]; 5506 T>G [S1760A]; 8801 A>G [Q2858R]; 9345 G>A [del exon 23]) remained unclassified.

## DISCUSSION

Our study shows that multifactorial likelihood analysis, in combination with prior probabilities based on evolutionary and physicochemical conservation, is a valuable tool to evaluate clinically problematic *BRCA1/2* sequence variants. This approach classified 72% of the 22 UVs studied. Only two *BRCA1* UVs (5632 T>A [V1838E] and IVS18+1 G>T [del exon 18]) were classified as pathogenic, and there was strong evidence for pathogenicity (86% posterior probability of being deleterious) for one *BRCA2* variant, 9345 G>A (del exon 23). This highlights the general understanding that the majority of UVs are individually of little clinical importance, based on the fact that the number of UVs far exceeds the proportion of families linked to *BRCA1* or *BRCA2* but with no identified pathogenic mutation. Both *BRCA1* variants fall within or span the BRCT motifs, domains well recognized to be critical for *BRCA1* function, while *BRCA2* exon 23 spans the BUBR1A/FilaminA/BCCI *p*-alpha interaction sites and contains residues conserved to *Arabidopsis*.

The interpretation from protein modeling data was also consistent with final classification. *BRCA1* V1838E was predicted to have a serious effect on protein structure, and this variant was classified as pathogenic. Likewise, evidence from modeling of *BRCA2* 5506 T>G (S1760A) was more equivocal, and this variant remained unclassified. Results from control screening supported our previous suggestion that this approach might be a simple adjunct to UV evaluation for nonfounder populations,<sup>3</sup> with a neutral/LCS classification for the single UV observed in controls. While underpowered to evaluate pathogenicity of rare variants, it is a useful prescreen to exclude common variation in little studied population groups.

As shown previously, the most informative components of the multifactorial likelihood predictions were co-occurrence and pedigree analyses. Co-occurrence scores alone were sufficient to classify six UVs (27%), and odds from segregation analysis classified eight UVs (36%), including three variants in cis. A total of 11 UVs (50%) could be classified considering only co-occurrence and segregation data.

Information from tumors was available for most UVs, but only added weight to the final classification for the subset of pathogenic *BRCA1* UVs. This indicates that ER and cytokeratin IHC would be very helpful in the classification of likely pathogenic *BRCA1* UVs. Although cytokeratin IHC is not yet routinely carried out in all diagnostic laboratories, the  $\kappa$  scores for agreement on staining positivity of CK5/6 and CK14 reported for a large data set were 0.74 to 0.82 (SE, 0.06 to 0.07), corresponding to concordance rates of 88% to 92%.<sup>6</sup> This suggests that cytokeratin staining could be introduced into routine histopathologic practice, especially in larger centers, to provide information for evaluation of *BRCA1* UVs. In our experience, CK5/6 and CK14 IHC can be implemented relatively easily if appropriate quality controls are in place—as most pathology laboratories practice routinely. We recognize that the IHC approach relies on the underlying assumption that missense and in-frame deletions will exhibit the same tumor histopathologic characteristics as truncating mutations,

but have confidence that it is valid for classification of high-risk *BRCA1* mutations from our published<sup>17a</sup> and unpublished data showing that functionally deficient *BRCA1* RING and BRCT domain missense substitutions exhibit immunohistopathologic profiles consistent with a *BRCA1* mutation.

Pathology scores for *BRCA2* UVs were not very informative individually, but in two instances contributed to the final classification. In this study, we did not include analysis of tumor loss of heterozygosity for *BRCA1* and *BRCA2*, as done previously,<sup>3</sup> because analysis of this and our previous data set revealed statistically significant increased loss of the variant compared with what is expected for the underlying hypothesis used previously to calculate likelihood estimates.

Sequence alignment analysis was useful for refining the classification, especially for *BRCA2* variants. The increased value of this approach over our previous study<sup>3</sup> was largely due to increasing the informativeness with additional sequences in the *BRCA2* alignment, and the incorporation of estimates of prior probabilities based on both alignment and location within functionally important domains. *BRCA1* and *BRCA2* protein multiple sequence alignments used for this analysis (or updated versions thereof) are publically available through the web-based program and alignments at <http://agvgd.iarc.fr> (International Agency for Research on Cancer; Align GVHD). Thus, investigators could use this resource to estimate prior probabilities for any *BRCA1* or *BRCA2* missense substitutions of interest. However, the clinical application of the methods presented in this study should be undertaken only with consideration of the caveats associated with the approaches, principally that prior probabilities and likelihood ratio estimates were derived assuming that variants were either neutral or of similar high risk as the average *BRCA1/2* mutation,<sup>18</sup> and from analysis of single data sets for the prior probability, co-occurrence and histopathology components, and that likelihoods are valid for patients with familial breast cancer. Moreover, it is likely that further development of sequence information-based estimates of prior probability to incorporate the physicochemical characteristics of unclassified variants as assessed by Align GVGD analysis<sup>14</sup> will refine different classes of substitutions, and thus improve the precision of the prior probabilities associated with individual substitutions.

Few of the UVs studied here have been analyzed previously, and none using the range of approaches we examined. Comparisons to larger studies with several points of evidence are not inconsistent with our findings. Co-occurrence and evolutionary conservation analysis of *BRCA1* missense variants<sup>14,19</sup> classified *BRCA1* Y179C, F486L, N550H, and P1099L as neutral/LCS, and another study found F486L to be absent in 1054 ethnicity-matched controls, but failed to classify it using a multifactorial approach based on odds derived from tumor information, bilaterality, and family history of ovarian cancer.<sup>20</sup> The *BRCA1* T826K and *BRCA2* S1172L, T2515I, and A2717S UVs were all reported to be present in the proband from at least one German family with breast/breast-ovarian cancer but absent from 200 ethnically matched controls.<sup>21</sup> The *BRCA2* K2950N variant was observed in a patient with familial prostate

cancer, but also in two of 340 normal chromosomes, supporting the neutral/LCS classification from this study.<sup>22</sup>

Few variants have been studied functionally. The splice site variant *BRCA2* 9345 G>than A has been reported to cause an in-frame deletion of exon 23 from reverse transcription polymerase chain reaction studies of peripheral blood lymphocyte RNA.<sup>23</sup> The T2515I amino acid substitution was shown to be associated with normal expression levels, complete ablation of cell survival activity, partial inactivation of homologous recombination and centrosome regulatory functions, and a cytoplasmic-nuclear localization profile intermediate between the aberrant predominantly nuclear localization of known mutations and the cytoplasmic localization for wild-type *BRCA2*.<sup>24</sup> These data suggest this alteration has a subtle effect on *BRCA2* function, and is thus unlikely to be associated with a high risk of cancer. Indeed, the same study reported low odds in favor of causality from segregation analysis,<sup>24</sup> and our multifactorial analysis classified T2515I as neutral/LCS.

In summary, we have provided evidence for the classification of 16 of 22 different *BRCA1* or *BRCA2* sequence variants. Our findings support the general understanding<sup>13</sup> that the majority of unclassified variants of *BRCA1* and *BRCA2* are not associated with a high risk of disease as associated with classical truncating mutations. Classification of individual variants is necessary to identify the subset demonstrating features of classical mutations with high risk of disease, and is considerably improved by analysis of ER, CK5/6, and CK14 tumor expression, and updated methods to estimate the clinical relevance of amino acid evolutionary conservation and position. These methodologies may be easily implemented, and together with supporting information provided by additional studies such as functional assays, may assist genetic counseling of unclassified sequence variants.

#### AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

#### AUTHOR CONTRIBUTIONS

**Conception and design:** Amanda B. Spurdle

**Financial support:** Amanda B. Spurdle, Melissa A. Brown

**Administrative support:** Amanda B. Spurdle, Sue Healey

**Provision of study materials or patients:** John L. Hopper

**Collection and assembly of data:** Amanda B. Spurdle, Sunil R. Lakhani, Sue Healey, Suzanne Parry, Leonard M. Da Silva, Ross Brinkworth, Davit Babikyan, Georgia Chenevix-Trench, David E. Goldgar

**Data analysis and interpretation:** Amanda B. Spurdle, Sunil R. Lakhani, Sue Healey, Suzanne Parry, Ross Brinkworth, Melissa A. Brown, Georgia Chenevix-Trench, Sean V. Tavtigian, David E. Goldgar

**Manuscript writing:** Amanda B. Spurdle, Sunil R. Lakhani, Sue Healey, Georgia Chenevix-Trench, Sean V. Tavtigian, David E. Goldgar

**Final approval of manuscript:** Amanda B. Spurdle, Sunil R. Lakhani, Sue Healey, Suzanne Parry, Leonard M. Da Silva, Ross Brinkworth, John L. Hopper, Melissa A. Brown, Davit Babikyan, Georgia Chenevix-Trench, Sean V. Tavtigian, David E. Goldgar

## REFERENCES

1. National Human Genome Research Institute: An open access on-line breast cancer mutation data base. <http://research.nhgri.nih.gov/projects/bic>
2. Goldgar DE, Easton DF, Deffenbaugh AM, et al: Integrated evaluation of DNA sequence variants of unknown clinical significance: Application to BRCA1 and BRCA2. *Am J Hum Genet* 75:535-544, 2004
3. Chenevix-Trench G, Healey S, Lakhani S, et al: Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer Res* 66:2019-2027, 2006
4. Foulkes WD, Brunet JS, Stefansson IM, et al: The prognostic implication of the basal-like (cyclinEhigh/p27low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. *Cancer Res* 64:830-835, 2004
5. Foulkes WD, Stefansson IM, Chappuis PO, et al: Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst* 95:1482-1485, 2003
6. Lakhani SR, Reis-Filho JS, Fulford L, et al: Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin Cancer Res* 11:5175-5180, 2005
7. Rodriguez-Pinilla SM, Sarrio D, Honrado E, et al: Vimentin and laminin expression is associated with basal-like phenotype in both sporadic and BRCA1-associated breast carcinomas. *J Clin Pathol* 60:1006-1012, 2007
8. Mann GJ, Thorne H, Balleine RL, et al: Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Res* 8:R12, 2006
9. Lakhani SR, Van De Vijver MJ, Jacquemier J, et al: The pathology of familial breast cancer: Predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol* 20:2310-2318, 2002
10. Brzovic PS, Meza JE, King MC, et al: BRCA1 RING domain cancer-predisposing mutations: Structural consequences and effects on protein-protein interactions. *J Biol Chem* 276:41399-41406, 2001
11. Williams RS, Green R, Glover JN: Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nat Struct Biol* 8:838-842, 2001
12. Shin DS, Pellegrini L, Daniels DS, et al: Full-length archaeal Rad51 structure and mutants: Mechanisms for RAD51 assembly and control by BRCA2. *Embo J* 22:4566-4576, 2003
13. Easton DF, Deffenbaugh AM, Pruss D, et al: A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. *Am J Hum Genet* 81:873-883, 2007
14. Tavtigian SV, Deffenbaugh AM, Yin L, et al: Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* 43:295-305, 2006
15. Lakhani SR, Jacquemier J, Sloane JP, et al: Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst* 90:1138-1145, 1998
16. Lakhani SR, Gusterson BA, Jacquemier J, et al: The pathology of familial breast cancer: Histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. *Clin Cancer Res* 6:782-789, 2000
17. Thompson D, Easton DF, Goldgar DE: A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* 73:652-655, 2003
- 17a. Lovelock PK, Spurdle AB, Mok M, et al: Pathogenicity of the BRCA1 unclassified sequence variants R1699Q, and A1708V: Assessment using functional assays and multifactorial likelihood analysis. *Breast Cancer Res* 9:R82, 2007
18. Antoniou A, Pharoah PD, Narod S, et al: Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. *Am J Hum Genet* 72:1117-1130, 2003
19. Judkins T, Hendrickson BC, Deffenbaugh AM, et al: Application of embryonic lethal or other obvious phenotypes to characterize the clinical significance of genetic variants found in trans with known deleterious mutations. *Cancer Res* 65:10096-10103, 2005
20. Osorio A, Milne RL, Honrado E, et al: Classification of missense variants of unknown significance in BRCA1 based on clinical and tumor information. *Hum Mutat* 28:477-485, 2007
21. Meindl A: Comprehensive analysis of 989 patients with breast or ovarian cancer provides BRCA1 and BRCA2 mutation profiles and frequencies for the German population. *Int J Cancer* 97:472-480, 2002
22. Gayther SA, de Foy KA, Harrington P, et al: The frequency of germ-line mutations in the breast cancer predisposition genes BRCA1 and BRCA2 in familial prostate cancer: The Cancer Res Campaign/British Prostate Group United Kingdom Familial Prostate Cancer Study Collaborators. *Cancer Res* 60:4513-4518, 2000
23. Peelen T, van Vliet M, Bosch A, et al: Screening for BRCA2 mutations in 81 Dutch breast-ovarian cancer families. *Br J Cancer* 82:151-156, 2000
24. Wu K, Hinson SR, Ohashi A, et al: Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. *Cancer Res* 65:417-426, 2005

---

**Acknowledgment**

We thank Heather Thorne, Eveline Niedermayr, the Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (kConFab) research nurses and staff, heads and staff of the Family Cancer Clinics, and the clinical follow-up study for their contributions to kConFab, and the many families who contribute to kConFab. We thank Amie Deffenbaugh for data on variant co-occurrence with mutations from the Myriad Genetic Laboratories database.