Published in final edited form as:

Hum Mutat. 2011 June ; 32(6): 678–687. doi:10.1002/humu.21495.

Splicing and Multifactorial Analysis of Intronic *BRCA1* and *BRCA2* Sequence Variants Identifies Clinically Significant Splicing Aberrations up to 12 Nucleotides from the Intron/Exon Boundary

Phillip J. Whiley^{1,2}, Lucia Guidugli³, Logan C. Walker¹, Sue Healey¹, Bryony A. Thompson¹, Sunil R. Lakhani⁴, Leonard M. Da Silva⁴, kConFab Investigators⁵, Sean V. Tavtigian⁶, David E. Goldgar⁷, Melissa A. Brown^{2,8}, Fergus J. Couch³, and Amanda B. Spurdle^{1*}

¹Queensland Institute of Medical Research, Brisbane, Australia

²School of Chemistry and Molecular Sciences, University of Queensland, Brisbane, Australia

³Laboratory of Medicine and Pathology, Mayo Clinic, Rochester, Minnesota

⁴Molecular & Cellular Pathology, University of Queensland Centre for Clinical Research (UQCCR), School of Medicine, University of Queensland, and Pathology Queensland: The Royal Brisbane and Women's Hospital, Brisbane, Australia

⁵Peter MacCallum Cancer Centre, Melbourne, Australia

⁶Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah

⁷Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah

⁸University of Queensland Diamantina Institute, Brisbane, Australia

Abstract

Clinical management of breast cancer families is complicated by identification of *BRCA1* and *BRCA2* sequence alterations of unknown significance. Molecular assays evaluating the effect of intronic variants on native splicing can help determine their clinical relevance. Twentysix intronic *BRCA1/2* variants ranging from the consensus dinucleotides in the splice acceptor or donor to 53 nucleotides into the intron were identified in multiple-case families. The effect of the variants on splicing was assessed using HSF matrices, MaxEntScan and NNsplice, followed by analysis of mRNA from lymphoblastoid cell lines. A total of 12 variants were associated with splicing aberrations predicted to result in production of truncated proteins, including a variant located 12

© 2011 Wiley-Liss, Inc.

*Correspondence to: Amanda B. Spurdle, Queensland Institute of Medical Research, 300 Herston Rd, Herston, QLD, 4006, Australia. Amanda.Spurdle@qimr.edu.au.

The first two authors contributed equally to this article.

The last two authors contributed equally to this article.

Conflicts of Interest: The authors have no potential conflicts of interest with study outcomes.

Additional Supporting Information may be found in the online version of this article.

nucleotides into the intron. The posterior probability of pathogenicity was estimated using a multifactorial likelihood approach, and provided a pathogenic or likely pathogenic classification for seven of the 12 spliceogenic variants. The apparent disparity between experimental evidence and the multifactorial predictions is likely due to several factors, including a paucity of likelihood information and a nonspecific prior probability applied for intronic variants outside the consensus dinucleotides. Development of prior probabilities of pathogenicity incorporating bioinformatic prediction of splicing aberrations should improve identification of functionally relevant variants and enhance multifactorial likelihood analysis of intronic variants.

Keywords

BRCA1; *BRCA2*; clinical significance; variants

Introduction

Sequencing of the cancer predisposition genes *BRCA1* (MIM# 113705) and *BRCA2* (MIM# 600185) is often undertaken for families presenting with multiple breast and/or ovarian cancer cases, in order to determine the most appropriate options for clinical management [Schwartz et al., 2008]. For clinicians, formulating advice based on individual genetic sequence is complicated by the incidence of sequence alterations of unknown clinical significance, often termed unclassified variants (UVs) or variants of uncertain clinical significance (VUS). At present, 1,785 distinct sequence variants in *BRCA1/2* on the Breast Cancer Information Core database (<http://research.nhgri.nih.gov/bic/>) fall in this category and include missense changes, small in-frame insertions or deletions, potential splice-site alterations, and also variants within noncoding and intronic regions. Synthesizing data derived from tumor characteristics and family-based genetic data using a multifactorial likelihood modelling approach provides a strong platform to estimate disease association [Chenevix-Trench et al., 2006; Easton et al., 2007; Goldgar et al., 2004; Lakhani et al., 1998; Spurdle et al., 2008b; Tavtigian et al., 2008]. However, for variants where limited data from individuals within a family is available for such modeling, functional analysis, particularly analysis of splicing aberrations may be used to infer pathogenicity [Spurdle et al., 2008a; Walker et al., 2010].

Intronic sequence variants play a key role in the regulation of pre-mRNA splicing [Chen et al., 2006; Kwong et al., 2008], and there is a firmly established link between aberrant splicing and predisposition to multiple human diseases [Lopez, 1998; Srebrow and Kornblihtt, 2006]. Therefore, there is a strong rationale for investigating the role of *BRCA1* and *BRCA2* intronic variants in the production of aberrant mRNA transcripts using bioinformatic analysis and in vitro assays, for the purpose of defining their clinical relevance. Intronic variants within highly conserved donor or acceptor dinucleotide sequences at splice junctions have the greatest potential to affect splicing by disrupting splicing motifs [House and Lynch, 2008], and such variants are generally considered pathogenic on the basis of the position alone. Intronic sequence variants in close proximity to the donor or acceptor dinucleotides have also been shown to alter normal splicing by disrupting the splice recognition motif or by promoting use of cryptic splice sites not

recognized or little recognized during normal splicing [Pagani and Baralle, 2004]. Lastly, variants deep into intronic regions have been reported to create de novo splice sites resulting in intronic insertions or to disrupt putative splice enhancers/silencers resulting in dysregulation of normal splicing [Chen et al., 2006; Davis et al., 2009; Harland et al., 2001; Homolova et al., 2010; Matsushima et al., 1995; Pagani et al., 2002; Rio Frio et al., 2009; Yu et al., 2008]. Although bioinformatic methods have some value in assessing the effect of a particular variant on splicing, they are currently unable to accurately predict the use of cryptic sites and thus define the resulting aberration [Claes et al., 2003; Vreeswijk et al., 2009].

This study was undertaken to assess pathogenicity of 26 intronic *BRCA1* and *BRCA2* variants using a combination of mRNA splicing assays and multifactorial likelihood analysis [Goldgar et al., 2004]. The variants investigated ranged from those within consensus 5' GT and 3' AG dinucleotides to variants located 53 bp into the intron.

Materials and Methods

Twenty-six *BRCA1/2* variants were included in this study: seven were identified in seven families from Australia and 19 variants from 22 families from the United States (see Supp. Table S1 for HGVS and BIC nomenclature with nucleotide numbering starting at the first transcribed base of *BRCA1* [GenBank NM_007294.2] and *BRCA2* [NM_000059.1]). The RNA used for in vitro analysis was derived from lymphoblastoid cell lines (LCLs). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence.

Subject Selection

Probands from Australian families were ascertained as eligible for research by the Kathleen Cunningham Foundation Consortium for Research into Familial Breast Cancer (kConFab, <http://www.kconfab.org/Index.shtml>), a repository of genetic, epidemiological, medical and psychosocial data, and biospecimens available to researchers for investigation into the familial aspects of breast cancer. Probands from U.S. families were identified in family cancer clinics and were enrolled in studies of *BRCA1* and *BRCA2* unclassified variants at the Mayo Clinic, following detection of an unclassified sequence variant during clinical testing. Additional US family members were enrolled separately and DNA samples extracted from blood were sequenced to verify the presence of the specific family variant. All participants provided written informed consent.

Bioinformatic Analysis of Variant Sequences

We utilized Human Splicing Finder version 2.4 (www.umd.be/HSF/), which evaluates splicing signals present in any human gene by using matrices to predict 5' and 3' splice sites and splice regulatory sites using different algorithms, including Human Splicing Finder matrices and MaxEntScan [Desmet et al., 2009; Yeo and Burge, 2004]. Variant nomenclature was input into HSF as is shown in Table 1, apart from insertion or deletion variants, for which the exact sequence was input consecutively for wild-type and variant sequences. We determined the difference between variant and wild-type output scores as a

proportion of wild-type scores for HSF matrices and MaxEntScan. NNSplice (http://www.fruitfly.org/seq_tools/splice.html) [Reese et al., 1997] was used to assess the effect of variants on native donor and acceptor sequences only, with exact sequence input consecutively for wild-type and variant sequences, and a minimum score of 0.1 set for both 5' and 3' splice sites. For variants that were assessed on the basis of an interruption at the intron–exon junction bracketed percentages refer to the variation between the variant score and the known intron–exon boundary. For variants assessed on the basis of the potential to create a de novo splice site, bracketed percentages in Table 1 refer to the variation between the wild-type sequence and the variant sequence at that position.

Identification of Splicing Aberrations

RNA was extracted from cyclohexamide-treated and untreated LCLs, using RNeasy Mini Kit (Qiagen, Doncaster, Victoria, Australia) according to the manufacturer's instructions (Australian variants), or Trizol (Invitrogen, Carlsbad, CA) (US variants). Cycloheximide (1 mg/ 10 ml) was added to the medium 4 hr before harvesting the cells to prevent degradation of unstable transcripts by nonsensemediated decay (NMD). Each RNA sample was treated with DNase to reduce DNA contamination using DNA-free kit (Ambion, Austin, TX). Complementary DNA (cDNA) was synthesized from 500 ng of DNase-treated RNA using Superscript III First-Strand Synthesis System (Invitrogen). Cdna was used as a template in PCR reactions with specific primers targeting the potential splice-sites (Supp. Table S2). Each RT-PCR analysis included a set of 14 nonvariant control LCLs.

PCR for Australian and US variants were performed under the conditions presented in Supp. Table S2. PCR products were purified using QIAquick PCR Purification Kit (Qiagen) and sequenced using Big-Dye Terminator version 3.1 sequencing chemistry under the following conditions: 96°C for 1 min, followed by 25 cycles of 96°C for 10 sec, 50°C for 5 sec, and 60°C for 4 min. Samples were run on an ABI 377 sequencer (Applied Biosystems, Bedford, MA). Each experiment was repeated by regrowing the LCL from frozen stock for each sample assayed, for repeat cDNA synthesis, RT-PCR, gel electrophoresis, and sequencing. PCR products from *BRCA1* c.4185+9C>T, *BRCA1* c.5194–12G>A, and *BRCA2* c.516+1G>T were cloned using pGEM®-T Vector (Promega, Madison, WI) and verified by sequencing. All RT-PCR studies were performed in duplicate.

A touchdown protocol was used to increase the specificity and sensitivity in PCR amplification of the samples carrying the c.593+4A>G and the c.4675+1G>A variants [Korbie and Mattick, 2008]. PCR products were separated on 2% agarose gels. PCR products from *BRCA2* c.7618–1G>A, *BRCA2* c.8953+1G>T, *BRCA1* c.4675+1G>A and c.593+4A>G were cloned using Topo TA cloning Kit (Invitrogen) for verification by sequencing.

Multifactorial Likelihood Classification

Likelihood ratios for segregation were derived by Bayes factor analysis as described previously [Spurdle et al., 2008b; Thompson et al., 2003]. Information for the pathology component of the model was derived from pathology reports or pathology review of breast tumor sections performed as part of kConFab core activities or by two pathologists (SRL or

LMDS) specifically for this project. Following the methods described in Spurdle et al. [2008a], estimates of the likelihood of *BRCA1* or *BRCA2* mutation status were derived for variant carriers using the available information on histopathologic features reported to be associated with *BRCA1* mutation status (cytokeratin [CK] 5/6, CK14, and estrogen receptor) and *BRCA2* mutation status (percentage of tubule formation) [Lakhani et al., 2002; Spurdle et al., 2008b]. Likelihood ratios reported for family history were based on the statistical model developed by Easton et al. [2007], which was derived from the Myriad Genetics Laboratories dataset of 70,000 *BRCA1* and *BRCA2* tests. Likelihood scores for co-occurrence with a pathogenic mutation were derived as previously described [Goldgar et al., 2004; Spurdle et al., 2008b], from the same dataset. Probabilities were derived for each of the components included in the study, under the assumption that each factor was statistically independent. The individual likelihood ratios were multiplied to calculate an overall multifactorial likelihood ratio. Bayes rule was then used to calculate a posterior probability that the variant was pathogenic from the multifactorial likelihood ratio and the prior probability (see below). Variants were classified according to the five class IARC quantitative scheme [Plon et al., 2008], based on the posterior probability.

Note: Prior probabilities for intronic variants were derived from previous heterogeneity analyses of *BRCA1/2* variants by domain and position (see Table 5 from [Easton et al., 2007]). In brief, a consecutive series of *BRCA1/2* sequence results from a large dataset acquired through Myriad Genetic Laboratories was stratified by mutation type and position, specifically: Consensus Splice Site, missense variant or in-frame insertion-deletion, or “Other,” which included intronic variants outside the consensus splice sites, 5’UTR, and other variants that did not fit the previous two categories. In parallel, using the same dataset, the personal or family history profile for known pathogenic mutations in *BRCA1/2* was compared to noncarriers to determine family history profiles that could best predict mutation status using logistic regression. The predicted probabilities of each individual found to carry one of the variants described above (consensus splice site, missense, in-frame insertion-deletion, or “Other,” was then used to construct a likelihood ratio of the probability of the observed personal and family history given the variant was pathogenic compared to that under the hypothesis that it was a neutral variant. The likelihood ratios of each variant within a class of interest (e.g., all consensus splice site variants) were used to obtain the maximum likelihood estimate of the proportion of variants in the class that are pathogenic. This proportion then serves as the prior probability that a new variant observed in that class is pathogenic. From the results from this analysis (shown in Table 5 of the publication by Easton et al. [2007]), variants in the “other” class (including intronic variants outside the donor or acceptor dinucleotide) have a prior probability of 0.26 (0.15–0.39), whereas variants in the donor or acceptor dinucleotide sites have a prior probability of 1.0 (0.91–1.0). In this study we conservatively assigned a prior probability of 0.96 for consensus site changes, based on the midpoint of the estimate ranges (0.9–1.0).

Results

In Silico and In Vitro Analysis

Variants were initially analyzed using bioinformatic predictions before assessment in vitro using RNA splicing analysis. Results from Human Splicing Finder version 2.4 (www.umd.be/HSF/), which incorporates MaxEntScan, are presented in Table 1, with a summary of the results from in vitro analysis. The algorithms generate consensus values of potential splice sites and search for branch points. Variants within the algorithm window may be assessed for destruction of existing splice sites or creation of de novo splice sites.

Six variants altered the consensus splicing acceptor or donor dinucleotides. Each variant resulted in two different aberrant products, including different combinations of whole exon skipping events and partial exon deletions (Table 1, and Figs. 1 and 2). One of these aberrations was an in-frame transcript. *BRCA1* c.4675+1G>A resulted in exon 15 skipping, observed as a major band by gel electrophoresis (Fig. 1C). *BRCA2* c.516+1G>T (Fig. 2B) was associated with two exon skipping events: exon 6 (observed as a major band) and exon 5/6 (observed as a minor band), in addition to barely detectable levels of the naturally occurring isoform (39bp of exon 6 with exon 7) that was observed in controls. Similarly, *BRCA2* c.7007+1G>C resulted in two aberrant transcripts (exon 12/13, observed as a minor band, and exon 13), both of which are predicted to encode truncated proteins (Fig. 2C). *BRCA2* c.7618-1G>A did not produce a single whole exon deletion: this variant resulted in two aberrant products due to the recognition of alternative sites within exon 16 (observed as a minor band) and a second within exon 17 (Fig. 2D). *BRCA2* c.8953+1G>T created an exon 22 transcript and a second transcript deleting only 31 bp of exon 22 (Fig. 2E). *BRCA2* c.9257-1G>C resulted in the production of an exon 25 transcript, and a second in-frame aberrant transcript skipping 27 bases of exon 25 (Fig. 2F). In each case where a variant occurred at the 5' end of the intron and exon skipping resulted, the preceding exon was lost.

Another four variants created whole exon deletions attributable to disruption of intronic consensus splice-site recognition sequences (*BRCA1* c.593+4A>G, Fig. 1A; *BRCA1* c.5467+5G>C, Fig. 1E; *BRCA2* c.9501+3A>T, Fig. 2G, *BRCA2* c.426-12_8del5, Fig. 2A). In each instance the whole exon deletion resulted in loss of the open reading frame and a transcript encoding a truncated protein. Both *BRCA1* c.5467+5G>C and *BRCA2* c.9501+3A>T appeared to be associated with major aberrant transcripts, whereas this was not as obvious for the exon 5 deletion associated with *BRCA2* c.426-12_8del5.

In silico prediction suggested that de novo splice sites were created by *BRCA1* c.4484+2ins8 and *BRCA1* c.5194-12G>A (Table 1). Aberrant splice products were not resolved by gel electrophoresis of RT-PCR products (Fig. 1B and D), but the alterations predicted were insertions of only 8 and 10 nucleotides in size, respectively. Direct sequencing of the amplified product clearly showed an eight-nucleotide insertion associated with *BRCA1* c.4484+2ins8. Cloning and sequencing of RT-PCR products from *BRCA1* c.5194-12G>A identified a 10-nucleotide insertion. A third variant predicted to create a small insertion (*BRCA1* c.4185+9C>T) was also analyzed by sequencing of cloned RT-PCR fragments. Recombinant DNA from nine of 10 colonies showed wild-type sequence, and one clone

harbored a deletion of exon 13 that was not observed by gel electrophoresis. It is probable that the exon 13 deletion is a low abundance isoform unrelated to the variant under study, particularly because the *BRCA1* c.4185+9C>T variant occurs at the 5' end of intron 12 and it is unclear how this variant would be associated with exon13 transcript production.

In vitro results were also compared to bioinformatic scores from all three programs to determine if application of post hoc thresholds improved sensitivity and specificity of predictions. As shown in Table 1, interruption of intron-exon junction was best predicted using a threshold of -5% variation for HSF (as used previously) [Walker et al., 2010], -15% for MaxEntScan, and -10% for NNSplice. Eleven of 12 variants assessed for interruption of intron-exon junction had predictions from all three programs that were consistent with in vitro results, and the final variant *BRCA2* c.7435+6G>A was predicted to interrupt the consensus site with -16.5% variation for MaxEntScan only (and not the other two programs used). Results for all 14 variants assessed for creation of a de novo acceptor or donor were predicted accurately by applying very stringent post hoc thresholds of a minimum increase of 60% for HSF and 200% for MaxEntScan.

Multifactorial Likelihood Analysis and Comparison to Interpretation of Pathogenicity Based on In Vitro Results

Multifactorial likelihood analysis was conducted on 15 families for whom data were available (Table 2). Variants were categorized based on the final posterior probability, according to the classification system for sequence variants proposed by the 2008 IARC working group on unclassified genetic variants [Plon et al., 2008]. Seven variants had posterior probabilities >0.95 that categorized them as likely pathogenic or pathogenic (class 4 and 5, respectively) (Table 2). This included all variants occurring at the 5' or 3' splice consensus dinucleotide, reflecting the assignment of a prior probability of 0.96 and evidence from at least one component in the model. *BRCA1* c.5194-12G>A also reached a posterior probability of >0.99 (class 5, pathogenic), with substantial segregation and family history likelihood scores of 532.1 and 363.1, respectively (Table 2).

Five variants were classified as class 3 (uncertain; posterior probability of 0.05-0.949) and three variants were considered class 2 (likely not pathogenic, posterior probability of 0.001-0.049) (Table 2). It is possible that four of these eight class 2/class 3 variants could be raised to at least class 4 (likely pathogenic), or possibly class 5 (pathogenic), on the basis of interpretation of qualitative splicing data according to simplistic application of guidelines published for clinical interpretation of splicing aberrations [Spurdle et al., 2008a]. Namely, a minor or major truncating aberration that was absent in controls was detected in vitro for *BRCA1* c.4484+2ins8, *BRCA1* c.593+4A>G, and *BRCA1* c.5467+5G>C; *BRCA2* C.9501+3A>T and *BRCA2* c.426-12_8del5 (for which multifactorial data was unavailable).

Discussion

This study has assessed the possible pathogenicity of 26 intronic *BRCA1* or *BRCA2* variants using mRNA based splicing assays, and where possible has compared these results to predictions from multifactorial likelihood analysis. Aberrant splicing products that encode truncated proteins were observed for 12 of these variants. Where there was a suggestion of

upregulation of alternative splicing, it occurred in addition to clearly observed, variant-specific truncating aberrations. Multifactorial analysis was possible for 15 of the variants assayed in vitro, including 11 of 12 variants with associated aberrant splice products.

Multifactorial results provided a pathogenic or likely pathogenic classification for seven of these 11 variants. Six were located in the consensus dinucleotides, confirming the very high likelihood that variants in these positions will alter splicing and are associated with disease. The remaining variant *BRCA1* c.5194-12G>A occurred 12 nucleotides within an intron, a depth from the exonic boundary that is not routinely considered as potentially pathogenic in the clinical setting. Unfortunately, given that the coverage of intronic sequences by clinical sites can allegedly vary from ~20 nucleotides to more than 100, the frequency of functionally relevant downstream variants in multiple case cancer families is difficult to assess at this time. However, there are two *BRCA1/2* variants at a similar depth that are currently reported on BIC (<http://research.nhgri.nih.gov/bic/>) to be of clinical significance (*BRCA1* c.213-11T>G [Friedman et al., 1995] and *BRCA1* c.213-12A>G [Hoffman et al., 1998]), and a few reports in the literature (*BRCA1* c.5406+32A>T [Matsushima et al., 1995], and *BRCA1* c.81-9C>G and *BRCA2* c.9649-12T>G [Joosse et al, 2010]). These findings together suggest that it is important to investigate the potential of intronic variants to enhance existing cryptic motifs, or to create de novo splicing sites, and thus alter cancer risk.

There were four variants with associated aberrant transcripts but classified as uncertain (*BRCA1* c.4484+2ins8, *BRCA2* c.9501+3A>T) or likely not pathogenic (*BRCA1* c.593+4A>G, *BRCA1* c.5467+5G>C) on the basis of the posterior probability of pathogenicity, with the posterior probability of 0.0447 for *BRCA1* c.5467+5G>C close to the upper boundary of 0.049 currently designated for class 2 variants. These four variants were all located outside the consensus dinucleotides, and a combination of several factors may partly explain the apparent disparity between evidence derived from experimental methods and the multifactorial model. First, there was unfortunately a paucity of information for multifactorial analysis of these variants, with a maximum of two genotypes per family available for segregation analysis, and tumor histopathology scores could be derived from only one affected carrier in each family. Family history and co-occurrence data were available for only two of these variants. It will thus be necessary to revisit multifactorial analysis for these variants, and such efforts will be facilitated by the recent establishment of the ENIGMA consortium (Evidence-based Network for the Interpretation of Germline Mutant Alleles <http://enigmaconsortium.org>) for large-scale collaborative studies of *BRCA1* and *BRCA2* sequence variants. Second, the prior probability of pathogenicity of 0.26 applied for nonconsensus intronic variants is rather nonspecific: it covers the general class termed “other” that includes all intronic variants outside the consensus dinucleotides and untranslated regions [Easton et al., 2007], and it currently does not incorporate bioinformatic prediction to improve discrimination between functionally relevant and functionally irrelevant intronic variants. Prior probabilities for this class of variant are likely to be enhanced by integration of in silico prediction methods that assess likelihood of splicing disruption, and/or effects on transcription and mRNA stability. The approach to calibrating prior probabilities for in silico splice predictions need not differ conceptually from that used

to calibrate the A-GVGD methods currently used to assign prior probabilities to missense variants [Spurdle et al., 2008b], which are based on position in the protein, evolutionary sequence conservation, and the physicochemical difference between the missense residue and the range of variation observed at its position in the protein, without directly assessing abrogation of function due to missense substitution. Third, we cannot exclude the possibility that in vitro assays conducted in LCLs may not accurately reflect splicing aberrations or levels of aberrations in target tissue (breast or ovary). Although none of the splicing aberrations identified in the mRNA from LCLs of variant carriers reported in Table 1 have been identified as naturally occurring isoforms in breast tissue [Munnes et al., 2000], we do note that the exon 9 deletion observed to occur in the carrier of *BRCA1* c.593+4A>G has previously been reported to occur in lymphocytes [Munnes et al., 2000]. This might suggest that the exon 9 deletion observed experimentally reflects tissue mRNA source and not variant status, and explain the disparity between the perceived splice aberration and the low posterior probability of pathogenicity for this variant. Nevertheless, our RT-PCR assays on mRNA from LCLs did not detect the exon 9 deletion in a large set of nonvariant carrying controls and we thus still cannot exclude the possibility that *BRCA1* c.593+4A>G may also result in expression of the exon 9 deletion in breast tissue that does not normally express this transcript. Future analysis of paired normal and cancer breast tissue from *BRCA1* c.593+4A>G carriers may help assess the relationship between *BRCA1* c.593+4A>G, transcript expression, and tumor development, when there is consensus about the appropriate reference normal breast tissue cell type (e.g., epithelial, myoepithelial, etc.) to use for such comparative assays. Continued attempts to assess risk via multifactorial approaches are thus essential to clarify risk associated with this variant.

Only one other variant assayed, *BRCA2* c.426–12_8del5, showed evidence for an aberration that might be considered clinically significant on the basis of the molecular defect observed. As for the four other variants described above, multifactorial analysis using multiple points of evidence will be helpful to confirm the clinical relevance. The resources available through ENIGMA (<http://www.enigmaconsortium.org/>), a consortium specifically established to encourage collaborative studies of variants in *BRCA1/2*, will be integral to maximize collection of family and biological material for future classification of these and other variants.

Similar to the widely accepted classification of nonsense or stop mutations as pathogenic on the basis of sequence information alone, there is some rationale to suggest that all of the variants observed to produce splicing aberrations could be considered pathogenic on the basis of the splicing results alone. Indeed, a five-class qualitative classification scheme for interpretation of in vitro splicing assays suggested that a variant allele that produces major transcript(s) carrying a premature stop codon or an in-frame deletion disrupting known functional domains could reasonably be considered to be class 5 pathogenic [Goldgar et al., 2008]. At this point in time, however, there has been no clarification as to what defines “major,” and what methods might best measure the abundance of aberrant isoforms relative to full-length transcript. The conservative viewpoint would thus be that (1) the subset of variants with splicing aberrations but conflicting or insufficient supporting evidence from multifactorial analysis (class 1 to class 3) undergo further quantitative studies to clarify the

functional consequences of the sequence variant in question, alongside extended genetic and pathology studies, and (2) if necessary, be raised to class 3 until the disparity is resolved. However, it is interesting to note that *BRCA1* c.593+4A>G was shown to produce relatively low levels of aberrant transcripts using a semiquantitative approach, and this variant does have the lowest posterior probability of all variants identified to be spliceogenic in this study.

It is also possible that other variants with no associated splicing aberrations may result in allelic imbalance created by transcriptional dysregulation, mRNA instability or mRNA processing, a possibility that will require quantitative assays of common exonic polymorphisms to assess [Caux-Moncoutier et al., 2009]. Although the relationship between percent reduction in wild-type transcript and risk is not yet established, it would be logical to assume at minimum that failure to express full-length transcript by the variant allele should be interpreted as total loss of native expression of this allele, with similar consequences to truncating variants that are considered pathogenic on the basis of sequence information alone. Such quantitative studies and further multifactorial studies will be important to clarify their role in disease. Once this has been established, it should be possible to incorporate such information in multifactorial analysis as a separate component, as has been proposed for assays of protein function [Couch et al., 2008].

Overall, the combination of splice prediction programs we used accurately predicted all aberrations detected in vitro. We thus suggest that bioinformatic predictions could be incorporated into algorithms that estimate the prior probability of pathogenicity for intronic variants, to improve multifactorial likelihood analysis. As a technical point, we show the value of bioinformatic prediction in experimental design, specifically the selection of appropriate assay methods to detect aberrations caused by *BRCA1* c.4484+2ins8 and *BRCA1* c.5194–12G>A. Moreover, we provide information from our post hoc analysis to suggest thresholds to maximize sensitivity and specificity. Interpretation of scores is challenging for creation of donor and acceptors, because this requires considering change in score at the variant position (does the variant actually create a new site?), and how this variant site score compares to the consensus site score. Results from this dataset suggested that de novo sites occurred when there was a very large increase in score at the variant position, and a variant score that was similar but not necessarily greater than the consensus score. This is exemplified by *BRCA1* c.5194–12G>A, which resulted in an increased score of 64% for HSF and 237% for MaxEntScan, and a score of 4.59 versus proximal consensus site score of 9.36. Analysis of further large datasets will be helpful to refine these suggested thresholds.

Overall, these findings emphasize the need to apply a comprehensive bioinformatic and experimental approach to assess the impact of intronic variants on native splicing and thus assist in their clinical classification. Importantly, our data suggest that such assays should be extended beyond the consensus splice sites, and may be relevant to variants up to 12 nucleotides into the intron. Although a recent study reported that deeply intronic variants of *BRCA1/2* were very unlikely to be associated with splicing aberrations [Caux-Moncoutier et al., 2009], we have shown that some intronic variants may create de novo splice sites, and as exemplified by *BRCA1* c.5194–12G>A, are associated with the same level of risk as seen for known pathogenic mutations. In addition, the comparison of results from in vitro assays to

bioinformatic predictions will allow accumulation of information to refine the bioinformatic techniques for application in future clinical evaluation of intronic variants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge the participation of the families concerned. We thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow-Up Study for their contributions to this resource, and the many families who contribute to kConFab. We thank Maxime Vallee for helpful discussions. We also acknowledge Linda Wadum, Kiley Johnson, Jennifer Mentlick, and Mary Karaus for their efforts to recruit the US-based families to these studies. We thank Myriad Genetic Laboratories for information used to derive family history scores and investigate co-occurrence of variants with pathogenic mutations. P.W. was awarded a scholarship by the QIMR Higher Degrees Committee. A.B.S. is an NHMRC Senior Research Fellow, L.Da.S. was supported by a fellowship from the Ludwig Institute for Cancer Research. L.W. is a John Gavin postdoctoral fellow. L.G. is supported by a fellowship from the Komen Foundation for the Cure.

Contract grant sponsor: The National Health and Medical Research Council (NHMRC); Contract grant number: ID442970; Contract grant sponsor: NIH; Contract grant number: CA116167; Contract grant sponsor: NIH Breast Cancer Specialized Program of research Excellence (SPORE); Contract grant number: P50 CA116201; Contract grant sponsors: National Breast Cancer Foundation (to kConFab); The NHMRC (to kConFab); The Queensland Cancer Fund (to kConFab); The Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia (to kConFab); The Cancer Foundation of Western Australia (to kConFab); NHMRC; Contract grant numbers: 145684; 288704 (to the kConFab Clinical Follow-Up Study).

References

- Caux-Moncoutier V, Pages-Berhouet S, Michaux D, Asselain B, Castera L, De Pauw A, Buecher B, Gauthier-Villars M, Stoppa-Lyonnet D, Houdayer C. Impact of BRCA1 and BRCA2 variants on splicing: clues from an allelic imbalance study. *Eur J Hum Genet.* 2009; 17:1471–1480. [PubMed: 19471317]
- Chen X, Truong TTN, Weaver J, Bove BA, Cattie K, Armstrong BA, Daly MB, Godwin AK. Intronic alterations in BRCA1 and BRCA2: effect on mRNA splicing fidelity and expression. *Hum Mutat.* 2006; 27:427–435. [PubMed: 16619214]
- Chenevix-Trench G, Healey S, Lakhani S, Waring P, Cummings M, Brinkworth R, Deffenbaugh AM, Burbidge LA, Pruss D, Judkins T, Scholl T, Bekessy A, Marsh A, Lovelock P, Wong M, Tesoriero A, Renard H, Southey M, Hopper JL, Yannoukakos K, Brown M, Easton D, Tavtigian SV, Goldgar D, Spurdle AB. kConFab Investigators. Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer Res.* 2006; 66:2019–2027. [PubMed: 16489001]
- Claes K, Poppe B, Machackova E, Coene I, Foretova L, De Paepe A, Messiaen L. Differentiating pathogenic mutations from polymorphic alterations in the splice sites of BRCA1 and BRCA2. *Genes, Chromosomes Cancer.* 2003; 37:314–320. [PubMed: 12759930]
- Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, de Wind N. Assessment of functional effects of unclassified genetic variants. *Hum Mutat.* 2008; 29:1314–1326. [PubMed: 18951449]
- Davis RL, Homer VM, George PM, Brennan SO. A deep intronic mutation in FGB creates a consensus exonic splicing enhancer motif that results in afibrinogenemia caused by aberrant mRNA splicing, which can be corrected in vitro with antisense oligonucleotide treatment. *Hum Mutat.* 2009; 30:221–227. [PubMed: 18853456]
- Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 2009; 37:e67. [PubMed: 19339519]

- Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro AN, Iversen ES, Couch FJ, Goldgar DE. 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.* 2007; 81:873–883. [PubMed: 17924331]
- Friedman LS, Szabo CI, Ostermeyer EA, Dowd P, Butler L, Park T, Lee MK, Goode EL, Rowell SE, King MC. Novel inherited mutations and variable expressivity of BRCA1 alleles, including the founder mutation 185delAG in Ashkenazi Jewish families. *Am J Hum Genet.* 1995; 57:1284–1297. [PubMed: 8533757]
- Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Hum Mutat.* 2008; 29:1265–1272. [PubMed: 18951437]
- Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro ANA, Tavtigian SV, Couch FJ. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet.* 2004; 75:535–544. [PubMed: 15290653]
- Harland M, Mistry S, Bishop DT, Bishop JA. A deep intronic mutation in CDKN2A is associated with disease in a subset of melanoma pedigrees. *Hum Mol Genet.* 2001; 10:2679–2686. [PubMed: 11726555]
- Hoffman JD, Hallam SE, Venne VL, Lyon E, Ward K. Implications of a novel cryptic splice site in the BRCA1 gene. *Am J Med Genet.* 1998; 80:140–144. [PubMed: 9805131]
- Homolova K, Zavadakova P, Doktor TK, Schroeder LD, Kozich V, Andresen BS. The deep intronic c. 903+469T>C mutation in the MTRR gene creates an SF2/ASF binding exonic splicing enhancer, which leads to pseudoexon activation and causes the cblE type of homocystinuria. *Hum Mutat.* 2010; 31:437–444. [PubMed: 20120036]
- House AE, Lynch KW. Regulation of alternative splicing: more than just the ABCs. *J Biol Chem.* 2008; 283:1217–1221. [PubMed: 18024429]
- Joosse SA, Brandwijk KI, Devilee P, Wesseling J, Hogervorst FB, Verhoef S, Nederlof PM. Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH. *Breast Cancer Res Treat.* 2010 Jul 8. Epub ahead of print.
- Korbie DJ, Mattick JS. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc.* 2008; 3:1452–1456. [PubMed: 18772872]
- Kwong A, Wong LP, Chan KY, Ma ES, Khoo US, Ford JM. Characterization of the pathogenic mechanism of a novel BRCA2 variant in a Chinese family. *Fam Cancer.* 2008; 7:125–133. [PubMed: 17657584]
- Lakhani SR, Jacquemier J, Sloane JP, Gusterson BA, Anderson TJ, van de Vijver MJ, Farid LM, Venter D, Antoniou A, Storer-Isser A, Smyth E, Steel CM, Haite N, Scott RJ, Goldgar D, Neuhausen S, Daly PA, Ormiston W, McManus R, Scherneck S, Ponder BA, Ford D, Peto J, Stoppa-Lyonnet D, Bignon YJ, Struwing JP, Spurr NK, Bishop DT, Klijn JG, Devilee P, Cornelisse CJ, Lasset C, Lenoir G, Barkardottir RB, Egilsson V, Hamann U, Chang-Claude J, Sobol H, Weber B, Stratton MR, Easton DF. Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst.* 1998; 90:1138–1145. [PubMed: 9701363]
- Lakhani SR, Van De Vijver MJ, Jacquemier J, Anderson TJ, Osin PP, McGuffog L, Easton DF. 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.* 2002; 20:2310–2318. [PubMed: 11981002]
- Lopez AJ. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu Rev Genet.* 1998; 32:279–305. [PubMed: 9928482]
- Matsushima M, Kobayashi K, Emi M, Saito H, Saito J, Suzumori K, Nakamura Y. Mutation analysis of the BRCA1 gene in 76 Japanese ovarian cancer patients: four germline mutations, but no evidence of somatic mutation. *Hum Mol Genet.* 1995; 4:1953–1956. [PubMed: 8595420]
- Munnes M, Zuther I, Schmitz B. A novel insertional mutation and differentially spliced mRNAs in the human BRCA1 gene. *Gene Funct Dis.* 2000; 1:38–47.
- Pagani F, Baralle FE. Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet.* 2004; 5:389–396. [PubMed: 15168696]

- Pagani F, Buratti E, Stuani C, Bendix R, Dork T, Baralle FE. A new type of mutation causes a splicing defect in ATM. *Nat Genet.* 2002; 30:426–429. [PubMed: 11889466]
- Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat.* 2008; 29:1282–1291. [PubMed: 18951446]
- Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol.* 1997; 4:311–323. [PubMed: 9278062]
- Rio Frio T, McGee TL, Wade NM, Iseli C, Beckmann JS, Berson EL, Rivolta C. A single-base substitution within an intronic repetitive element causes dominant retinitis pigmentosa with reduced penetrance. *Hum Mutat.* 2009; 30:1340–1347. [PubMed: 19618371]
- Schwartz GF, Hughes KS, Lynch HT, Fabian CJ, Fentiman IS, Robson ME, Domchek SM, Hartmann LC, Holland R, Winchester DJ. Proceedings of the international consensus conference on breast cancer risk, genetics, & risk management, April, 2007. *Cancer.* 2008; 113:2627–2637. [PubMed: 18853415]
- Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat.* 2008a; 29:1304–1313. [PubMed: 18951448]
- Spurdle AB, Lakhani SR, Healey S, Parry S, Da Silva LM, Brinkworth R, Hopper JL, Brown MA, Babikyan D, Chenevix-Trench G, Tavtigian SV, Goldgar DE. kConFab Investigators. Clinical classification of BRCA1 and BRCA2 DNA sequence variants: the value of cytokeratin profiles and evolutionary analysis—a report from the kConFab Investigators. *J Clin Oncol.* 2008b; 26:1657–1663. [PubMed: 18375895]
- Srebrow A, Kornblihtt AR. The connection between splicing and cancer. *J Cell Sci.* 2006; 119(Pt 13): 2635–2641. [PubMed: 16787944]
- Tavtigian SV, Greenblatt MS, Goldgar DE, Boffetta P. Assessing pathogenicity: overview of results from the IARC Unclassified Genetic Variants Working Group. *Hum Mutat.* 2008; 29:1261–1264. [PubMed: 18951436]
- 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.* 2003; 73:652–655. [PubMed: 12900794]
- Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, Bakker E, van Asperen CJ, Devilee P. Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. *Hum Mutat.* 2009; 30:107–114. [PubMed: 18693280]
- Walker LC, Whiley PJ, Couch FJ, Farrugia DJ, Healey S, Eccles DM, Lin F, Butler SA, Goff SA, Thompson BA, Lakhani SR, Da Silva LM, Tavtigian SV, Goldgar DE, Brown MA, Spurdle AB. kConFab Investigators. Detection of splicing aberrations caused by BRCA1 and BRCA2 sequence variants encoding missense substitutions: implications for prediction of pathogenicity. *Hum Mutat.* 2010; 31:E1484–E1505. [PubMed: 20513136]
- Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 2004; 11:377–394. [PubMed: 15285897]
- Yu Y, Maroney PA, Denker JA, Zhang XH, Dybkov O, Luhrmann R, Jankowsky E, Chasin LA, Nilsen TW. Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition. *Cell.* 2008; 135:1224–1236. [PubMed: 19109894]

Figure 1.

Instances of aberrant splicing arising from *BRCA1* intronic sequence variants detected by RT-PCR. VC+ and VC- represent RT-PCR on mRNA from cycloheximide variant carrier treated and untreated LCLs, respectively, and nonvariant carrying controls are represented by C- or C+ according cycloheximide treatment status. **A:** *BRCA1* c.593+4A>G displays a minor exon9 product confirmed by subcloning and sequencing of PCR products. The exon9/10 naturally occurring isoform is represented by the 362-bp fragment. **B:** *BRCA1* c.4484+2ins8 resulted in the comigration of two bands: a full-length product and a second containing an eight nucleotide insertion, characterized by directly sequencing the PCR product. **C:** *BRCA1* c.4675+1G>A exhibits an exon 15 deletion. **D:** *BRCA1* c.5194-12G>A results in a predicted insertion of 10 nucleotides into exon 20 undetectable by gel electrophoresis but subsequently identified by cloning and sequencing. **E:** *BRCA1* c.5467+5G>C is associated with exon 23. [Color figures can be viewed in the online issue, which is available at www.wiley.com/humanmutation.]

Figure 2. Instances of aberrant splicing arising from *BRCA2* intronic sequence variants detected by RT-PCR. VC+ and VC- represent RT-PCR on mRNA from cycloheximide variant carrier treated and untreated LCLs, respectively, and nonvariant carrying controls are represented by C- or C+ according cycloheximide treatment status. **A:** *BRCA2* c.426-12_8del5 is associated with exon 5. The exon 7/39 bp of exon 6 common isoform is also observed in controls. **B:** *BRCA2* c.516+1G>T results in two aberrant splice products; exon 5/6 and exon 6. The exon 7/39 bp of exon 6 common isoform is also observed in controls. **C:** *BRCA2* c.7007+1G>C results in exon 12 and exon 12/13 aberrations. **D:** *BRCA2* c.7618-1G>A creates a 44 nucleotide exon 16 deletion and exon 16/69 bp of exon 17. **E:** *BRCA2* c.8953+1G>T creates a minor 31-bp nucleotide deletion of exon 22 and a major exon 22 splice product. **F:** *BRCA2* c.9257-1G>C creates an in-frame 27 nucleotide deletion from exon 25 and exon 25. **G:** *BRCA2* c.9501+3A>T is associated with exon 25. [Color figures can be viewed in the online issue, which is available at www.wiley.com/humanmutation.]

Table 1

Bioinformatic Prediction Scores and In Vitro Splicing Results

	Human splicing finder		MaxEntScan		NNsplice		Basis for prediction scores recorded	Changes consensus donor/acceptor?	In vitro splicing result ^b	Consistency of bioinformatic predictions with splicing ^c results applying post hoc assumptions
	Variant	Proximal consensus site	Variant	Proximal consensus site	Variant	Proximal consensus site				
BRCA1										
c.593-4A>G	90.5 (−8.4%)	98.84	8.56 (−19.8%)	10.67	0.89 (−11%)	1.00	Interruption of intron–exon junction	Yes	exon 9 deletion	Yes—3/3 predicted aberration
c.4185+9C>T	78.5 (51.9%)	85.5	2.99 (162.8%)	8.59	—	—	Creation of de novo donor	No	wild type	Yes—2/2 predicted no aberration
c.4484+2ins GGAAGGT	96.71 (296.26%)	96.71	10.60 (224.5%)	10.57	—	—	Creation of de novo donor	No	8-bp insertion into exon 14	Yes—2/2 predicted aberration
c.4675+1G>A	56.53 (−32.19%)	83.37	−1.33 (−119.44%)	6.84	<0.1 (<−43%)	0.44	Interruption of intron–exon junction	Yes	exon 15 deletion, 11-bp deletion from exon 15	Yes—3/3 predicted aberration
c.5152+10A>G	NSC	82.4	NSC	7.96	—	—	Creation of de novo donor	No	wild type	Yes—2/2 predicted no aberration
c.5194-12G>A	74.1 (64.1%)	86.33	4.59 (236.6%)	9.36	—	—	Creation of de novo acceptor	No	10-bp insertion into exon 20	Yes—2/2 predicted aberration
c.5278−14C>G	NSC	93.65	12.43 (−4.9%)	13.07	—	—	Creation of de novo acceptor	No	wild type	Yes—2/2 predicted no aberration
c.5406+9T>C	70.99 (2.87%)	83.72	0.66 (125.58%)	9.49	—	—	Creation of de novo donor	No	wild type	Yes—2/2 predicted no aberration
c.5467+5G>C	80.35 (−13.0%)	92.36	6.05 (−35.2%)	9.33	0.30 (−69.7%)	0.99	Interruption of intron–exon junction	Yes	exon 23 deletion	Yes—3/3 predicted aberration
c.81−20C>T	NSC	77.27	NSC	7.05	—	—	Creation of de novo acceptor	No	wild type	Yes—2/2 predicted no aberration
BRCA2										
c.426-12_8delGTTTT	56.74 (−28.65)	77.67	−6.72 (−174.75%)	8.99	0.79 (−20.2%)	0.99	Interruption of intron–exon junction	Yes	exon 5 deletion	Yes—3/3 predicted aberration
c.426−37T>A	NSC	79.52	1.28 (−46.67%)	8.99	—	—	Creation of de novo acceptor	No	wild type	Yes—2/2 predicted no aberration
c.516+1G>T	60.05 (−30.89%)	86.88	0.38 (−95.72%)	8.88	site abolished	0.98	Interruption of intron–exon junction	Yes	exon 5 deletion, exon 5 and 6 deletion	Yes—3/3 predicted aberration
c.682-12delTA	NSC	86.68	NSC	7.06	—	—	Creation of de novo acceptor	No	wild type	Yes—2/2 predicted no aberration
C.7007+1G>C	64.1 (−29.5%)	90.93	2.26 (−78.5%)	10.53	site abolished	0.99	Interruption of intron–exon junction	Yes	exon 13 deletion and exon 12 and 13 deletion	Yes—3/3 predicted aberration
c.7435+6G>A	74.56 (−0.5%)	74.96	4.71 (−16.5%)	5.64	0.39 (56%)	0.25	Interruption of intron–exon junction	No	wild type	Yes—1/3 predicted aberration
c.7435+53C>T	NSC	74.96	NSC	5.64	—	—	Creation of de novo donor	No	wild type	Yes—2/2 predicted no aberration
c.7436−14T>G	NSC	78.79	5.58 (8.14%)	5.16	—	—	Creation of de novo acceptor	No	wild type	Yes—2/2 predicted no aberration
c.7618-1G>A	50.16 (−36.6%)	79.1	−1.63 (−122.9%)	7.11	site abolished	0.69	Interruption of intron–exon junction	Yes	45-bp deletion from exon 16, exon 16 and 69-bp deletion from exon 17	Yes—3/3 predicted aberration
c.7805+6C>G	74.96 (0.3%)	74.72	5.9 (26.07%)	4.68	0.24 (0.04%)	0.25	Interruption of intron–exon junction	No	wild type	Yes—3/3 predicted no aberration
c.8487+8G>A	NSC	88.86	NSC	9.46	—	—	Creation of de novo donor	No	wild type	Yes—2/2 predicted no aberration

	Human splicing finder		MaxEntScan		NNSplice		Basis for prediction scores recorded	Changes consensus donor/acceptor?	In vitro splicing result ^b Intron 22 deletion, and 31-bp deletion of exon 22	Consistency of bioinformatic predictions with splicing ^c results applying post hoc assumptions
	Variant	Proximal consensus site	Variant	Proximal consensus site	Variant	Proximal consensus site				
C.8953+1G>T	73.16 (−26.8%)	100.0	2.35 (−78.4%)	10.86	site abolished	0.79	Interruption of intron–exon junction	Yes	exon 22 deletion, and 31-bp deletion of exon 22	Yes—3/3 predicted aberration
c.9257−16T>C	NSC	87.14	NSC	13.32			Creation of de novo acceptor	No	wild type	Yes—2/2 predicted no aberration
C.9257−1G>C	58.19 (−33.2%)	−87.14	5.26 (−60.5%)	13.32	site abolished	0.98	Interruption of intron–exon junction	Yes	exon 25 deletion, 27-bp deletion from exon 25 ^a	Yes—3/3 predicted aberration
c.9501+9A>C	73.2 (−10.7%)	96.31	0.89 (−28.8%)	10.28	—	—	Creation of de novo donor	No	wild type	Yes—2/2 predicted no aberration
c.9501+3A>T	91.28 (−5.2%)	96.31	4.36 (−57.59%)	10.28	0.26 (−73.7%)	0.99	Interruption of intron–exon junction	Yes	exon 25 deletion	Yes—3/3 predicted aberration

Bracketed percentages refer to the difference between variant and wild-type scores as a proportion of the wild-type score. NSC, no sites created (no scores provided by bioinformatic program output). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence. *BRCA1* (Gen Bank NM_007294.2) and *BRCA2* (NM_000059.1).

^aAll aberrations result in loss of the open reading frame except for the 27-bp deletion from exon 25 for *BRCA2* c.9257−1G>C. Variants noted in bold result in splicing aberrations.

^cFor each of the RT-PCR assays performed, a wild-type splice profile was observed for 14 controls confirmed not to carry the variant.

^cBased on post hoc assumptions as follows:

- Positive scores for interruption of intron–exon junction indicate that a variant is less likely to - interrupt a native consensus site.
- Negative scores for creation of de novo sites indicate that the variant is less likely to produce a de novo site.
- Variation in scores of at least −5% for HSF, −15% for MaxEntScan, and −10% for NNSplice were interpreted as loss of a site.

Minimum increase in scores of 60% for HSF and 200% for MaxEntScan were interpreted as gain of de novo acceptors or donor.

Table 2

BRCA1 and BRCA2 Variant Multifactorial Likelihood Analysis Results

	Prior probability of pathogenicity	Likelihood scores				Posterior probability of pathogenicity from multifactorial likelihood analysis	Evidence from multifactorial analysis ^b	In vitro splicing results ^c	Experimental method	Concordance between multifactorial and in vitro analysis of pathogenicity
		Segregation ^a	Tumor Histopathology	Family history	Co-occurrence with a deleterious mutation					
BRCA1										
c.593+4A>G	0.26	0.62	0.18	0.20	1.10	0.0244	0.009	Exon skipping	RT-PCR/subcloning	Truncating splice aberration without supporting multifactorial data
c.4185+9C>T	0.26	0.28	N/A	N/A	N/A	0.28	0.09	wild type	RT-PCR, sequencing/ subcloning	Unresolved multifactorial analysis
c.4484+2ins8	0.26	0.95	2.58	N/A	N/A	2.454	0.46	Intronic insertion	RT-PCR, sequencing/ subcloning	Truncating splice aberration without supporting multifactorial data
C.4675+1G>A	0.96	30.8	N/A	128.8	1.55	6109.2	0.99	Multiple aberrant transcripts	RT-PCR/subcloning	Concordant
c.5194-12G>A	0.26	532.1	0.76	363.1	1.26	184817.2	0.99	Intronic insertion	RT-PCR, sequencing/ subcloning	Concordant
c.5278-14C>G	0.26	N/A	N/A	0.14	1.58	0.22	0.07	wild type	RT-PCR	Unresolved multifactorial analysis
c.5467+5G>C	0.26	0.73	0.18	N/A	N/A	0.133	0.0447	Exon skipping	RT-PCR, sequencing	Truncating splice aberration without supporting multifactorial data
BRCA2										
c.516+1G>T	0.96	1.1	N/A	N/A	N/A	1.1	0.96	Multiple aberrant transcripts	RT-PCR, sequencing/ subcloning	Concordant
c.682-12delTA	0.26	N/A	N/A	1.15	1.15	1.32	0.32	wild type	RT-PCR	Unresolved multifactorial analysis
C.7007+1G>C	0.96	3.9	N/A	0.56	1.12	2.45	0.98	Multiple aberrant transcripts	RT-PCR/sequencing	Concordant
c.7436-14T>G	0.26	0.37	N/A	1.05	0.05	0.02	0.007	wild type	RT-PCR	Concordant
c.7618-1G>A	0.96	N/A	N/A	32.40	2.90	93.20	0.99	Multiple aberrant transcripts	RT-PCR/subcloning	Concordant
C.8953+1G>T	0.96	N/A	N/A	12.02	1.17	14.06	0.99	Multiple aberrant transcripts	RT-PCR/sequencing	Concordant
C.9257-1G>C	0.96	10.1	N/A	49.0	1.15	568.15	0.99	Multiple aberrant transcripts	RT-PCR/sequencing	Concordant
c.9501+3A>T	0.26	0.74	1.19	2.29	2.52	5.1	0.64	Exon skipping	RT-PCR/ sequencing	Truncating splice aberration without supporting multifactorial data

N/A, not applicable, no data available for this variant.

^aBased on Bayes scores from this study and published data from nonoverlapping families Chen et al. [2006].

^bAs per Plon et al. [2008] and Spurdle et al. [2008a].

^c Variants noted in bold result in splicing aberrations leading to a transcript encoding for a truncated protein. All in vitro splicing assays were carried out in cycloheximide treated LCLs. Multifactorial likelihood data was unavailable for *BRCA1* c.5152+10A>G, c.5406+9T>C, c.5278-14C>G, c.81-20C>T, *BRCA2* c.7435+6G>A, c.7436+53C>T, c.7805+6C>G, c.8487+8G>A, c.9257-16T>C, and c.9501+9A>C displaying a wild-type splice profile and *BRCA2* c.426-12_8del5 which produced a splicing aberration (Fig. 1F). Example calculation of posterior probability for *BRCA1* c.593+4A>G; the Prior Probability for c.593+4A>G is 0.26, because this is an intronic variant outside the consensus dinucleotide (see Materials and Methods). The Odds for Causality are 0.0244:1 (or equivalently 41:1 against causality), calculated as the product of the individual statistically independent components (LR Co-occurrence [1.10] × LR Tumor Histopathology [0.18] × LR Segregation [0.62] × LR Family History [0.20]). The Posterior Probability for *BRCA1* c.593+4A>G is 0.009 = Posterior Odds/ (Posterior Odds + 1), where the Posterior Odds = Prior Probability (0.26) × Odds for causality (0.0244) × (1/1-prior probability). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence. *BRCA1* (GenBank NM_007294.2) and *BRCA2* (NM_000059.1).