Alternative Splicing and Molecular Characterization of Splice Site Variants: \textit{BRCA1} c.591C\textgreater T as a Case Study

Vanessa Dosil, Alicia Tosar, Carmen Cañadas, Pedro Pérez-Segura, Eduardo Díaz-Rubio, Trinidad Caldés, and Miguel de la Hoya

**BACKGROUND:** Deleterious mutations in \textit{BRCA1} (breast cancer 1, early onset; MIM 113705) increase human breast and ovarian cancer [HB(O)C] risk; however, many variants cannot be readily classified as deleterious or neutral. Unclassified variants (UVs) pose serious problems in genetic counseling. RNA-splicing analysis is essential for the assessment of many UVs.

**METHODS:** Denaturing gradient gel electrophoresis was used to genotype the \textit{BRCA1} c.591C\textgreater T variant in 685 index cases of HB(O)C families, 326 sporadic breast cancer cases, and 450 healthy controls from Spain. In silico tools were used to predict the effect of the c.591C\textgreater T variant on splicing. In vitro splicing analysis was performed in 7 c.591C\textgreater T carriers and 10 noncarriers. cDNAs were PCR-amplified with primers designed to detect \textit{BRCA1} alternative splicing isoforms. The products were analyzed by capillary electrophoresis. Peak areas were used to quantify the relative abundance of each isoform. Sequencing through exonic single-nucleotide polymorphisms (SNPs) enabled us to discriminate wild-type and variant transcripts.

**RESULTS:** c.591C\textgreater T was detected in HB(O)C cases (1.5%), breast cancer cases (0.3%), and controls (0.9%). c.591C\textgreater T induced \textit{BRCA1} exon 9 skipping and modified the relative expression of Δ(9,10), Δ(9,10,11B), Δ11B, and full-length isoforms. The mean ratio of Δ(9,10) to the full-length isoform increased from 0.25 in noncarriers to 1.5 in carriers. The mean Δ(9,10,11B)/Δ11B ratio increased from 0.2 to 4. Overall expression levels of c.591C\textgreater T and wild-type alleles were similar.

**CONCLUSIONS:** Our data support a nonpathogenic role for the \textit{BRCA1} c.591C\textgreater T variant. Naturally occurring alternative splicing isoforms need to be considered when assessing the role of \textit{BRCA1} UVs on splicing. Deleterious germ-line mutations in \textit{BRCA1} (breast cancer 1, early onset; MIM 113705) confer a high risk of human breast and ovarian cancer [HB(O)C].\textsuperscript{3} \textit{BRCA1} genetic testing has revealed a large number of sequence variants, but only frameshifts and premature stop codons are generally assumed to be pathogenic without further analysis (1). Therefore, a growing number of variants are being detected that cannot be readily distinguished as either disease-causing or neutral. In the Breast Cancer Information Core (BIC) database, 691 such variants were listed as of October 2009 (2). These so-called unclassified variants (UVs) pose serious problems in genetic counseling, ranging from the impossibility of accurate risk assessment to psychological aspects (3).

Several approaches have been proposed to reclassify \textit{BRCA1} UVs into either disease-causing or neutral variants (1). Often, testing the effect of UVs on RNA splicing becomes essential to assess their pathogenicity. For instance, exon skipping or the activation of a cryptic splice site (leading to the deletion/insertion of exonic/intronic sequences) provides strong evidence that the variant is pathogenic. On the contrary, typical RNA processing suggests that the variant is neutral (4–14). Surprisingly, the existence of naturally occurring \textit{BRCA1} splicing isoforms (15) is mostly overlooked when the role of \textit{BRCA1} variants on splicing is tested.

We report the identification and characterization of allele-specific splicing (ASS) linked to the nonpathogenic \textit{BRCA1} variant c.591C\textgreater T. Our study illustrates the importance of taking into consideration naturally
occurring \textit{BRCA1} splicing isoforms when assessing the pathogenicity of \textit{BRCA1} UVs affecting splicing.

\section*{Materials and Methods}

\subsection*{Participant Recruitment}
Index cases of HB(O)C from 685 families were identified by the Familial Cancer Clinic of our hospital during the period 1998–2009. HB(O)C families fulfilled at least one of the following inclusion criteria: 3 breast and/or ovarian cancer cases in 2 generations of the same parental branch, 2 breast and/or ovarian cancer cases diagnosed at <50 years of age in the same parental branch, or 1 breast cancer case diagnosed at <35 years of age.

We recruited 450 healthy female blood donors visiting the blood bank of our hospital during the period 2002–2009. None of the blood donors reported any history of cancer in first-degree relatives.

Sporadic breast cancer cases \((n = 326)\) were identified by the Oncology Department of our hospital during the period 2002–2008. All female breast cancer cases classified as sporadic were eligible for the present study. We classified breast cancer cases as sporadic if no first-degree relatives with breast cancer had been reported.

All participants provided written informed consent. The institutional ethics committee approved the study.

\subsection*{BRCA1 Genotyping}
Genomic DNA was extracted from peripheral blood lymphocytes in a fully automated nucleic acid-purification workstation (MagNA Pure Compact System; Roche Diagnostics). MagNA Pure Compact DNA Isolation Kits were used according to the manufacturer's protocol. Index cases from HB(O)C families were scanned for the presence of \textit{BRCA1} germ-line mutations by denaturing gradient gel electrophoresis and sequencing, as previously described \((16)\). We used denaturing gradient gel electrophoresis to genotype c.591C>T \((p.197=)\), c.5123C>A \((p.A1708E)\), and 14 \textit{BRCA1} tagging single-nucleotide polymorphisms (SNPs) in sporadic breast cancer cases and in healthy controls. \textit{BRCA1} tagging SNPs have been described previously \((17)\). Human Genome Variation Society guidelines (http://www.hgvs.org) were used for \textit{BRCA1} mutation nomenclature.

\subsection*{In Silico Splicing Analysis}
The following splice-site analysis tools were used: \((a)\) Berkeley Drosophila Genome Project (BDGP) (http://www.fruitfly.org/seq_tools/splice.html); \((b)\) NetGene2 (NG2) (http://www.cbs.dtu.dk/services/NetGene2); and \((c)\) Human Splicing Finder (HSF) (http://www.umdb.be/HSF/).

Human default parameter settings were used in all analyses. BDGP, NG2, and HSF matrices were used to analyze the effect of the c.591C>T variant on donor site efficiency. In addition, the following HSF matrices were used to analyze the effect of the c.591C>T variant on the following putative splicing regulatory sequences: HSF integrated matrices for serine/arginine-rich proteins (SRp40, SC35, SF2/ASF, SF2/ASF IgM/BRCA1, and SRp55), exonic splicing enhancer (ESE) motifs from HSF (ESE-HSFs), RESCUE ESE hexamers (RESCUE-ESE), putative 8-mer ESEs (PESEs) and putative 8-mer exonic splicing silencers (PESSs), exon-identity and intron-identity elements (EIIEs and IIIEs), heterogeneous nuclear ribonucleoprotein–binding motifs, and Fas exonic splicing silencers. A comprehensive description of these splicing regulatory elements can be found elsewhere \((18)\).

The analysis was performed with a 400-bp sequence centered on the c.591C>T variant. In accordance with other studies \((19)\), a scoring decrease \(>10\%\) was considered predictive of splicing alterations.

\subsection*{cDNA Synthesis}
We analyzed 7 c.591C>T carriers \((3\) index cases, \(2\) relatives, and \(2\) healthy controls) and 10 noncarriers \((7\) healthy controls, \(2\) index cases carrying \textit{BRCA1} disease-causing mutation c.421-1G>A, and \(1\) index case carrying \textit{BRCA1} disease-causing mutation p.L1974X). Total RNA was extracted from peripheral blood lymphocytes in a MagNA Pure Compact workstation. MagNA Pure Compact RNA Isolation Kits were used according to the manufacturer’s protocol. First-strand cDNA was synthesized with random hexamers (Promega) and a mixture of reverse transcriptases from avian myeloblastosis virus and Moloney murine leukemia virus (Promega).

\subsection*{PCR Amplification of \textit{BRCA1} Splicing Isoforms}
To amplify different \textit{BRCA1} splicing isoforms, we designed primers with the following sequences (5' to 3'): CATCCAAAGTGATGGCTCAGA (exon7fw), TGTGCAACTCTCTAACCTTG (exon8fw), ACACTGCAATGGCTGCTTTGTG (exon8-11Afw), and TGTCATTTGCAATTGGTGCTTG (exon10fw), and TGTCATTTGCAATTGGTGCTTG (exon11Bfw). All primers were designed with the open source software, Primer3 \((20)\). Reverse primers were labeled with 6-carboxyfluorescein (FAM). PCRs were performed with FastStart Taq DNA polymerase (Roche).
Characterization of Splice Site Variants

Table 1. Splice-site prediction analysis of the BRCA1 variant c.591C>T. a,b

<table>
<thead>
<tr>
<th>Wild type</th>
<th>Score</th>
<th>Variant</th>
<th>Score</th>
<th>Variation, %</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAGgtgagt</td>
<td>98.84</td>
<td>TAGgtgagt</td>
<td>96.86</td>
<td>-2</td>
<td>HSF matrices</td>
</tr>
<tr>
<td>CAGgtgagt</td>
<td>10.67</td>
<td>TAGgtgagt</td>
<td>8.83</td>
<td>-17.2</td>
<td>MESb</td>
</tr>
<tr>
<td>CAGgtgagt</td>
<td>1.00</td>
<td>TAGgtgagt</td>
<td>0.99</td>
<td>-0.01</td>
<td>BDGP</td>
</tr>
<tr>
<td>CAGgtgagt</td>
<td>0.81</td>
<td>TAGgtgagt</td>
<td>0.67</td>
<td>-17.2</td>
<td>NG2</td>
</tr>
</tbody>
</table>

a The 3’ end of exon 9 is displayed in capital letters. The polymorphic position is displayed in boldface.

b MES, MaxEnt scoring.

Results

To date, we have performed BRCA1 mutation scanning in a cohort of 685 index cases of HB(O)C families from Spain. The BRCA1 c.591C>T variant was identified in 10 index cases (1.5%). In addition, we have detected the variant in 4 of 450 healthy controls (0.9%) and in 1 of 326 (0.3%) cases of sporadic breast cancer.

BRCA1 alleles are described almost completely by 10 canonical haplotypes defined by 14 tagging SNPs (17); however, haplotypes 1 and 2 are the commonest BRCA1 haplotypes (at least in Western populations), accounting for roughly 78% of all chromosomes (17). In our cohort, the c.591C>T variant was associated only with haplotype 1 (data not shown); however, most haplotype 1 alleles do not contain this variant. Interestingly, the disease-causing mutation c.5123C>A was detected only in index cases carrying the c.591C>T variant. Segregation analysis confirmed that c.5123C>A was linked to c.591C>T, at least in our population.

The c.591C>T variant is located at position −3 of the intron 9 donor splice site. We used 4 splice site-analysis tools to predict the effect on splicing, with conflicting results (Table 1). The donor splice site scoring was either decreased (MaxEnt scoring and NG2) or preserved (HSF matrices and BDGP). In addition, the c.591C>T variant was associated with potential splicing enhancers and the creation of splicing silencers (Table 2). RESCUE-ESE, PESE, EIE, and ESE-HSF matrices found no differences between variant and wild-type sequences (data not shown).

To investigate a possible effect of the c.591C>T variant on splicing, we analyzed cDNA samples derived from 7 carriers and 10 noncarriers. After cDNA amplification with primers exon8fw and exon10rev, we observed the expected 201-bp product corresponding to the full-length BRCA1 transcript in all tested samples. An additional product of 155 bp was observed only in c.591C>T carriers (Fig. 1A). Because BRCA1 exon 9 spans 46 nucleotides, the data were compatible with
Clinical Chemistry 56:1 (2010)

Table 2. Predicted effect of BRCA1 variant c.591C>T on splicing regulatory motifs. a

<table>
<thead>
<tr>
<th>Role</th>
<th>Motif</th>
<th>Wild type</th>
<th>Score</th>
<th>Variant</th>
<th>Score</th>
<th>Variation, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEb</td>
<td>SRp40</td>
<td>TTGCAgg</td>
<td>82.51</td>
<td>TTGTAgg</td>
<td>0</td>
<td>−100 (site broken)</td>
</tr>
<tr>
<td>SE</td>
<td>SRp55</td>
<td>TGCAgg</td>
<td>75.65</td>
<td>TGTA gg</td>
<td>0</td>
<td>−100 (site broken)</td>
</tr>
<tr>
<td>SE</td>
<td>SF2/ASF</td>
<td>TGCAgg</td>
<td>72.77</td>
<td>TGTA gg</td>
<td>0</td>
<td>−100 (site broken)</td>
</tr>
<tr>
<td>SS</td>
<td>Fas-ESS</td>
<td>Absent</td>
<td>NA</td>
<td>TGTAgg</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>Fas-ESS</td>
<td>Absent</td>
<td>NA</td>
<td>GTAgg</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>PESS</td>
<td>Absent</td>
<td>NA</td>
<td>TATTGTA</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>PESS</td>
<td>Absent</td>
<td>NA</td>
<td>TATTGTA</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>IIEs</td>
<td>Absent</td>
<td>NA</td>
<td>ATTTGA</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>IIEs</td>
<td>Absent</td>
<td>NA</td>
<td>TGTAgg</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>IIEs</td>
<td>Absent</td>
<td>NA</td>
<td>TGTAgg</td>
<td>NA</td>
<td>New site</td>
</tr>
<tr>
<td>SS</td>
<td>hnRNP</td>
<td>CAggtg</td>
<td>70.24</td>
<td>TAggtg</td>
<td>85.48</td>
<td>+21.69</td>
</tr>
</tbody>
</table>

a The 3’ end of exon 9 is displayed in capital letters. The polymorphic position is displayed in boldface.

b SE, splicing enhancer; SS, splicing silencer; NA, not applicable; FAS-ESS, Fas exonic splicing silencer; hnRNP, heterogeneous nuclear ribonucleoprotein. SRp40, SRp55, and SF2/ASF are serine/arginine-rich proteins.

exon 9 skipping. The BRCA1 Δ9 signal was consistently lower than the full-length signal in all carriers (Fig. 1A). The mean (SD) normalized signal ratio was 0.22 (0.05). Exon 9 skipping was confirmed by direct sequencing (Fig. 1B). To assess any contribution of the variant allele to the population of full-length transcripts, we sequenced PCR products obtained with a forward primer designed to anneal across the exon 8–exon 9 junction (exon8-9fw). The chromatogram showed that full-length transcripts arise predominantly, but not exclusively, from the wild-type allele (Fig. 1C). The data suggest that the c.591C>T variant decreases the efficiency of, but does not abolish, the intron 9 donor splice site.

BRCA1 splicing variants Δ(9,10) and Δ(9,10,11B) do not require a functional intron 9 donor site for proper splicing, at least in theory; however, we were interested in investigating any effect of the c.591C>T variant on these isoforms and so conducted additional experiments. After cDNA amplification with the primers exon7fw and exon11Brev, we observed double peaks of 446/449 bp and 323/326 bp that corresponded to full-length and Δ(9,10) isoforms, respectively, in all investigated samples (Fig. 2). In c.591C>T carriers, we also observed an extra double peak of 400/403 bp (compatible with exon 9 skipping) (Fig. 2). Sequencing (data not shown) confirmed that the double peaks corresponded to wobble splicing occurring at the 5’ end of BRCA1 exon 8 (15).

cDNA amplification with primers exon8fw and exon12rev generated 2 PCR products of 396 bp and 273 bp, corresponding to the Δ11B and Δ(9,10,11B) isoforms, respectively (Fig. 2). The products corresponding to the full-length and Δ(9,10) isoforms were too large to be amplified under the same conditions. An extra product of 350 bp was generated only in c.591C>T carriers, indicating that exon 9 skipping also occurs in the Δ11B isoform (Fig. 2).

Interestingly, the mean normalized signal ratios of the Δ(9,10) and Δ(9,10,11B) isoforms were higher in c.591C>T carriers. These findings were statistically significant (Fig. 2). The data suggested that c.591C>T modifies the relative expression level of Δ(9,10), Δ(9,10,11B), Δ11B, and full-length isoforms.

To demonstrate ASS associated with the variant allele, we selected 4 samples (2 c.591C>T carriers and 2 noncarrier healthy controls) heterozygous for BRCA1 haplotypes 1 and 2 and amplified cDNAs with 3 different forward primers (exon12fw; exon8–11Afw, a primer designed to anneal across the exon 8–exon 11A junction; and exon10fw) in combination with 1 reverse primer (exon13rev). Use of primers exon8–11Afw and exon10fw individually in combination with primer exon13rev amplified products specific for Δ(9,10,11B) and Δ(11B), respectively. All PCR products were sequenced with primer exon12fw. We used the BRCA1 tagging SNP rs1060915 located in exon 13 (c.4308T>C) to distinguish the expression of T (haplotype 1) and C (haplotype 2) alleles. Sequencing of the PCR fragment encompassing exons 12 and 13 indicated similar expression of the 2 alleles, irrespective of c.591C>T status (Fig. 3). With the exception of BRCA1-IRIS (21), both exon 12 and exon 13 are present in all BRCA1 isoforms that have been described.
Fig. 1. Detection of BRCA1 exon 9 skipping in c.591C>T carriers.
(A), Representative examples of cDNA amplification with BRCA1 primers exon8fw and exon10rev. (B), Exon 9 skipping demonstrated by direct sequencing. (C), Sequencing a PCR product generated with a forward primer annealing across the exon 8–exon 9 junction demonstrated that the variant allele produces full-length transcripts.
Therefore, the data indicated that the expression levels of wild-type and variant alleles were roughly equivalent. As expected, sequencing of Δ(9,10,11B) and Δ11B isoforms showed similar contributions from the 2 alleles in healthy control samples (Fig. 3). In c.591C>T carriers, sequencing of Δ(9,10,11B) products revealed a predominant contribution of the variant (haplotype 1) allele (Fig. 3). On the contrary, sequencing of Δ11B products revealed a predominant contribution of the wild-type (haplotype 2) allele (Fig. 3). Taken together, the data indicated that the population of messengers arising from the variant allele was characterized by a substantial reduction of full-length and Δ9 isoforms that was compensated by an increase in the population of Δ(9,10) and Δ(9,10,11B) isoforms, demonstrating ASS.

**Discussion**

In this study, we have shown that the BRCA1 c.591C>T variant is associated with ASS. Several groups have considered that this variant is most likely a neutral polymorphism (22–24). The variant is recorded 31 times in the BIC database (Cys197Cys according to the BIC traditional nomenclature) as a silent change of no clinical importance (2). Our study further supports a nonpathogenic role for this change. First, the variant is present in 0.9% of the healthy control population. Second, the expression levels of variant and wild-type alleles are similar. Lastly, c.5123C>A, a known BRCA1 disease-causing mutation (25), is observed in cis with c.591C>T in HB(O)C families.
To our knowledge, this report is the first of non-pathogenic ASS occurring in the \textit{BRCA1} gene. The ASS is characterized by a reduced population of full-length and $\Delta 11B$ transcripts, which is counterbalanced by an increased population of $\Delta (9,10)$ and $\Delta (9,10,11B)$ transcripts such that the overall expression level is similar to that of the wild-type allele. In addition, ASS is characterized by the presence of transcripts lacking exon 9. A schematic representation of c.591C$\rightarrow$T–associated ASS is shown in Fig. 4.

RNA-splicing analysis is essential for the assessment of many \textit{BRCA1} UVs (8, 12, 19, 26, 27). In most
cases, analysis is performed by amplifying cDNA with primers flanking the exon likely to be affected by the variant. This approach is designed to detect single-cassette exon events. Interestingly, it will produce misleading results if it is used to test the effect of the BRCA1 c.591C>T variant on splicing. This is the case because cDNA amplification with primers located in exons 8–10 (flanking exons) restricts the analysis to the full-length isoform, thus detecting exon 9 skipping (a frameshift alteration) but missing an increase in the skipping of exons 9–10 (a naturally occurring BRCA1 isoform preserving the open reading frame). The latter probably explains why the BRCA1 c.591C>T variant does not have a pathogenic effect (to be more precise, it does not cause an autosomal dominant HB(O)C syndrome), although a role as an allele with a low-penetrance risk cannot be fully discarded. Moreover, it is important to point out that ASS has been characterized in blood cells and does not necessarily reflect the situation in other tissues, such as the mammary and ovary epithelia.

Our study was not intended to investigate the molecular mechanisms responsible for the association between c.591C>T and ASS, so we can only speculate. In our opinion, the mere localization of the variant points to a reduced efficiency of the intron 9 splicing donor site as a likely causative mechanism, but other mechanisms are possible. For instance, the variant may target some splicing regulatory sequences. Both hypotheses find some support from in silico analysis (Tables 1 and 2). Regardless of the true mechanisms involved, we propose that inclusion/skipping of exons 9–10 are competing splicing reactions, so that hampering one of the reactions shifts the process to the alternative reaction. In this scenario, a sequence change slightly disrupting any splicing site (or splicing regulatory) sequence is more likely to affect the outcome of the splicing reaction.

At least 33 different BRCA1 variants recorded in the BIC database (2) are located in the genomic region encompassing exons 9 and 10. Of these variants, only 4 are considered disease-causing mutations: 3 exonic variants that have been reported only once (c.625_626ins20, c.668_669insA, c.668_669delA) and 1 intronic variant reported 11 times (c.594-2A>C). The pathogenic nature of the latter is supported by experimental data demonstrating exon 10 skipping (12); however, we suggest that the relevant outcome of intron 9 acceptor site deficiency could be (as is the case for intron 9 donor site deficiency) an increased population of Δ(9,10) transcripts, a possibility that has not yet been investigated.

Our study highlights the fact that splice variants other than the full-length isoform should be considered when analyzing the effect of certain BRCA1 variants (or variants in any other tumor suppressor gene) on splicing. This consideration is important because restricting the in vitro analysis to single-cassette events (namely, select primers for splicing analysis in flanking exons) may introduce bias that affects the interpretation of biological relevance. In addition, our study suggests that the false-negative rate of splicing-prediction programs may increase if alternative splicing variants (naturally occurring alternative splicing pathways) do exist.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures of Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: Research grant FIS 05/0864 ISCIII, Spanish Ministry of Science and Innovation; research grant RD 06/0020/0021 from RTICC, ISCIII, Spanish Ministry of Science and Innovation; research grants FMMA-06 and FMMA-08, Fundación Mutua Madrilena.
Expert Testimony: None declared.
Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: We thank the patients and their families for their cooperation.

References