Aberrant splicing induced by missense mutations in \textit{BRCA1}: clues from a humanized mouse model

Yongping Yang, Srividya Swaminathan, Betty K. Martin and Shyam K. Sharan*

Mouse Cancer Genetics Program, Center for Cancer Research, National Cancer Institute at Frederick, 1050 Boyles Street, Frederick, MD 21702, USA

Received April 15, 2003; Revised June 19, 2003; Accepted June 30, 2003

Numerous missense mutations in human \textit{BRCA1} gene have been linked to predisposition to breast cancer. However, the functional significance of the majority of these mutations remains unknown. We have examined the molecular basis for three such cancer-causing mutations. The first mutation, a T→G transversion in codon 64, is predicted to change a conserved cysteine residue to glycine in the RING finger domain of the 1863 amino acid BRCA1 protein. Using a humanized mouse model we demonstrate that this missense mutation actually results in a functionally null protein. This striking result occurs because the single base alteration generates a new 5’ splice site in exon 5 and also disrupts a putative exonic splicing enhancer motif. Consequently, the normal splice donor site is disrupted and an internal cryptic splice site is activated. This results in a 22-nucleotide deletion and the aberrant transcript is predicted to encode a severely truncated protein consisting of only 63 amino acids. To identify other missense mutations in \textit{BRCA1} that may result in aberrant splicing, we screened various mutations using the Genscan program. We demonstrate that at least two other missense mutations in codons 1495 and 1823 result in aberrant splicing due to the possible disruption of \textit{cis}-acting splicing regulatory elements. In conclusion, our study demonstrates for the first time the application of a humanized mouse model for functional analysis of human mutations in mice and also shows the need for a careful examination of the functional consequences of single base alterations and single nucleotide polymorphisms identified in human disease-causing genes.

INTRODUCTION

Germline mutations in human breast cancer susceptibility gene \textit{BRCA1} are responsible for approximately half of all familial hereditary breast cancer cases and also confer increased risk of ovarian, colon or prostate cancer (reviewed in 1). \textit{BRCA1} encodes a 1863 amino acid protein with an N-terminal RING finger domain and two BRCT-domains, also present in other proteins involved in DNA repair and cell-cycle regulation (2–4). The interaction of BRCA1 with RAD51, a key DNA repair protein, has revealed its role in recombinational repair of double strand breaks (5). BRCA1 is part of the RAD50-MRE11-NBS1 complex that is involved in non-homologous end-joining and homologous recombination (6). A role for BRCA1 in transcription-coupled repair of DNA that is damaged by oxidative agents has also been described (7). In addition, BRCA1 has been shown to be a target for phosphorylation by the ataxia telangiectasia mutated (ATM) and checkpoint kinase 2 (CHK2) proteins in response to DNA damage (8,9). BRCA1 also regulates the G2/M checkpoint in response to damaged DNA by activating CHK1 kinase (10). Taken together, BRCA1 appears to play a critical role in maintenance of genomic integrity and loss of its function results in tumorigenesis.

A large number of mutations have been identified in \textit{BRCA1} and individuals who inherit these mutations have a 60–80% risk of developing breast cancer by the age of 70 (11,12). These mutations are present throughout the length of the gene and are listed in the Breast Cancer Information Core database (http://research.nhgri.nih.gov/bic/). Most common mutations are either small insertions or deletions that alter the reading frame and are predicted to result in truncated BRCA1 protein. Various nonsense mutations and splice-site mutations with similar effect on the protein have also been identified. In addition to these clearly disruptive alterations, several missense mutations have also been identified along the length of the protein and account for about 25% of all the mutations reported (13). In the absence of a functional assay, the biological significance of these missense mutations remains largely unknown. It is puzzling that a single amino acid alteration in this relatively
large protein has the same effect as the complete absence of the protein. To explain how these missense mutations may result in tumorigenesis, we hypothesize that the protein has various functional domains that are involved in different biological processes and a mutation in any of these may result in tumorigenesis. The possibility that some of these single amino acid alterations may result in an unstable protein cannot be ruled out. However, this cannot be true for all missense mutations, as some of the variants have been shown to result in a stable protein (Yang and Sharan, unpublished data).

Owing to the relatively low (59%) sequence identity between human and mouse BRCA1 amino acid sequences, generating human missense mutations in mice may not be relevant for functional analysis (14). It is possible that missense mutations identified in humans, if generated in mouse Brcal, may have a different biological effect. This may be due to subtle differences in the protein folding/structure of the mouse and human BRCA1 proteins or due to suppression or enhancement of the phenotype by the non-conserved amino acids. We have previously described a humanized mouse model system that is ideal for in vivo functional analysis of human missense mutations in mice (15). In this model system, a human BRCA1 transgene functionally replaces the endogenous Brcal gene. We have shown that the wild-type human gene present in a bacterial artificial chromosome (BAC), when introduced into the germline of Brca1 knockout mice, can rescue the embryonic lethality of the homozygous mutant mice. In addition, the expression of the human gene mirrors the expression of the endogenous mouse gene (15). Here we describe the use of this model system to study the functional significance of a missense mutation in codon 64 that was predicted to change a single amino acid in the functionally important RING finger domain of BRCA1 (16). Our results show that the single nucleotide change disrupts normal splicing and results in an aberrant transcript that generates a functionally null protein. Furthermore, we have predicted and demonstrated similar effects on the splicing process by two additional missense mutations in codons 1495 and 1823 of BRCA1.

RESULTS

Generation of BRCA1 humanized mice with a 309T→G alteration in codon 64

The missense mutation at position 309 of the human BRCA1 cDNA (U14680) involved transversion of a T to G nucleotide (represented as 309T→G). This mutation in codon 64 (TGT→GGT) is predicted to change the seventh conserved cysteine residue to glycine in the Cys64-His-Cys64 RING finger domain (16). We generated this mutation in human BRCA1 gene present in a BAC clone, BAC RP11-81205, using an oligonucleotide-based recombineering method (17). BAC RP11-81205 has been shown to contain a 200 kb insert including 80 kb of BRCA1, 90 kb of upstream sequence and 30 kb of downstream sequence and can completely rescue the embryonic lethality of Brca11700T mutant mice (15). Brca11700T mutant allele was generated by inserting the neomycin resistance gene into exon 20 to disrupt the Brca1 gene (18).

Mice heterozygous for the Brca11700T mutation are viable and fertile but homozygous mutant animals die during early embryogenesis.

The modified BAC containing the 309T→G mutation was used to generate transgenic mice that were also heterozygous for the Brca11700T allele. Three independent Brca1 heterozygous transgenic founders (Brca1 Ko/+; Tg) were then crossed to Brca1 Ko/+ mice to obtain homozygous mutant Brca1 transgenic mice (Brca1 Ko/Ko; Tg/+ ) to examine the effect of the mutation in codon 64 of human BRCA1 in mice lacking endogenous Brca1 gene. We confirmed the expression of the BAC transgene in all the three lines by RT–PCR (Fig. 1A). Northern analysis revealed expression of the full-length human BRCA1 transcript in multiple tissues such as the kidney, ovary, spleen, testis and thymus of 6-week-old transgenic mice (Fig. 1B).

Analysis of 309T→G mutation in codon 64 in Brca1 Ko/Ko; Tg/+ mice

We failed to obtain any Brca1 Ko/Ko; Tg/+ mice (28 mice were expected out of 194 F2 offspring) when the Brca1 Ko/+; Tg/+ mice were crossed with Brca1 Ko/+ mice, suggesting that the 309T→G mutation in codon 64 resulted in embryonic lethality. To determine the phenotype associated with the Cys64Gly missense mutation, we first examined the embryos obtained from Brca1 Ko/+; Tg/+ mice that were mated to Brca1 Ko/+ mice at day 8.5 of gestation when the Brca1 Ko/Ko embryos first show a mutant phenotype. We obtained two phenotypic classes of embryos; those that appeared to undergo normal embryonic development and a few that appeared to be extremely reduced in size and resembled the homozygous mutant embryos (Fig. 1C). Genotypic analysis of the embryos revealed that the mutant embryos were either Brca1 Ko/Ko or Brca1 Ko/Ko; Tg/+ (Fig. 1D). These embryos were indistinguishable from each other.

These findings suggested that the RING finger domain of BRCA1 was required for normal embryonic development. Alternatively, the cysteine to glycine change affected the stability of the mutant protein. We examined the expression of the mutant protein in transgenic mice by western analysis using antibodies specific to human BRCA1. A 220 kDa wild-type BRCA1 protein was detected in tissues from mice carrying a wild-type BRCA1 transgene; however, no BRCA1 protein was detected in tissues from mice carrying the mutant BRCA1 transgene (Fig. 1E).

Unstable protein or aberrant transcript?

We examined the possibility that the mutant BRCA1 protein was not detected because the full-length protein was not synthesized due to the presence of additional mutations in the transcript. We amplified a region of BRCA1 cDNA that spanned exons 2–11 using RT–PCR for sequence analysis. Surprisingly, while the wild-type human BRCA1 amplified a 456 bp fragment, the mutant transcripts amplified a smaller PCR product (Fig. 2A). Sequence analysis of the two RT–PCR products revealed a 22-nucleotide deletion in the transcript generated from the mutant allele. These 22 nucleotides were from the 3′ end of exon 5 immediately downstream of codon 64.
Since the genomic sequence of the mutant allele showed the presence of 22 nucleotides, we speculated that these nucleotides were lost during splicing (Fig. 2D). The 22-nucleotide deletion results in a frame shift alteration after codon 63 and is predicted to generate a premature termination codon at position 80 (UGA) within exon 6. These results show that the missense mutation in codon 64 results in a functionally null protein due to aberrant splicing.

Effect of the 309T→G mutation in humans

We next asked whether a similar splicing defect occurred in human patients carrying the 309T→G mutation. We performed RT–PCR using total RNA obtained from Epstein–Barr virus-transformed lymphocytic cell lines derived from a family (Kindred 27) carrying the 309T→G mutation in BRCA1 (16). RT–PCR product obtained from cells without the mutation.
showed the expected 456 bp fragment. Cells derived from patients carrying the 309T→G mutation showed the presence of two RT–PCR products, a normal 456 bp fragment and a smaller 434 bp fragment, both of equal intensity (Fig. 3A). The presence of normal and aberrant cDNA fragments in the mutation carrier was consistent with the presence of both wild-type and mutant alleles in the lymphocytic cells. Sequence analysis of the two bands confirmed that the 434 bp fragment contained the 22-nucleotide deletion found in transgenic mice while the 456 bp fragment contained full-length exon 5 sequence (Fig. 3B). These findings clearly demonstrated that the 309T→G mutation in exon 5 results in a splicing defect in humans.

Other missense mutations in BRCA1 that affect splicing

To examine whether there were other missense mutations in the human BRCA1 gene that may also result in aberrant splicing, we examined various missense mutations deposited in the Breast Cancer Information core database using the Genscan program. This program predicts putative open reading frames utilizing various parameters, including the presence of splice sites (19). Based on the results of the Genscan analysis, we identified two missense mutations that could potentially disrupt a 5' splicing site. The first mutation is in codon 1495, a G to T change in nucleotide 4603 (4603G→T), the last nucleotide of exon 14. The 4603G→T mutation is predicted to change an arginine (AGG) residue to a methionine (ATG). The second missense mutation is predicted to change an alanine (GCA) to a threonine (ACA) in codon 1823, due to a G to A alteration in nucleotide 5586 (5586G→A), the last base of exon 23. According to the Genscan program, both missense mutations are predicted to result in skipping of their respective exons during splicing. Further support for the Genscan prediction was provided by a recent report showing that in a breast and ovarian cancer patient the 5586G→A mutation in codon 1823 results in the deletion of exon 23 (20).

To test the prediction that the exon 14 missense mutation results in exon skipping and to confirm the aberrant splicing due to the exon 23 missense mutation, we generated the two mutations in BAC RP11-812O5 in which a neomycin resistance gene (Neo) was introduced into the vector sequence (Figs 4A and 5A). The BACs containing the missense mutation were introduced into mouse embryonic stem (ES) cells. We then examined the BRCA1 transcripts generated from the two mutant alleles by RT–PCR. We determined the effect of the mutation in exon 14 using PCR primers to amplify a 347 bp fragment of the transcript spanning exons 13–15 (Fig. 4B).
RT–PCR product obtained from an ES cell clone expressing wild-type BRCA1 showed the expected size fragment. However, the RT–PCR product obtained from the mutant allele showed the presence of a single fragment smaller in size (220 bp) than the wild-type fragment (Fig. 4B). Sequence analysis revealed that the mutant transcript lacked 127 nucleotides encoded by exon 14, confirming the Genscan prediction (Fig. 4C). Deletion of exon 14 leads to a frame shift that causes premature truncation of the protein at position 1462 within exon 15.

Finally, we examined the effect of the missense mutation in codon 1823 in exon 23 by RT–PCR. A 279 bp RT–PCR product containing exons 20–24 was obtained from the wild-type BRCA1 gene (Fig. 5B). However, ES cell clones containing the mutant BRCA1 showed the presence of two fragments. One of the fragments was similar to the RT–PCR product obtained from wild-type BRCA1, in addition, there was a smaller, 212 bp fragment (Fig. 5B). These results were very similar to those observed in a human patient with this missense mutation where two different transcripts were identified, a full-length BRCA1 transcript generated by the wild-type allele and a smaller transcript lacking exon 23 from the allele with the 5586G→A mutation (20). We confirmed the deletion of exon 23 by sequencing the 212 bp RT–PCR fragment (Fig. 5C). This variant of the BRCA1 transcript is expected to generate a truncated protein due to a new termination codon (UGA) at the position 1813 within exon 24.

The presence of the longer RT–PCR product in ES cells that contain only the mutant human BRCA1 allele was puzzling (Fig. 5B). Sequence analysis confirmed that the upper fragment represented two different but equally represented BRCA1 transcripts (Fig. 5C). One of the variants showed the presence of full-length exon 23 including the 5586G→A mutated nucleotide. This transcript is predicted to result in a full-length BRCA1 protein with an alanine to threonine change in codon 1823. In addition, a novel splice variant of BRCA1 was also identified. This variant showed the presence of full-length exon 23 but also included the first five nucleotides of intron 23 (Fig. 5C). The introduction of five additional nucleotides in the transcripts is expected to cause a frame shift creating a premature truncation of the protein at the position 1836 within exon 24.

Based on densitometric analysis of the RT–PCR products on an agarose gel, these two transcripts constitute a small percentage (~20%) of the total BRCA1 transcripts expressed (Fig. 5B). This suggests that the exon 23-skipped variant is the predominant expressed transcript. However, the preferential amplification of the smaller RT–PCR product cannot be ruled out.

**DISCUSSION**

We have used a humanized mouse model system to examine the functional significance of a missense mutation in BRCA1. This mutation, caused by a T→G transversion in codon 64, was predicted to change a cysteine residue to glycine (16). However, we have demonstrated that the missense mutation prevents translation of the full-length BRCA1 protein due to aberrant splicing of the transcript. We have confirmed that this mutation results in a splicing defect in humans too. Using mouse ES cells expressing human BRCA1 present in a BAC, we have demonstrated that two other missense mutations in codons 1495 and 1823 also result in aberrant splicing due to possible disruption of cis-acting splicing regulatory elements.
Although the missense mutation in codon 1823 is listed as a putative missense mutation in the BIC database, a recent report has shown that in a breast and ovarian cancer patient this mutation results in skipping of exon 23 due to aberrant splicing (20). Our analysis of the codon 1823 mutation in the human gene expressed in mouse ES cells is in accordance with this finding. In addition, we have identified two minor splicing variants that were not discovered in the human patient expressing the wild-type and the mutant alleles. Since the size of the RT–PCR products generated by the variants containing exon 23 is similar to those generated by the wild-type transcript, the apparently normal sized transcripts were not analyzed. However, in the mouse ES cells expressing only the mutant \textit{BRCA1} transcript, the presence of RT–PCR products similar to those from wild-type allele clearly suggested the presence of additional splice variants. We have confirmed the presence of the splice variant containing the first five nucleotides of intron 23 in the human patient sample by sequencing individually cloned RT–PCR products (data not shown). Our results clearly demonstrate the relevance of using mouse ES cells in the analysis of splicing variants.

Mutations that affect splicing of the pre-mRNA are estimated to account for $\sim$15% of human genetic disorders (21). However, this estimate is based on mutations that map to the classical splice-site sequences. The presence of cis-acting elements like the exonic splicing enhancer (ESE) and the exonic splicing silencer (ESS) recognized by the cellular splicing machinery play a critical role in splicing of various exons present in the pre-mRNA (reviewed in 22). The ESE motifs are recognized by a family of serine/arginine-rich splicing factors (known as the SR proteins) such as SF2/ASF, SC35, SRp40 and SRp55 proteins (23,24). More recently, mutations in these regulatory sequences have been shown to disrupt normal splicing (reviewed in 25,26). One such mutation has been reported in the \textit{BRCA1} gene where a nonsense mutation in codon 1694 in exon 18 causes skipping of the entire exon 18 due to disruption of an SF2/ASF enhancer motif (27,28). An evaluation of 50 point mutations in various human genes that result in exon skipping showed that more than half of them affected the SR protein binding motifs (29). Further support for this prediction is provided by the fact that 48–50% of mutations in neurofibromatosis type 1 (\textit{NF1}) and \textit{ATM} result in defective splicing (30,31). Recently, a missense mutation in codon 2722 of \textit{BRCA2}, predicted to disrupt an ESE motif, was shown to result in exon skipping (32).

We have evaluated the effect of the three missense mutations on putative SR protein binding sites by utilizing the ESE motif-scoring matrices (33–35). ESE motif-scoring matrices predicted the presence of multiple high-score motifs of SF2/ASF, SC35, SRp40 and SRp55 within exons 5, 14 and 23. The 309T$\rightarrow$G mutation in codon 1495 in exon 14 causes skipping of the entire exon 18 due to disruption of an SF2/ASF enhancer motif (27,28). An evaluation of 50 point mutations in various human genes that result in exon skipping showed that more than half of them affected the SR protein binding motifs (29). Further support for this prediction is provided by the fact that 48–50% of mutations in neurofibromatosis type 1 (\textit{NF1}) and \textit{ATM} result in defective splicing (30,31). Recently, a missense mutation in codon 2722 of \textit{BRCA2}, predicted to disrupt an ESE motif, was shown to result in exon skipping (32).

We have evaluated the effect of the three missense mutation on putative SR protein binding sites by utilizing the ESE motif-scoring matrices (33–35). ESE motif-scoring matrices predicted the presence of multiple high-score motifs of SF2/ASF, SC35, SRp40 and SRp55 within exons 5, 14 and 23. The 309T$\rightarrow$G mutation in codon 1495 in exon 14 causes skipping of the entire exon 18 due to disruption of an SF2/ASF enhancer motif (27,28). An evaluation of 50 point mutations in various human genes that result in exon skipping showed that more than half of them affected the SR protein binding motifs (29). Further support for this prediction is provided by the fact that 48–50% of mutations in neurofibromatosis type 1 (\textit{NF1}) and \textit{ATM} result in defective splicing (30,31). Recently, a missense mutation in codon 2722 of \textit{BRCA2}, predicted to disrupt an ESE motif, was shown to result in exon skipping (32).

We have evaluated the effect of the three missense mutation on putative SR protein binding sites by utilizing the ESE motif-scoring matrices (33–35). ESE motif-scoring matrices predicted the presence of multiple high-score motifs of SF2/ASF, SC35, SRp40 and SRp55 within exons 5, 14 and 23. The 309T$\rightarrow$G mutation in codon 1495 in exon 14 causes skipping of the entire exon 18 due to disruption of an SF2/ASF enhancer motif (27,28). An evaluation of 50 point mutations in various human genes that result in exon skipping showed that more than half of them affected the SR protein binding motifs (29). Further support for this prediction is provided by the fact that 48–50% of mutations in neurofibromatosis type 1 (\textit{NF1}) and \textit{ATM} result in defective splicing (30,31). Recently, a missense mutation in codon 2722 of \textit{BRCA2}, predicted to disrupt an ESE motif, was shown to result in exon skipping (32).

We have evaluated the effect of the three missense mutation on putative SR protein binding sites by utilizing the ESE motif-scoring matrices (33–35). ESE motif-scoring matrices predicted the presence of multiple high-score motifs of SF2/ASF, SC35, SRp40 and SRp55 within exons 5, 14 and 23. The 309T$\rightarrow$G mutation in codon 1495 in exon 14 causes skipping of the entire exon 18 due to disruption of an SF2/ASF enhancer motif (27,28). An evaluation of 50 point mutations in various human genes that result in exon skipping showed that more than half of them affected the SR protein binding motifs (29). Further support for this prediction is provided by the fact that 48–50% of mutations in neurofibromatosis type 1 (\textit{NF1}) and \textit{ATM} result in defective splicing (30,31). Recently, a missense mutation in codon 2722 of \textit{BRCA2}, predicted to disrupt an ESE motif, was shown to result in exon skipping (32).

We have evaluated the effect of the three missense mutation on putative SR protein binding sites by utilizing the ESE motif-scoring matrices (33–35). ESE motif-scoring matrices predicted the presence of multiple high-score motifs of SF2/ASF, SC35, SRp40 and SRp55 within exons 5, 14 and 23. The 309T$\rightarrow$G mutation in codon 1495 in exon 14 causes skipping of the entire exon 18 due to disruption of an SF2/ASF enhancer motif (27,28). An evaluation of 50 point mutations in various human genes that result in exon skipping showed that more than half of them affected the SR protein binding motifs (29). Further support for this prediction is provided by the fact that 48–50% of mutations in neurofibromatosis type 1 (\textit{NF1}) and \textit{ATM} result in defective splicing (30,31). Recently, a missense mutation in codon 2722 of \textit{BRCA2}, predicted to disrupt an ESE motif, was shown to result in exon skipping (32).
complementary to the U1 snRNA (36). U1 snRNA is the RNA component of the U1 small nuclear ribonucleoprotein (U1 snRNP), which has an essential role in determining the 5' splicing site and also in the assembly of the splicing complex. The 309T→G transversion creates a new cryptic 5' splicing site motif. The sequence near codon 64 is complementary to nine out of 10 bases of U1 snRNA splice processing sequence at positions 1–10 of U1 snRNA (Fig. 6C). On the other hand the normal 5' splicing site of exon 5 shows complementarity to only seven of the first 10 nucleotides of the U1 snRNA (Fig. 6C). This may explain the dominance of the new splice site within exon 5.

The 4603G→T mutation in exon 14 disrupts a single U1 snRNA binding sequence (Fig. 6C). Although eight out of 10 nucleotides of the mutant BRCA1 transcript are complementary to the U1 snRNA sequence, the 4603G→T transversion disrupts the complementarity of the ninth nucleotide of the U1 snRNA within the transcript. Proper interaction between nucleotides 5–9 of the U1 snRNA with their complementary bases in the transcript is required for their interaction (37). Loss of this binding may account for the skipping of exon 14 during splicing. Similarly, mutation in codon 1823, located in the last base of exon 23, may explain generation of transcripts that skip this exon. Based on our findings, although splicing at the 3' end of exon 23 is markedly reduced, some transcripts do show normal 5' splicing of exon 23, in spite of the G→A mutation in the U1 snRNA binding site (Fig. 6C). We predict that the purine-to-purine transition mutation may not completely abolish U1 snRNA binding due to wobble pairing. This is in contrast to purine-to-pyrimidine transversion mutation in codon 1495, which completely abrogates the use of constitutive 5' splice site of exon 14. However, these predictions have to be tested to confirm the precise mechanism underlying the splicing defects due to single nucleotide changes in exons 5, 14 and 23.

The application of the humanized mouse model described here clearly demonstrates the usefulness of this model system to study the functional significance of point mutations identified in human disease causing genes. In spite of the amino acid sequence divergence, the mouse and human BRCA1 genes show strong functional conservation and can genetically complement each other. With the identification of a large number of mutations in BRCA1 and other human disease genes, the humanized model system should prove to be helpful in understanding their (those that result in stable protein) in vivo functional significance. In addition, using the ES cell based in vitro approach described here, the molecular and biochemical effects of the mutation can be assessed easily. Such an approach can be very helpful when it is not feasible to obtain samples from human patients for such studies. Our findings also warrant a careful investigation of various single base alterations and single-nucleotide polymorphisms (SNPs) that have been identified in various human disease-causing genes.

**MATERIALS AND METHODS**

**Generation of 309T→G mutation in BRCA1 in BAC by recombineering**

309T→G mutation in codon 64 of BRCA1 was generated in the BAC RP11-812O5 using the oligonucleotide-based BAC
recombineering technique (17). This method is based on the generation of BAC recombinants by the bacteriophage lambda Red recombination system present in DY380 cells when single-stranded oligonucleotides are used as targeting vectors (38,39). Recombineering was performed as described previously (17,40).

Figure 6. Effect of missense mutations in exon 5 and 23 on high score SR motifs and the effect of the three missense mutations on the binding of the U1 snRNA to the exonic 5′ splice site. The SR motif analysis was performed using ESEfinder (http://exon.cshl.org/ESE/index.html). High-score motifs are shown in red for SF2/ASF, blue for SC35, green for SRp40 and yellow for SRp55. The width of each bar along the x-axis represents the position of the motif in the DNA sequence. The height of each bar on the y-axis represents the numerical score. (A) Comparison of SR motifs in wild-type (WT) and 309T→G mutant exon 5 sequences. T to G change in nucleotide 56 of exon 5 (shown by an arrow) results in disruption of an SRp55 motif. (B) Comparison of SR motifs in wild-type (WT) and 5586G→A mutant exon 23 sequences. G to A change in the last nucleotide of exon 23 (shown by an arrow) results in generation of a new high score SRp55 motif. The SRp55 may result in activation of a cryptic splicing site in intron 23 that results in inclusion of five intronic nucleotides (underlined) in that transcript. (C) The pairing of various 5′ splice sites (5′ splice site of exon 5, cryptic 5′ splice site created by 309T→G missense mutation in exon 5, 5′ splice site of exon 14 containing 4603G→T mutation in the last nucleotide and 5′ splice site of exon 23 with 5586G→A mutation in the last nucleotide) with the U1 snRNA sequence is shown. A vertical line shows complementary base pairing and ‘+’ shows non-complementary pairing. The exonic sequences in BRCA1 transcript are shown in upper case and the intronic sequences are in lower case. The 5′ splice site consensus sequence is shown at the top. ‘/’ represents the splicing site.

Briefly, 10 ml of DY380 cells containing BAC RP11-812O5, grown at 32°C to an OD600 of 0.6, were induced at 42°C for 15 min. After being chilled on ice for 15 min, the cells were washed with ice-cold sterile water three times and resuspended in 50 μl of ice-cold sterile water. A 300 ng aliquot of single-stranded targeting vector was
immediately electroporated into these cells. The targeting vector, a 100mer synthetic oligonucleotide containing human BRCA1 sequence (5'-GCGATCTGACGCTACAGGGAGATGCTAGGT-3') with a single base change (T→G) in the middle (in bold) was used to generate 309T→G mutation in codon 64 in exon 5.

After electroporation cells were resuspended in 1 ml of SOC medium and incubated at 32°C for 1.5 h. The cells (~10^8) were diluted and plated in 2.2 ml-deep well plates (Marsh Biomedical Products Inc.) at about 10 cells per well in 500 μl LB media (containing chloramphenicol to maintain the BAC). The cells were grown at 32°C for about 24 h. After 24 h, 10 μl of culture from each pool were analyzed by PCR. To detect the BAC clones containing the point mutation, we utilized mismatch primers designed to amplify only the recombinants as described previously (17). A detection primer containing a two-base mismatch to the wild-type sequence but only a penultimate base mismatch to the mutated sequence at its 3' end, 5'-ACCTTTGGTTATATCATCTTGC-3', was utilized along with a forward primer, 5'-ATGCTATTACCACACTTTC-3', in a two-step PCR. The PCR condition included denaturation for 4 min at 94°C followed by 40 cycles of 94°C for 15 s and 60°C for 1 min (a common annealing and extension temperature), and a final extension at 72°C for 7 min. These primers amplified a 505 bp product. Once the positive pools were identified, cells were plated at low dilutions and isolated clones were screened as described above. Finally, the presence of missense mutation was confirmed by sequencing.

**Generation of 4603G→T and 5586G→A mutations in BRCA1**

These two mutations were generated in a BRCA1 containing BAC by a new two-step 'hit and fix' method, which is a modification of the method described previously (41). The two-step method has been designed to overcome the occasional non-specific amplification by the mismatch primer. In the first step of this method, a stretch of six to 20 nucleotides was randomly changed around the base where the mutation was being generated. In the second step, the modified bases generated in the first step were changed back to the original sequence except for the desired mutation. Since several nucleotides were changed in each of the two steps, the recombinant BACs were identified by standard PCR methods using a primer specific for the altered bases. Another modification involved the use of 180mer single-stranded targeting vectors instead of 100mer. These targeting vectors were generated as described previously by using two 100mer oligonucleotides with 20 complementary bases at the 3' end in a PCR reaction (17). Sequence of the targeting vectors and primers used for screening will be provided upon request. Generation of correct mutations in the BAC clones was confirmed by sequencing using BigDye termination sequencing reagents (Perkin Elmer) and ABI sequencer following manufacturer's instructions.

**Generation of humanized mice with 309T→G mutation**

BAC DNA was extracted using standard alkaline lysis method and then subjected to Cesium Chloride gradient on an ultracentrifuge. Supercoiled DNA was dialyzed overnight in 1x TE (10 mM Tris pH 7.6, 1 mM EDTA) buffer and diluted to 0.5–1.0 ng/μl for pronuclear microinjection. Fertilized eggs for microinjection were obtained from wild-type females that were superovulated and mated with Brca1 Ko/+ males.

Transgenic founders were identified by Southern analysis of BamHI-digested tail DNA using the chloramphenicol resistance gene present in the BAC vector, as probe. Mice were genotyped for the mutation in the endogenous Brca1 gene as described previously (15). Presence of Brca1 was assessed by PCR analysis of genomic DNA from transgenic mice using primers from the 5' untranslated region, portion of exon 11 present in the middle of Brca1 and 3' untranslated region. To obtain Brca1 Ko/Ko; Tg/+ mice, Brca1 Ko/+; Tg/+ founders or F1 animals were crossed with Brca1 Ko/+ mice.

**RT–PCR analysis**

Total RNA was isolated from various tissues of 6–8-week-old transgenic or non-transgenic mice using RNA-Beool following manufacturer's directions (Tel-Test Inc.). Twenty micrograms of each RNA sample were loaded on an agarose gel and transferred to Hybond N+ membrane (Amersham Pharmacia Biotech) following standard protocol (42). The northern blot was hybridized with a human BRCA1-specific single strand radio-labeled anti-sense probe generated by PCR using a reverse primer (5'-ACTGGAGCCACTTTGTTAG-3') and BRCA1 cDNA fragment 2743-3291 as a template in the presence of γ-32P-dCTP.
sequenced using BigDye termination sequencing reagents (Perkin Elmer). When individual RT–PCR products had to be sequenced, they were first sub-cloned using pGEM-T vector System II kit (Promega). Plasmid DNA from 15–25 individual clones containing the RT–PCR products were purified and the inserts were sequenced using T7 sequencing primer.

**Immuno-precipitation and western blot analysis**

Fresh mouse testes were homogenized in the lysis buffer (RIPA, 1% Triton X100, 0.2% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 0.1 mg/ml PMSF) containing protease inhibitor complex tablet (Roche Biotech). Homogenates were spun twice at 4°C, 12,000g for 15 min. The lysates were incubated with human-specific anti-BRCA1 Ab-3 monoclonal antibody (Oncogene) at 2 µg of antibody in 1 ml of tissue lysate at 4°C for 3 h, then Protein A-agarose was added and further incubated overnight. Immune-bound complexes were washed three times with lysis buffer, and separated by 6% SDS–PAGE and blotted to a nitrocellulose membrane.

Immunoblot analysis to detect the human BRCA1 protein was performed by incubating 4 µg of rabbit anti-BRCA1 polyclonal antibody (Ab-3, NeoMarkers) as primary antibody in 1 ml of Tris buffered saline (TBS) containing 5% non-fat dry milk at 4°C overnight. The bound antibodies were detected with incubation with enzyme conjugated anti-rabbit IgG : HRP and in 1 ml of Tris buffered saline (TBS) containing 5% non-fat dry milk. Both anti-BRCA1 antibodies were against a 17mer synthetic peptide corresponding to amino acids 1847–1863 of human BRCA1.

**Embryo collection and genotyping**

**Brcal Ko/+; Tg/+ or Brcal Ko/+ females were mated to Brcal Ko/+ or Brcal Ko/++; Tg/+ males, to obtain embryos, which were dissected from the uterine horns at various days of gestation. The day of vaginal plug was considered day 0.5. For genotyping of the embryos either the yolk-sac or the entire embryo was lysed in 20–40 µl of PCR lysis buffer (50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM MgCl2, 0.01 mg/ml gelatin, 0.45% Tween 20 and 0.45% NP40) containing 1 mg/ml Proteinase K at 55°C for 2–4 h. PCR primers and reaction conditions were used as described previously (18). Presence of the transgene was detected by PCR using a forward primer 5’-ACTGGCCAGTGATCATGAC-3’ and a reverse primer 5’-CTTTTTGTTTATTTCTCATGACCAC-3’ specific to the human **BRCAL** gene.

**Human lymphocytic cell lines**

Immortalized human lymphocytic cell lines containing 309T→G mutation in the **BRCAL** gene were obtained from L. Brody. The cells were cultured as described previously (16).

**ES cells containing human **BRCAL** BAC**

A Neo gene under the control of PGK promoter was introduced into the BAC vector by recombining as described previously (Court et al., manuscript in preparation). The BAC DNA was purified using plasmid purification kit (Qiagen) and resuspended in 1 × PBS at low concentration (100 ng/µl). Twenty micrograms of BAC-RP11–812O5 containing the missense mutations (4603G→T and 5586G→A) were electroporated into mouse AB2.2 ES cells and selected for resistance to G418 following standard protocol (43). G418 resistant clones were confirmed by hybridizing a Southern blot containing EcoRI digested ES cell DNA with **BRCAL**-specific probes. A 1265 bp probe from the 5’ end of **BRCAL** gene was generated by PCR using primers 5’-GTCTTCGATAACTGGCC-3’ and 5’-TTTTTCAACGGCAAGAGCA-3’ and a 244 bp probe from the 3’ end of **BRCAL** gene was generated by PCR using primers 5’-AGTGGACACTCTACCAGTG-3’, 5’-TTAAGG-GACCCCTGCATAGC-3’.

**ESE motif analysis**

Wild-type and mutant **BRCAL** exons 5, 14 and 23 were analyzed for putative ESE motifs using the web-based ESE finder located at http://exon.cshl.org/ESE/index.html.

**ACKNOWLEDGEMENTS**

We thank Drs Larry Brody and Barbara Weber for the human lymphocytic cell lines used in the study and to Dr Drakoulis Yannoukakos and Angela Ladopoulou for providing patient cDNA sample with missense mutation in codon 1823. We also thank Drs Jairaj Acharya, Usha Acharya, Larry Brody, Neal Copeland, Ira Daar and Sergey Kuznetsov for helpful discussion and critical review of the manuscript. We thank Linda Cleveland for help with sequencing; Dr Keiko Akagi for help with bio-informatics; Publication Department of NCI-Frederick for the illustrations. This research was supported by the National Cancer Institute, Department of Human and Health Services.

**REFERENCES**


