# Confirmation of *BRCA1* by analysis of germline mutations linked to breast and ovarian cancer in ten families

Lori S. Friedman, Elizabeth A. Ostermeyer, Csilla I. Szabo, Patrick Dowd, Eric D. Lynch, Sarah E. Rowell & Mary-Claire King

We provide genetic evidence supporting the identity of the candidate gene for *BRCA1* through the characterization of germline mutations in 63 breast cancer patients and 10 ovarian cancer patients in ten families with cancer linked to chromosome 17q21. Nine different mutations were detected by screening *BRCA1* DNA and RNA by single-strand conformation polymorphism analysis and direct sequencing. Seven mutations lead to protein truncations at sites throughout the gene. One missense mutation (which occurred independently in two families) leads to loss of a cysteine in the zinc binding domain. An intronic single basepair substitution destroys an acceptor site and activates a cryptic splice site, leading to a 59 basepair insertion and chain termination. The four families with both breast and ovarian cancer had chain termination mutations in the N-terminal half of the protein.

*Department of Molecular & Cell Biology and School of Public Health, University of California, Berkeley, California 94720, USA*

In October 1994, a candidate gene for *BRCA1*, which is responsible for inherited predisposition to breast and ovarian cancers in some families, was isolated by positional cloning[1]. In this report, we confirm that *BRCA1* is this predisposing gene by analysing germline mutations in the families that originally defined the linked phenotype[2], as well as in other families with breast cancer, and often ovarian cancer, linked to chromosome 17q21.

For this purpose, the *BRCA1* gene was screened using single strand conformation polymorphism (SSCP) analysis in 20 *BRCA1*-linked families, using both genomic DNA and cDNA prepared from lymphoblast RNA. Criteria for distinguishing cancer-predisposing mutations from polymorphisms in *BRCA1* were (i) cosegregation of the variant with breast cancer, and with ovarian cancer, if it appeared; (ii) absence of the variant in control chromosomes; and (iii) amino acid substitution in, or truncation of, the BRCA1 protein encoded by the variant sequence.

### Detection of mutations in genomic and cDNA

Primers used to screen genomic DNA and cDNA for variation in *BRCA1* by SSCP analysis are indicated in Table 1. All primer pairs were used to screen all 20 families by SSCP analysis. Primers flanking exons 2, 3, 4 (an Alu sequence not found in most cDNA clones), 5, 8, 9, 10 and 12–24 are from the ftp file at morgan.med.utah.edu[1]. New primers were designed to flank exon 1, which is 100 basepairs (bp) 5' of the start site and for which surrounding genomic sequence was reported in Genbank (L182093) as an STS on 17q21 (ref. 3). In order to screen splice junctions of exons 6 and 7, intron 6 was sequenced and primers defined to yield 200–250 bp amplified products flanking exons 6 and 7 independently. Primers were also designed to amplify exon 11 in 250–300 bp fragments, using the *BRCA1* cDNA sequence published as Genbank accession number U14680 (Table 1). New SSCP primers were designed to amplify from cDNA, and are designated c1-c11 in Table 1.

Each variant detected by SSCP analysis was sequenced using several templates: the variant band from the SSCP gel, genomic DNA and/or cDNA from the proband, and genomic DNA from family members. The 18 mutations (M) or polymorphisms (P) detected in genomic and/or cDNA are also indicated in Table 1 next to the primers with which each was observed.

### Analysis of *BRCA1* mutations

Two types of *BRCA1* variants fulfilling the criteria for cancer-predisposing mutations are illustrated by families 1 and 84. In family 1, an SSCP pattern reflecting a small deletion cosegregates with breast and ovarian cancer and the *BRCA1*-linked haplotype (Fig. 1). The haplotype analysis predicted that persons #1 and #8 did not inherit the linked chromosome, and the mutation analysis confirmed their *BRCA1* sequence was wild type. Thus their breast cancers at ages 71 and 45 are sporadic. The variant sequence shown has a 2 bp deletion at nucleotides (nts) 2800–2801 (codon 894), creating a stop at codon 901.

In family 84, an SSCP pattern representing a single nucleotide substitution cosegregates with breast cancer and the *BRCA1*-linked haplotype (Fig. 2). The variant genomic sequence has a thymine to guanine mutation at nt 300 of sequence U14680, resulting in a Cys61Gly substitution. This mutation removes the penultimate cysteine of the putative C3HC4 zinc-binding motif[4]—the

**Table 1 *BRCA1* genomic DNA and cDNA primers, polymorphisms and mutations**

| Primer pair | Tm | Size (bp) | DNA | U14680 bp | Forward primer | Reverse primer | Variant (bp) |
|---|---|---|---|---|---|---|---|
| exon1 | 55 | 315 | genomic | 1-100 | 5'- TAG CCC TTG GTT TCC GTG -3' | 5'- TCA CAA CGC CTT ACG CCT C -3' | |
| c1 | 55 | 253 | cDNA | 28-280 | 5'- ACC AGG CTG TGG GGT TTC -3' | 5'- TGG TTG AGA AGT TTC AGC ATG -3' | |
| exon2 | 60 | 250 | genomic | 101-199 | 5'- GAA GTT GTC ATT TTA TAA ACC TTT -3' | 5'- TGT CTT TTC TTC CCT AGT ATG T -3' | |
| exon3 | 60 | 300 | genomic | 200-253 | 5'- TCC TGA CAC AGC AGA CAT TTA -3' | 5'- TTG GAT TTT CGT TCT CAC TTA -3' | |
| c2 | 55 | 319 | cDNA | 212-530 | 5'- CAA GGA ACC TGT CTC CAC AAA GTG -3' | 5'- AAG TCT TTT GGC ACG GTT TCT G -3' | M(300); M(332) |
| exon4 | 60 | 200 | genomic | Alu | 5'- GTC AAA GAG ATA GAA TGT GAG C -3' | 5'- CCC GTC TCT ACA GAA AAC AC -3' | |
| exon5 | 60 | 200 | genomic | 254-331 | 5'- CTC TTA AGG GCA GTT GTG AG -3' | 5'- TTC CTA CTG TGG TTG CTT CC -3' | M(300) |
| exon 6 | 55 | 200 | genomic | 332-420 | 5'- CTT ATT TTA GTG TCC TTA AAA GG -3' | 5'- TTT CAT GGA CAG CAC TTG AGT G -3' | M(332) |
| exon 7 | 55 | 250 | genomic | 421-560 | 5'- CAC AAC AAA GAG CAT ACA TAG GG -3' | 5'- TCG GGT TCA CTC TGT AGA AG -3' | |
| c3 | 55 | 258 | cDNA | 390-647 | 5'- TGT GCT TTT CAG CTT GAC ACA GG -3' | 5'- CGT CTT TTG AGG TTG TAT CCG CTG -3' | |
| exon8 | 60 | 220 | genomic | 561-665 | 5'- TGT TAG CTG ACT GAT GAT GGT -3' | 5'- ATC CAG CAA TTA TTA TTA AAT AC -3' | P(intron) |
| c4 | 55 | 298 | cDNA | 581-878 | 5'- CCA ACT CTC TAA CCT TGG AAC TGT G -3' | 5'- TGC ACG CTT CTC AGT GGT GTT C -3' | |
| exon9 | 60 | 200 | genomic | 666-712 | 5'- CCA CAG TAG ATG CTC AGT AAA TA -3' | 5'- TAG GAA AAT ACC AGC TTC ATA GA -3' | |
| exon10 | 60 | 220 | genomic | 713-788 | 5'- TGG TCA GCT TTC TGT AAT CG -3' | 5'- GTA TCT ACC CAC TCT CTT CTT CAG -3' | |
| exon11Ai | 55 | 301 | genomic | 789-1090 | 5'- GGA ATT AAA TGA AAG AGT ATG AGC -3' | 5'- CTT CCA GCC CAT CTG TTA TGT TG -3' | |
| exon11A | 55 | 309 | both | 858-1166 | 5'- AAC ACC ACT GAG AAG CGT GCA G -3' | 5'- CTC ACA CAG GGG ATC AGC ATT C -3' | |
| exon11B | 55 | 300 | both | 1068-1363 | 5'- CAA CAT AAC AGA TGG GCT GGA AG -3' | 5'- ACG TCC AAT ACA TCA GCT ACT TTG G -3' | P(1186) |
| exon11B2 | 55 | 209 | both | 1209-1417 | 5'- CCT AGA GAT ACT GAA GAT GTT CCT TGG -3' | 5'- GCC AGT AAG TCT ATT TTC TCT GAA GAA CC -3' | |
| exon11C | 55 | 295 | both | 1299-1593 | 5'- GGT TCT GAT GAC TCA CAT GAT GGG -3' | 5'- TCT GTG GCT CAG TAA CAA ATG CTC -3' | |
| exon11D | 55 | 254 | both | 1505-1758 | 5'- GAA AAC CTA TCG GAA GAA GGC AAG -3' | 5'- TCA TCA CTT GAC CAT TCT GCT CC -3' | |
| exon11D2 | 55 | 272 | both | 1584-1855 | 5'- GAG CCA CAG ATA ATA CAA GAG CGT C -3' | 5'- GCA GAT TCT TTT TCG AGT GAT TCT ATT GGG -3' | |
| exon11E | 55 | 269 | both | 1716-1985 | 5'- ATC AGG GAA CTA ACC AAA CGG AG -3' | 5'- CGC ATG AAT ATG CCT GGT AGA AG -3' | |
| exon11E2 | 55 | 410 | both | 1718-2127 | 5'- TCA GGG AAC TAA CCA AAC GGA G -3' | 5'- CCA TGA GTT GTA GGT TTC TGC TG -3' | |
| exon11F | 55 | 273 | both | 1947-2219 | 5'- AGG CTG AGG AGG AAG TCT TCT ACC -3' | 5'- CAG CTC TGG GAA AGT ATC GCT G -3' | |
| exon11G | 55 | 319 | both | 2142-2460 | 5'- GCA ACT GGA GCC AAG AAG AGT AAC -3' | 5'- CCT GAG TGC CAT AAT CAG TAC CAG G -3' | P(2201); P(2430); M(2415) |
| exon11H | 55 | 312 | both | 2198-2509 | 5'- CAG CGA TAC TTT CCC AGA GCT G -3' | 5'- TCT GTT TTT GCC TTC CCT AGA GTG -3' | P(2430); M(2415) |
| exon11H2 | 55 | 286 | both | 2335-2620 | 5'- AAG TGT CTA ATA ATG CTG AAG ACC CC -3' | 5'- CCC AAT GGA TAC TTA AAG CCT TCT G -3' | P(2430); M(2415) |
| exon11I | 55 | 280 | both | 2485-2764 | 5'- GCA CTC TAG GGA AGG CAA AAA CAG -3' | 5'- CAT TCC TCT TCT GCA TTT CCT GG -3' | P(2430) |
| exon11I2 | 55 | 333 | both | 2598-2930 | 5'- GAA GGC TTT AAG TAT CCA TTG GG -3' | 5'- CTT ATC TTT CTG ACC AAC CAC AGG -3' | M(2800); M(2863) |
| exon11J | 55 | 288 | both | 2716-3003 | 5'- GCC AGT CAT TTG CTC CGT TTT C -3' | 5'- CGT TGC CTC TGA ACT GAG ATG ATA G -3' | M(2800); M(2863) |
| exon11K | 55 | 305 | both | 2897-3201 | 5'- TGC AGG CTT TCC TGT GGT TG -3' | 5'- GGC TAA TTG TGC TCA CTG TAC TTG G -3' | |
| exon11K2 | 55 | 296 | both | 2998-3293 | 5'- GCA ACG AAA CTG GAC TCA TTA CTC -3' | 5'- AAT ACT GGA GCC CAC TTC ATT AGT AC -3' | P(3232) |
| exon11L | 55 | 270 | both | 3138-3438 | 5'- TCA ATG TCA CCT GAA AGA GAA ATG G -3' | 5'- CAG GAT GCT TAC AAT TAC TTC CAG G -3' | P(3232); M(3238) |
| exon11M | 55 | 270 | both | 3357-3662 | 5'- TTG AAT GCT ATG CTT AGA TTA GGG G -3' | 5'- GAC GCT TTT GCT AAA AAC AGC AG -3' | |
| exon11N | 55 | 253 | both | 3552-3804 | 5'- GTT TGT TCT GAG ACA CCT GAT GAC C -3' | 5'- AGT GTT GGA AGC AGG GAA GCT C -3' | P(3667); M(3726) |
| exon11O | 55 | 289 | both | 3683-3791 | 5'- GAG TCC TAG CCC TTT CAC CCA TAC -3' | 5'- GTG ATG TTC CTG AGA TGC CTT TG -3' | |
| exon11P | 55 | 314 | both | 3857-4170 | 5'- CGT TGC TAC CGA GTG TCT GTC TAA G -3' | 5'- AGC CCG TTC CTC TTT CTT CAT C -3 | |
| c5-exon11Pi | 55 | 200 | genomic | 4105-4215 | 5'- AAA GCC AGG GAG TTG GTC TGA G -3' | 5'- GTG CTC CCA AAA GCA TAA A -3' | M(4184) |
| c5 | 55 | 312 | cDNA | 4105-4416 | 5'- AAA GCC AGG GAG TTG GTC TGA G -3' | 5'- TGA TGG AAG GGT AGC TGT TAG AAG G -3' | M(4184) |
| exon12 | 60 | 220 | genomic | 4216-4302 | 5'- GTC CTG CCA ATG AGA AGA AA -3' | 5'- TGT CAG CAA ACC TAA GAA TGT -3' | P(intron) |
| exon13 | 60 | 280 | genomic | 4303-4476 | 5'- AAT GGA AAG CTT CTC AAA GTA -3' | 5'- ATG TTG GAG CTA GGT CCT TAC -3' | |
| c6 | 55 | 287 | cDNA | 4323-4609 | 5'- CAT AAC CTG ATA AAG CTC CAG CAG G -3' | 5'- GAT GAC CTT TCC ACT CCT GGT TC -3' | |
| c7 | 55 | 290 | cDNA | 4445-4734 | 5'- GCG AAA TCC AGA ACA AAG CAC ATC -3' | 5'- GCT GTT GCT CCT CCA CAT CAA C -3' | |
| exon14 | 60 | 250 | genomic | 4477-4603 | 5'- CTA ACC TGA ATT ATC ACT ATC A -3' | 5'- GTG TAT AAA TGC CTG TAT GCA -3' | |
| exon15 | 60 | 250 | genomic | 4604-4794 | 5'- TGG CTG CCC AGG AAG TAT G -3' | 5'- AAC CAG AAT ATC TTT ATG TAG GA -3' | |
| c8 | 55 | 266 | cDNA | 4645-4910 | 5'- ACA TGC ACA GTT GCT CTG GGA G -3' | 5'- GGT TGA AGA TGG TAT GTT GCC AAC -3' | |
| exon16 | 60 | 375 | genomic | 4795-5105 | 5'- AAT TCT TAA CAG AGA CCA GAA C -3' | 5'- AAA ACT CTT TCC AGA ATG TTG T -3' | P(4956) |
| c9 | 55 | 307 | cDNA | 4837-5143 | 5'- ATG ACC CTG AAT CTG ATC CTT CTG -3' | 5'- GTG ATG TGG TGT TTT CTG GCA AAC -3' | P(4956) |
| c10 | 55 | 290 | cDNA | 5092-5381 | 5'- CCC CAG AAG AAT TTA TGC TCG TG -3' | 5'- TTC TCT TGC TCG CTT TGG ACC -3' | |
| exon17 | 60 | 350 | genomic | 5106-5193 | 5'- GTG TAG AAC GTG CAG GAT TG -3' | 5'- TCG CCT CAT GTG GTT TTA -3' | P(intron) |
| exon18 | 60 | 350 | genomic | 5194-5273 | 5'- GGC TCT TTA GCT TCT TAG GAC -3' | 5'- GAG ACC ATT TTC CCA GCA TC -3' | |
| exon19 | 60 | 220 | genomic | 5274-5310 | 5'- CTG TCA TTC TTC CTG TGC TC -3' | 5'- CAT TGT TAA GGA AAG TGG TGC -3' | |
| exon20 | 60 | 220 | genomic | 5311-5396 | 5'- ATA TGA CGT GTC TGC TCC AC -3' | 5'- GGG AAT CCA AAT TAC ACA GC -3' | |
| c11 | 55 | 318 | cDNA | 5331-5648 | 5'- GGA GAT GTG GTC AAT GGA AGA AAC -3' | 5'- TGC TAC ACT GTC CAA CAC CCA CTC -3' | |
| exon21 | 60 | 275 | genomic | 5397-5451 | 5'- AAG CTC TTC CTT TTT GAA AGT C -3' | 5'- GTA GAG AAA TAG AAT AGC CTC T -3' | |
| exon22 | 60 | 275 | genomic | 5452-5526 | 5'- TCC CAT TGA GAG GTC TTG CT -3' | 5'- GAG AAG ACT TCT GAG GCT AC -3' | |
| exon23 | 60 | 250 | genomic | 5527-5586 | 5'- CAG AGC AAG ACC CTG TCT C -3' | 5'- ACT GTG CTA CTC AAG CAC CA -3' | |
| exon24 | 60 | 275 | genomic | 5587-5711 | 5'- ATG AAT TGA CAC TAA TCT CTG C -3' | 5'- GTA GCC AGG ACA GTA GAA GGA -3' | M(5677) |

The names of 58 primer pairs used to screen 20 *BRCA1*-linked families by SSCP analysis are listed in the first column. $T_m$ indicates the annealing temperature for each PCR primer pair. The size of the PCR product is given in basepairs. The source of DNA template used for each primer pair was genomic DNA, cDNA prepared from lymphocyte RNA, or both genomic DNA and cDNA. Listed in column five are the nucleotides of the *BRCA1* cDNA sequence (U14680) that are analysed by the corresponding primer pair. Nucleotide locations of polymorphisms (P) and mutations (M) are indicated at their position in the *BRCA1* cDNA sequence.

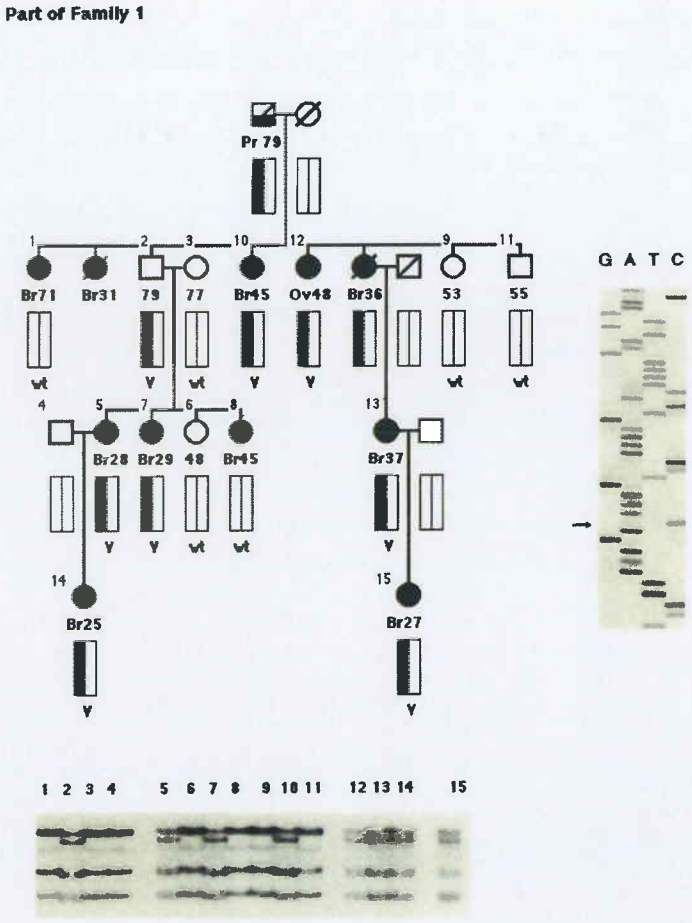


Fig. 1 Frameshift mutation in *BRCA1* linked to breast and ovarian cancer in family 1. Symbols under each relative indicate chromosome 17q21 haplotypes linked (filled rectangles) or not linked (open rectangles) to *BRCA1*. Cancers are indicated as follows: Br, breast cancer; Ov, ovarian cancer; Pr, prostate cancer. Ages at diagnosis are also indicated. V indicates variant sequence and wt indicates wildtype sequence for *BRCA1*, based on SSCP analysis of 15 relatives, shown below the pedigree. The mutant sequence, shown on the right, has a 2 bp deletion (AA) at nts 2800–2801 (arrow), leading to a stop at codon 901. The wild type sequence is TCCTTAAAGAAACAAAGTCCAAAGTC-ACTTTTGAATGTGAACA, with the mutated sequence underlined.

only recognizable domain so far in the BRCA1 protein. This variant does not appear in 120 control chromosomes, suggesting that it is not simply a rare polymorphism.

In all, we identified mutations cosegregating with breast and ovarian cancer and linked to *BRCA1* in ten families (Table 2). None of these mutations has been previously observed (D. Goldgar, personal communication)[1]. The ten families have nine different mutations: the same thymine to guanine substitution at nt 300 (the Cys61Gly mutation) appeared twice, in families 84 and 4. This mutation probably arose independently in the two families, because chromosome 17q21 marker alleles defining the *BRCA1* haplotypes of the two families differ. The ancestry of family 84 is Polish and of family 4 German; insofar as the families are aware, they are not related. The only other missense mutation in our families leads to the substitution of asparagine for serine at codon 1040 in all patients with breast cancer linked to *BRCA1* in family 14. The functional significance of this mutation is currently unknown and might represent either a disease-predisposing allele or a rare polymorphism.

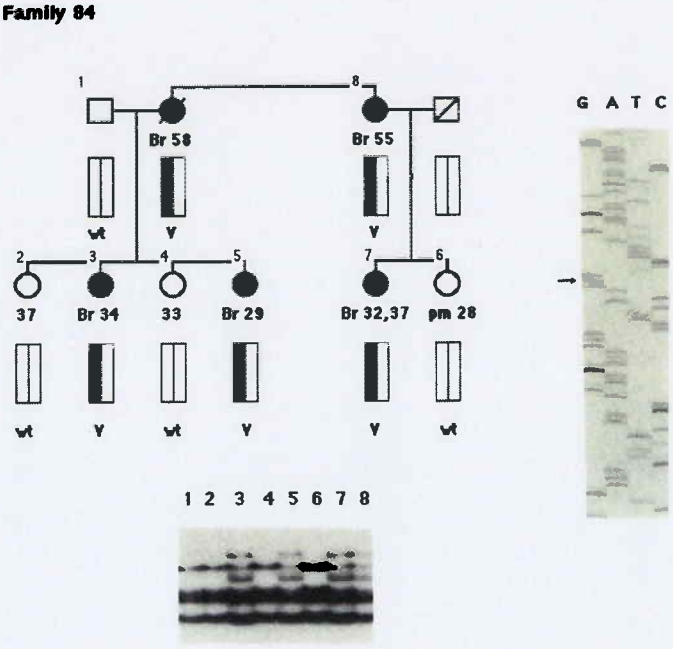


Fig. 2 Missense mutation in the zinc-binding domain of BRCA1 linked to breast cancer in Family 84. Symbols are as described in Fig. 1; individual #6 had a prophylactic mastectomy at age 28, ten years prior to entrance into the study. The mutant sequence, shown on the right, illustrates the T to G substitution at nt 300 (arrow), which leads to replacement of the penultimate Cys in the zinc binding domain with a Gly residue. This Cys61Gly mutation also occurred, probably independently, in Family 4. The wild type sequence is GCTGAAACTTCTCAACCAGAAGAAA-GGGCCTTCACAGTGTGTCCTTTATGTAAGAATGATATA-ACCAAAAGGTAT, with the mutated sequence underlined.

**Table 2a Germline mutations cosegregating with breast and ovarian cancer in *BRCA1*-linked families**

| Family | *BRCA1* nt | Exon | Mutation | Amino acid change | Predicted effect |
|---|---|---|---|---|---|
| 4 | 300 | 5 | T to G substitution (TGT→GGT) | Cys61Gly | lose zinc-binding motif |
| 84 | 300 | 5 | T to G substitution (TGT→GGT) | Cys61Gly | lose zinc-binding motif |
| 82 | 332 | Intron 5 | T to G substitution→59 bp insertion | 75Stop | protein truncation |
| 3 | 2415 | 11 | AG deletion | Ser766Stop | protein truncation |
| 1 | 2800 | 11 | AA deletion | 901Stop | protein truncation |
| 102 | 2863 | 11 | TC deletion | Ser915Stop | protein truncation |
| 14 | 3238 | 11 | G to A substitution (AGC→AAC) | Ser1040Asn | missense |
| 74 | 3726 | 11 | C to T substitution (CGA→TGA) | Arg1203Stop | protein truncation |
| 2 | 4184 | 11 | TCAA deletion | 1364Stop | protein truncation |
| 77 | 5677 | 24 | A insertion | Tyr1853Stop | protein truncation |

**Table 2b Sites, ages at diagnosis and laterality of cancers associated with *BRCA1* mutations**

| Family | Breast cancer | | Ovarian cancer | | Other cancers with with *BRCA1* mutations |
|---|---|---|---|---|---|
| | Cases (bilateral) | Avg. age at dx | Cases | Avg. age at dx | |
| 4 | 10 (1) | 39.7 | | | |
| 84 | 5 (1) | 41.6 | | | |
| 82 | 6 (2) | 40.3 | 2 | 52.1 | Small intestine (dx 46) |
| 3 | 3 (0) | 37.3 | 4 | 46.3 | |
| 1 | 16 (0) | 37.8 | 2 | 54.5 | 3 Prostate (dx 57, 77, 79) |
| 102 | 5 (0) | 45.2 | 3 | 51.7 | |
| 14 | 6 (2) | 51.1 | | | |
| 74 | 6 (1) | 45.5 | | | Kidney (dx 55); Colon (dx. 40) |
| 2 | 5 (1) | 38.2 | | | |
| 77 | 7 (1) | 32.9 | | | |

However, we have not found the mutation in 120 control chromosomes.

The remaining seven mutations all lead to truncated proteins, mostly due to small insertions or deletions of 1 to 4 bp (Table 2). Four mutations are immediate stops; the others are frameshifts leading to stops after 4, 7 and 9 codons, respectively. In family 3, one AG doublet is deleted from an $(AG)_3$ repeat; the other small length changes appeared to be random throughout the sequence. In family 82, a thymine to guanine substitution occurs within the consensus sequence 5' of the exon 6 splice acceptor site (Fig. 3). This substitution leads to the preferential activation in intron 5 of a cryptic splice acceptor[4]. This generates a mutant transcript containing a 59 bp insertion of intronic sequence 5' of exon 6, causing a stop at codon 75. Both the wild type and mutant transcripts are detected in lymphocyte cDNA with primer pair c2.

One variant observed only at the cDNA level could not yet be classified as a predisposing mutation. In two patients from family 5 with *BRCA1*-linked breast and ovarian cancer, cDNA prepared from lymphoblast RNA lacked exon 3. (No RNA from breast tumour is available for this family.) Exon 3 splice junction sequences are normal in genomic DNA from these patients. The absence of exon 3 has not been observed in lymphoblast cDNA from controls or from other breast cancer families. This case differs from the family 82 variant in that in family 5, an entire exon is

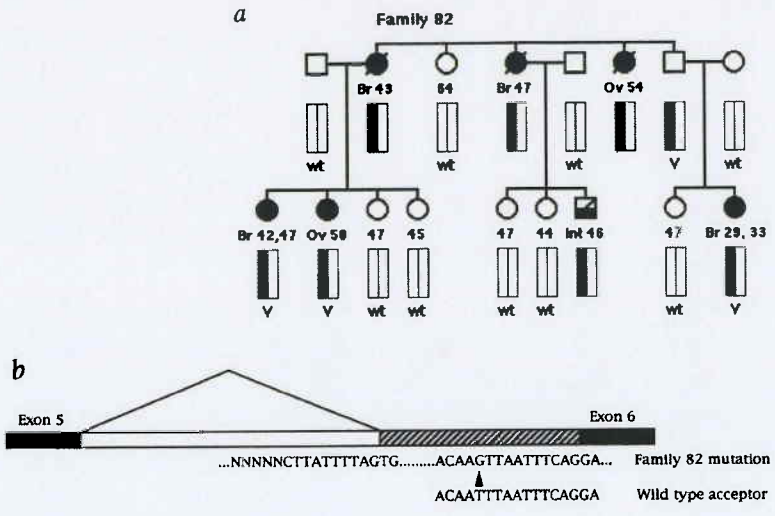


Fig. 3 Aberrant splicing leading to a truncated BRCA1 protein in Family 82. *a*, *BRCA1*-linked patients and unaffected relatives of Family 82. Mutant (V) or wildtype (wt) sequence was confirmed in living relatives. (Symbols are as given in Fig. 1). *b*, Aberrant splicing of exons 5 and 6 in Family 82. A cryptic splice acceptor site in intron 5 (open box) leads to insertion of 59 bp (hatched box) into the mutant transcript. Sequences of exons 5 and 6 (filled boxes) are normal. The cryptic acceptor is activated by a T to G substitution in the usual acceptor site. The concensus splice acceptor sequence is $(Py)_{12}$NCAG/N.

**Table 3 Neutral polymorphisms in *BRCA1***

| *BRCA1* nt | Nucleotide variant | Amino acid variant | Allele frequencies |
|---|---|---|---|
| intron 8 | T deleted | non-coding | .67/.33 |
| 1186 | A/G | Gln356Arg | .82/.13 |
| 2201 | C/T | silent | |
| 2430 | T/C | silent | .67/.33 |
| 3232 | A/G | Glu1038Gly | .67/.33 |
| 3667 | A/G | Lys1183Arg | .67/.33 |
| intron 12 | CA deleted | non-coding | .95/.05 |
| 4956 | A/G | Ser1613Gly | .67/.33 |
| intron 16 or 17 | not sequenced | non-coding | .69/.31 |

missing and no intronic sequence has been introduced. It is possible that absence of exon 3 represents alternative splicing characteristic of lymphoblasts we immortalized for this family, rather than having any relationship to cancer predisposition. Alternatively, an inherited mutation in genomic DNA leading to consistent aberrant splicing would strongly suggest that the variant is disease-related.

### *BRCA1* polymorphisms

Nine polymorphisms were detected in *BRCA1*, six in the coding sequence and three in introns (Table 3). Four polymorphisms alter amino acid residues. Five polymorphisms are in complete disequilibrium with each other, and a sixth is in partial disequilibrium, based on genotypes of 42 chromosomes. Identification of these polymorphisms enabled us to screen for mutations causing loss of *BRCA1* transcript in the ten families for which no mutations were found by SSCP analysis of genomic DNA and cDNA.

Genomic DNAs from probands of these families were genotyped for polymorphisms in the *BRCA1* coding sequence. Probands of five families with undetected mutations were heterozygous for one or more polymorphisms. The cDNA of these five probands were also heterozygous, indicating that mRNA was being expressed from both *BRCA1* alleles. For the remaining five families, patients heterozygous for coding sequence polymorphisms will be identified and the same test carried out.

### Discussion

Most of the *BRCA1* mutations we have identified result in premature termination of protein translation. Such truncated proteins could conceivably compete with wildtype protein, thereby acting as dominant mutations. On the other hand, loss of the wildtype allele in the tumours of patients with *BRCA1* germline mutations would suggest the opposite role — that the termination mutation was recessive. Paraffin blocks with adequate malignant cells for loss-of-heterozygosity (LOH) studies were available for one tumour from family 1, five tumours from family 3 and two tumours from family 82. For all eight of these tumors, the wildtype *BRCA1* allele was lost, an observation consistent with recessive mutations[5].

The number of families with *BRCA1* mutations identified is still very small, so associations between genotype and phenotype are not at all certain. Cancer-related mutations occur throughout the *BRCA1* sequence (Table 2). In our series, all four families with both ovarian and breast cancer had truncation mutations in the first half of the protein. However, termination of the BRCA1 protein only 11 amino acids from the C-terminal end was associated with very early-onset breast cancer in family 77. Bilateral breast cancer was associated with mutations throughout the sequence, although ascertainment of bilaterality is biased in very high-risk families in that some women choose bilateral mastectomy following their first diagnosis of breast cancer. In this series, there is no association between age at onset and locale or type of mutation, but these families were selected for early age at breast cancer diagnosis and probably do not represent *BRCA1* patients generally. Better evaluation of outcomes associated with *BRCA1* mutations of different types and at different locations in the sequence will be based on identifying germline *BRCA1* mutations in population-based series of cases not selected for family history.

The mutations found, and not found, in these 20 families illustrate the complexities of screening for inherited *BRCA1* mutations in families, let alone in the general population[6,7]. First, mutations were identified in only 50% of the families in this series, although both genomic DNA and RNA were screened by SSCP analysis and all families have convincing evidence of *BRCA1* linkage[2,8]. The families for which *BRCA1* mutations have not yet been identified are as likely to carry mutations as those successfully screened. The posterior probability of linkage is ≥.95 for families with and without definable mutations. The two groups of families are similar in number and ages at diagnosis of breast and ovarian cancers linked to *BRCA1*; ovarian cancer occurred in four of ten families with mutations found and in five of ten families with mutations still undetected. Additional mutations may prove detectable by other screening techniques such as mismatch cleavage analysis, or by Southern or northern hybridization. However, it is likely that a large fraction of mutations will only be detectable when genomic sequence of the entire gene is available. The importance of genomic sequence is illustrated by the successful detection of two mutations in cDNA: the 59 bp intronic insertion in family 82 reported here and the loss of transcript in family 2035 previously described[1]. The intronic mutation leading to aberrant splicing in family 82 could only be identified from genomic sequence. In general, mutations in introns and promoter regions can only be identified reliably when sequence of the entire 80 kb *BRCA1* gene is known and critical regions identified therein. The alternative would be to screen breast or ovarian RNA for mutations using multiple techniques, an approach not feasible on a wide scale.

Second, most *BRCA1* variants are not predisposing mutations. Most individuals, including women at high risk of breast cancer, are heterozygous at one or more sites in the *BRCA1* gene. Most of these variants are polymorphisms, some of which will be revealed, after sequencing, to be noncoding or silent. However, in this series, four missense mutations were revealed that altered peptide sequence. Because they occurred at similar frequencies in cases and controls, we concluded that they are polymorphisms. Even by screening cases and controls, some rare polymorphisms may be mistakenly identified as predisposing mutations. The mutation at nt 3238 in family 14 may prove to be of this type, and we would not predict risk for an asymptomatic woman based on that

mutation without determining its biological function.

Third, there are many predisposing mutations in *BRCA1*, and the risks associated with most are not known. The mutations identified so far have high penetrance, because they were identified in families ascertained for very high lifetime risk of breast cancer. Whether other *BRCA1* mutations are associated with more moderate risk, possibly influenced by other genetic or environmental factors, remains to be seen.

## Methodology

**SSCP.** PCR was carried out in 50 µl volumes containing: 50 ng cDNA or genomic DNA, 1× PCR buffer (Boehringer Mannheim), 200 µM dATP, dGTP, dTTP (Boehringer Mannheim), 10 µM dCTP (Boehringer Mannheim), 50 pmoles each primer from Table 1 (Operon), 0.5 µCi $^{32}$P-dCTP (NEN Dupont), 1.25 U *Taq* DNA polymerase (Boehringer Mannheim). PCR conditions for primer pairs were 35 cycles of 94 °C, 45 s; annealing temperature ($T_m$ from Table 1) 30 s; 72 °C; 30 s. PCR template was lymphocyte cDNA or genomic DNA from members of *BRCA1*-linked families who carry the predisposing haplotype. Amplified samples were diluted 1:10 in formamide buffer (98% formamide, 10 mM EDTA pH 8, 0.05% Bromophenol blue, 0.05% Xylene cyanol), held at 95 °C for 5 min, then cooled rapidly to 4 °C. and held for 5 min. For each sample, 5 µl was loaded onto an SSCP gel and run at 6 W (constant power) for 16 h in 0.6× TBE at room temperature. (An 80 ml gel solution contains: 0.5× MDE (AT Biochem), 0.6× TBE, 160 µl 25% ammonium persulphate, 38 µl TEMED.) Gels were dried on a vacuum gel dryer and exposed to film for 12–24 h with an intensifying screen. Variant bands were cut out of the gel and rehydrated in 100 µl water.

**Sequencing.** PCR was carried out in 50 µl volumes containing: 50 ng DNA or cDNA or 1 µl of rehydrated SSCP gel fragment solution, 1× PCR buffer (Boehringer Mannheim), 200 µM dATP, dCTP, dTTP, dGTP (Boehringer Mannheim), 1.25 U *Taq* polymerase (Boehringer Mannheim), 50 pmoles each primer (Operon) that detected the SSCP variant. Cycling conditions were 35 cycles of 94 °C, 45 s, annealing temperature ($T_m$ in Table 1) 30 s, 72 °C, 30 s. Direct sequencing of the double-stranded DNA from the PCR amplification was performed with the USB PCR product sequencing kit, using $^{35}$S ATP as the isotope. Sequencing primers were the same as those used to amplify the template. Sequence was run on 6% acrylamide/bis gels at 70 W constant power for 2 h, and exposed to X-ray film for 12–24 h. Mutations were confirmed in multiple members of each family by amplifying and directly sequencing from genomic DNA.

Intron 6 sequence was obtained by amplifying genomic DNA using the primer pair exon 6&7 (ftp file at morgan.med.utah.edu), with the reaction mixture described above. Direct sequencing was done on the 800 bp product as described above.

**Acknowledgements**

*This work was supported in part by NIH grant R01 CA27632. L.S.F. is a Komen Foundation Fellow; C.I.S. is an NIH Fellow; M-C.K. is an American Cancer Society Professor.*

Received 8 November; accepted 9 November 1994.

1. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71 (1994).
2. Hall, J.M. *et al.* Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **250**, 1684–1689 (1990).
3. Albertsen, H.M. *et al.* A physical map and candidate genes in the BRCA1 region on chromosome 17q12–21. *Nature Genet.* **7**, 472–479 (1994).
4. Sharp, P.A. Split genes and RNA splicing. *Cell* **77**, 805–815 (1994).
5. Friedman, L.S. *et al.* The search for BRCA1 and the cloning of 22 genes from chromosome 17q21. *Cancer Res.* (in the press).
6. American Society of Human Genetics. Statement on genetic testing for breast and ovarian cancer. *Am. J. hum. Genet.* **55**, 860 (1994).
7. Rowell, S., Newman, B., Boyd, J. & King, M-C. Inherited predisposition to breast and ovarian cancer. *Am. J. hum. Genet.* **55**, 861–865 (1994).
8. Arena, J.F. *et al.* Inherited breast cancer: Confirmation of linkage to BRCA2, possible linkage to the estrogen receptor, and the clinical implications of genetic complexity. *J. Am. med. Assoc.* (in the press).