Am. J. Hum. Genet. 70:858–865, 2002

Distinct *BRCA1* Rearrangements Involving the *BRCA1* Pseudogene Suggest the Existence of a Recombination Hot Spot

Nadine Puget,[1,2] Sophie Gad,[3] Laure Perrin-Vidoz,[1] Olga M. Sinilnikova,[2] Dominique Stoppa-Lyonnet,[3] Gilbert M. Lenoir,[1,*] and Sylvie Mazoyer[1]

[1]Laboratoire de Génétique, Unité Mixte de Recherche 5641 CNRS, Université Claude Bernard, and [2]International Agency for Research on Cancer, Lyon, France; and [3]Service de Génétique Oncologique, Institut Curie, Paris

The 5′ end of the breast and ovarian cancer–susceptibility gene *BRCA1* has previously been shown to lie within a duplicated region of chromosome band 17q21. The duplicated region contains *BRCA1* exons 1A, 1B, and 2 and their surrounding introns; as a result, a *BRCA1* pseudogene (Ψ*BRCA1*) lies upstream of *BRCA1*. However, the sequence of this segment remained essentially unknown. We needed this information to investigate at the nucleotide level the germline deletions comprising *BRCA1* exons 1A, 1B, and 2, which we had previously identified in two families with breast and ovarian cancer. We have analyzed the recently deposited nucleotide sequence of the 1.0-Mb region upstream of *BRCA1*. We found that 14 blocks of homology between the tandemly repeated copies (cumulative length = 11.5 kb) show similarity of 77%–92%. Gaps between blocks result from insertion or deletion, usually of repetitive elements. *BRCA1* exon 1A and Ψ*BRCA1* exon 1A are 44.5 kb apart. In the two families with breast and ovarian cancer mentioned above, distinct homologous recombination events occurred between intron 2 of *BRCA1* and intron 2 of Ψ*BRCA1*, leading to 37-kb deletions. Breakpoint junctions were found to be located at close but distinct sites within segments that are 98% identical. The mutant alleles lack the *BRCA1* promoter and harbor a chimeric gene consisting of Ψ*BRCA1* exons 1A, 1B, and 2, which lacks the initiation codon, fused to *BRCA1* exons 3–24. Thus, we report a new mutational mechanism for the *BRCA1* gene. The presence of a large region homologous to *BRCA1* on the same chromosome appears to constitute a hot spot for recombination.

Introduction

More and more germline rearrangements have been identified in the *BRCA1* gene (MIM 113705) during the last 4 years, in families with breast and/or ovarian cancer (Petrij-Bosch et al. 1997; Puget et al. 1997; Swensen et al. 1997; Montagna et al. 1999; Puget et al. 1999*a,* 1999*b;* Payne et al. 2000; Rohlfs et al. 2000*a,* 2000*b;* Unger et al. 2000; Gad et al. 2001*a,* 2001*b*). This is due to the fact that a growing number of laboratories screening for mutations in *BRCA1* use techniques allowing their identification—that is, Southern blot analysis (Swensen et al. 1997; Montagna et al. 1999; Puget et al. 1999*a;* Rohlfs et al. 2000*b;* Unger et al. 2000; Lahti-Domenici et al. 2001), cDNA analysis (Petrij-Bosch et al. 1997; Puget et al. 1997, 1999*b;* Rohlfs et al. 2000*a*), DNA combing (Gad et al. 2001*a,* 2001*b*), long-range PCR (Payne et al. 2000), and semiquantitative PCR (Robinson et al. 2000). Indeed, 20 different large alterations have been reported, among which 14 are fully characterized. All these rearrangements share several characteristics with the far more numerous point mutations found in *BRCA1:* they are scattered throughout the gene, >60% are unique, and the recurrence of certain mutations is due to founder effects (Petrij-Bosch et al. 1997; The *BRCA1* Exon 13 Duplication Screening Group 2000; Rohlfs et al. 2000*b*). The breakpoint junction of only one rearrangement falls within an exonic sequence (Payne et al. 2000). In all other cases, they fall within intronic *Alu* sequences, which represent 41.5% of the *BRCA1* gene (Smith et al. 1996).

In a previous study (Puget et al. 1999*a*), we identified two different germline deletions that involve exons 1 and 2 and whose breakpoints could not be characterized because knowledge of the structure of the region upstream of *BRCA1* was not comprehensive. A *BRCA1* pseudogene, Ψ*BRCA1*, had been shown to lie ~30 kb upstream of *BRCA1* (Barker et al. 1996; Brown et al. 1996), because of the duplication of a region containing *BRCA1* exons 1 and 2. The first exons of the *NBR2* (next to *BRCA1* 2) gene, located between *BRCA1* and its pseudogene, are homologous to the first exons of the *NBR1* (next to *BRCA1* 1) gene, which lies head-to-head with Ψ*BRCA1* (Xu et al. 1997). The deletions we iden-

Received October 26, 2001; accepted for publication January 2, 2002; electronically published March 5, 2002.

Address for correspondence and reprints: Dr. Sylvie Mazoyer, Laboratoire de Génétique, Unité Mixte de Recherche 5641 CNRS, Faculté de Médecine, 8 avenue Rockefeller, 69373 Lyon cedex 08, France. E-mail: smazoyer@rockefeller.univ-lyon1.fr

* Present affiliation: Institut Gustave Roussy, Villejuif, France.

0002-9297/2002/7004-0004$15.00

tified in the two families with breast and ovarian cancer, F32 and F3514, did not include Ψ*BRCA1* exons 1 and 2 (Puget et al. 1999*a*) but were found to totally encompass *NBR2* (data not shown). We suspected that the existence of the duplicated regions could represent a recombination hot spot and could be involved in predisposition to breast and ovarian cancer in families F32 and F3514.

On the basis of a color bar code of the *BRCA1* region on combed DNA in carriers, we show evidence in favor of this hypothesis. The recent availability of the full sequence of the 1-Mb region upstream of *BRCA1* allowed us (1) to describe its precise organization, which revealed its high potential of recombination and (2) to use PCR to characterize the deletions' breakpoints. This showed that both recombinations occurred between intron 2 of the pseudogene and intron 2 of the gene, at close, but not identical, sites. These findings strengthen the hypothesis that the presence of a large region homologous to *BRCA1* and on the same chromosome may constitute a hot spot for recombination.

Subjects and Methods

Sequence Data Analyses

The sequence of BAC RP11-242D8 (GenBank accession number AC060780) was identified by performing a BLAST search with *BRCA1* intron 2, using a "human repeats" filter in the High-Throughput Genome Sequences database. The sequence was, at the time this manuscript was written, a "working draft" consisting of three contigs and not yet annotated. Because the first 39 kb of the first contig (152,199 bp) is the only one to contain *BRCA1* sequences (from exon 1A to most of intron 12), it is certain to be a clone end. Therefore, coordinates given in the present report will not change when the sequence is updated. Exons of *NBR1* and *NBR2* were positioned within the sequence by use of the SIM4 program, while "exons" 1A, 1B, and 2 of Ψ*BRCA1* were defined by means of homology with *BRCA1* exons, using the BLAST program. The extent of the similarity between gene and pseudogene was investigated using the LALIGN program. Blocks were arbitrarily defined as regions of homology within which gaps did not exceed 50 nucleotides. RepeatMasker was run, to identify interspersed repeat sequences. The sequence between the ends of Ψ*BRCA1* and *NBR2* (nucleotides 59596–67624, referring to the AC060780 sequence) was analyzed by use of the NIX analysis program (nucleotide identification of unknown sequences, U.K. Medical Research Council Human Genome Mapping Project), a Web-based package of gene-analysis software (including GRAIL, GENSCAN, Fgenes, Fex, Hexon, Genemark, Genefinder, BLAST, Polyah, RepeatMasker, and TRNAscan).

Color Bar Coding the BRCA1 *Region on Combed DNA from Families F32 and F3514*

DNA samples were extracted from lymphoblastoid cell lines of *BRCA1* carriers belonging to families F32 and F3514, as reported elsewhere (Michalet et al. 1997). Combing relies on homogeneous stretching of DNA molecules at a constant rate of 2 kb/μm (Bensimon et al. 1994; Michalet et al. 1997). FISH was performed on combed DNA, using six probes generated by random labeling (as described by Gad et al. [2001*a*]) of a PAC covering the whole *BRCA1* region (PAC103O14 from the chromosome 17 library of the Lawrence Livermore National Laboratory [Brown et al. 1995]), a rearranged cosmid covering the *BRCA1* promoter region (cosmid ICRFc105D06121 from the Ressourcen Zentrum Primary Database [Brown et al. 1996; Gad et al. 2001*a*]), and four long-range (LR) PCR products (Expand Long Template PCR System, Roche Diagnostics). Three PCR products covered exons of the *BRCA1* gene, and the remaining one covered exons 2–4 of the *NBR2* gene (Gad et al. 2001*a*; primers used are available at the Institut Curie Web site).

Full signals were observed under an epifluorescence DMRB microscope (Leica) and were captured with IPLab Spectrum-SU2 software (Vysis), using an NU 200 CCD camera (Photometrics). Image analyses were performed with CartographiX software (X. Michalet, Institut Pasteur, Paris), which permits the measurement of fragments of full signals with respect to the constant stretching rate of 2 kb/μm (Michalet et al. 1997; Gad et al. 2001*a*).

Cloning of Breakpoints in Families F32 and F3514

PCRs were performed using the Expand Long Template PCR System (Roche) according to the manufacturer's recommendations. Among *BRCA1* carriers in families F32 and F3514, forward primer pintron2.2F (ACC TAA AAT TCC TTC TGC TGG AC), which is located within intron 2 of the pseudogene at the beginning of block 11 (nucleotides 71699–71677), and reverse primer intron2.R2 (CCT TGG GCT AAC CAC TCT ACC), which is located in intron 2 (nucleotides 33948–33968), gave products of ~800 bp for F32 and F3514 carriers' DNA only, the normal fragment being far too large (37.8 kb) to be amplified. These fragments were then sequenced.

Screening for BRCA1 *Rearrangements Involving* ΨBRCA1 *in Families with Breast and Ovarian Cancer*

An original set of ~100 families, ascertained by H. T. Lynch (Creighton University School of Medicine), were

screened for mutations in the coding region and splice sites of the *BRCA1* and *BRCA2* genes by use of heteroduplex analysis and the protein truncation test (Mazoyer et al. 1996; Serova et al. 1996; Puget et al. 1997, 1999*a,* 1999*b*; Serova et al. 1997). Of the families that remained negative for mutations, 13 were used in this analysis. Families were selected according to the following criteria: (1) at least 5 cases of breast cancer, of which at least 4 were diagnosed at <60 years of age; (2) at least 4 cases of breast cancer, of which at least 2 were diagnosed at <45 years of age; (3) at least 1 case of ovarian cancer diagnosed at any age and 3 cases of breast cancer diagnosed at <60 years of age; or (4) at least 2 cases of ovarian cancer diagnosed at any age and 1 case of breast cancer diagnosed at <45 years of age or 2 cases of breast cancer diagnosed at any age. Among these families, seven are affected by breast cancer alone (one includes one case of male breast cancer), whereas the remaining six are families with breast and ovarian cancer (one includes one case of male breast cancer).

Screening for rearrangements involving Ψ*BRCA1* was conducted with the following primer pairs: NBR1.1AF (CCT GAG GCC TGA ATA TCA GC)/intron2.R4 (CTA TCC TCT CAA CGA CAC CGA T) to amplify from blocks 5–8; pintron2.7F (TCA AGG AAA TTT TCT TTT GTG C)/intron2.R7 (TGT GGA GTT TCC CCC ATT CT) to amplify blocks 10–14.

The expected size of the mutant product in the presence of a rearrangement is 2.4–2.9 kb for the first pair and ~7 kb for the second one. All screened DNA samples were subjected to a preliminary PCR with primers giving rise to a 7.2-kb fragment, in order to check their quality. A positive control—that is, carriers of *BRCA1* from F32 and F3514, could only be used with primers pintron2.7F and intron2.R7. More information about screening protocols, including the choice of primers, can be obtained upon request from the corresponding author.

Results

Structure of the Region Upstream of the BRCA1 *Gene*

The complete genomic sequence of the human *BRCA1* gene (GenBank accession number L78833), which begins 2,965 nucleotides upstream of exon 1A and therefore contains *NBR2* exons 1A and 1B, was made available in 1996, 2 years after the *BRCA1* cDNA was cloned. Comparison of this sequence with available *NBR1* and Ψ*BRCA1* sequences (U72483 and U77841) showed an overall similarity of 90.3%, scattered over 3,865 nucleotides. All attempts to amplify the region between Ψ*BRCA1* and *BRCA1* by long-range PCR with a wide collection of primers have been unsuccessful, either because this region was larger than the ~30 kb proposed or because it contained sequences difficult to amplify or for both reasons. In the same way, we were unable to amplify the region between Ψ*BRCA1* and *NBR2* or that between *NBR2* exon 2 and *BRCA1*.

To obtain more information on the structure of this region, we developed a restriction map of a 60-kb P1 clone (P1 1141 [Swensen et al. 1997]), from which we deduced that most of *BRCA1* intron 2 was likely to be contained within the duplicated region (data not shown). The deduction was later confirmed with the recently deposited sequence of BAC RP11-242D8 (GenBank accession number AC060780), which, at the present time, consists of three contigs, among which a 152,199-nucleotide piece comprises *NBR1*, Ψ*BRCA1*, *NBR2,* and 39,879 nucleotides of *BRCA1*. Analysis of this sequence revealed that *BRCA1* exon 1A and Ψ*BRCA1* exon 1A are 44,490 nucleotides apart. The duplicated segments extend over 14,491 and 24,970 bp in the *NBR2/BRCA1* and the *NBR1/*Ψ*BRCA1* regions, respectively, and are fragmented in 14 blocks of homology (fig. 1 and table 1), ranging from 77 to 1,555 bp (total 11,429 bp). The percentage of identity between gene and pseudogene within blocks varies from 76.9 to 92.1 (table 1). Gaps between blocks (61–7,457 bp) result from insertion or deletion mostly of repetitive elements in *BRCA1* or Ψ*BRCA1* (fig. 1), with two exceptions: (1) a 3,794-bp region containing repetitive elements, nonspecific sequences, and *NBR1* exon 2 (between blocks 1 and 2) and (2) a large P1 ribosomal protein pseudogene in the Ψ*BRCA1* region (between blocks 5 and 6). The sequence between Ψ*BRCA1* and *NBR2* is 8,029 nucleotides and was analyzed by the use of NIX, which incorporates a number of independent gene-prediction tools. Although a couple of genes were predicted (data not shown), no known transcribed sequence was identified. RepeatMasker indicates that interspersed repeats account for 56.5% of this region.

Cloning of the Breakpoints of the Deletions in Families F32 and F3514

In a previous study, we analyzed 78 DNA samples by quantitative Southern blot hybridization, to assess the proportion of rearrangements in the *BRCA1* mutation spectrum (Puget et al. 1999*a*). Densitometric analysis of the blots showed, in two American families with breast and ovarian cancer (F32 and F3514), deletions of *BRCA1* exons 1 and 2 but not of exon 3 or Ψ*BRCA1* exons 1 and 2 (Puget et al. 1999*a*). Allelotyping with three *BRCA1* intragenic markers revealed that these two families do not share a haplotype (Puget et al. 1999*a*). Further Southern blot analysis with more probes showed that these deletions totally encompass *NBR2* (data not shown). To estimate the extent of the deleted regions in F32 and F3514, the *BRCA1* gene region was analyzed using the dynamic molecular-combing technique in association with FISH

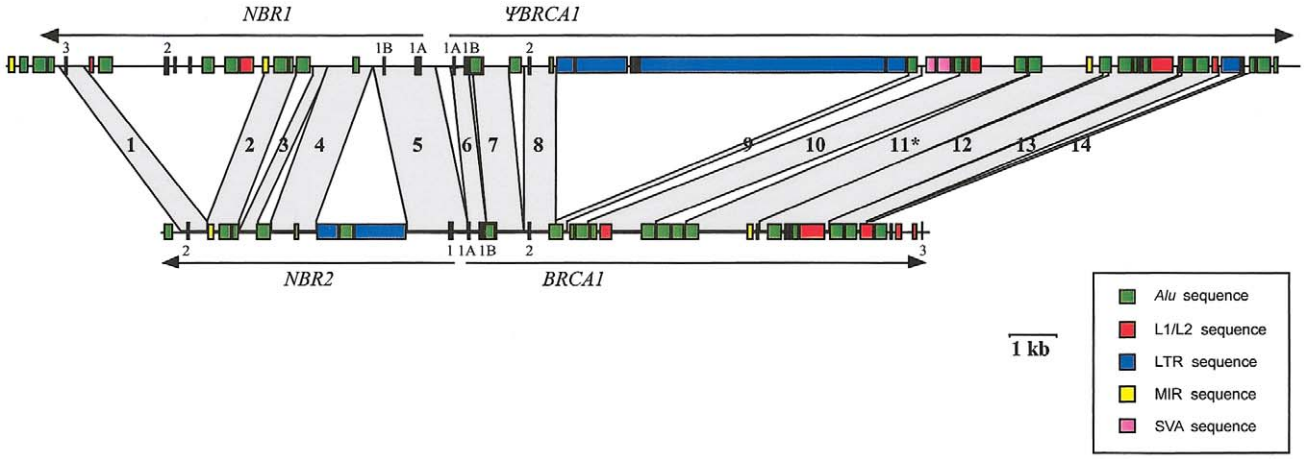

Figure 1 Schematic representation of the duplicated regions showing the homology between *NBR1* and *NBR2*, and between Ψ*BRCA1* and *BRCA1*. Blackened boxes represent exons, and shaded boxes represent blocks of homology, which were arbitrarily numbered. Repetitive elements, i.e., *Alu,* L1/L2, LTR, SVA, and MIR sequences are also shown. Block 11, which contains the breakpoints of both deletions in families F32 and F3514, is emphasized with an asterisk (*). The schema is drawn to scale.

(Bensimon et al. 1994; Michalet et al. 1997). Homogeneously stretched DNA molecules were visualized with a four-color bar code, allowing the analysis of the whole *BRCA1* and upstream region (Gad et al. 2001*a*). As shown in figure 2, two signal patterns were observed in the F32 and F3514 DNA samples: the abnormal allele showed a similar *BRCA1* bar code in both cases, with absence of hybridization of the *NBR2* LR2–4 and of the cosmid probe, which normally allows visualization of both Ψ*BRCA1* and *BRCA1* exons 1 and 2 because of cross-hybridization, on one of its targets. Taken together with the Southern blot results, these data suggested that a deletion comprising the region between Ψ*BRCA1* and *BRCA1* had occurred in F32 and F3514 and had led to the replacement of *BRCA1* exons 1 and 2 by Ψ*BRCA1* exons 1 and 2.

In the same way that we were unable to amplify the region between Ψ*BRCA1* and *BRCA1* by long-range PCR in control individuals, we were unable to amplify the mutant allele with various combinations of forward primers located in Ψ*BRCA1* exon 2 and reverse primers located in *BRCA1* exon 3 by long-range PCR, using a wide range of conditions, although the expected fragment was estimated to be <15 kb.

When sequence information concerning intron 2 of the *BRCA1* pseudogene became available, it gave us the opportunity to design primers that specifically hybridized the pseudogene. Such forward primers were used in conjunction with reverse primers specifically hybridizing the *BRCA1* gene. Primers pintron2.2F and intron2.R2 gave rise to an ~800-bp fragment when they were used with DNA from members of families F32 and F3514 who carried the deletion. No product was obtained when these primers were used with DNA from control individuals. Sequencing of this fragment revealed that a 36,934-bp fragment, comprising *BRCA1* exons 1 and 2 and the region located between Ψ*BRCA1* and *BRCA1*, is deleted in both families. In F32, the breakpoint junctions occurred in two regions of 237 bp within block 11 with 100% similarity (nucleotides 34439–34675 in *BRCA1* and nucleotides 71374–71610 in Ψ*BRCA1*, referring to the RP11-242D8 BAC sequence). The breakpoint junctions in F3514 occurred in 23-bp regions of perfect homology, located 77 bp apart (nucleotides 34,339–34,361 in *BRCA1* and nucleotides 71,274–71,296 in Ψ*BRCA1*). These regions do not contain any repetitive element.

Because of the deletions, carriers in families F32 and F3514 harbor a chimeric gene that consists of Ψ*BRCA1* exons 1A, 1B, and 2 fused to *BRCA1* exons 3–24 (fig. 3). As a consequence, the *BRCA1* promoter is absent from this mutant allele. It is not known yet whether the *BRCA1* pseudogene is transcribed, and therefore we do not know whether the mutant gene might be expressed. Nevertheless, all attempts to identify a mutant transcript were unsuccessful (data not shown). In any case, and although both exons 2 are highly similar, the methionine codon used as the *BRCA1* translation initiation codon is lost in the chimeric gene, because it is changed to an isoleucine codon in Ψ*BRCA1*. Missense mutations changing this methionine codon to a valine or an isoleucine codon have been reported in five families with breast and ovarian cancer (Breast Cancer Information Core). It implies that even if other methionine codons can be used for translation initiation; as has been suggested elsewhere (Liu et al. 2000), the resulting

Table 1

Blocks of Homology between *NBR1*/*NBR2* and Ψ*BRCA1*/*BRCA1*

Block Number	AC060780 Coordinates		Size of Block (bp)	Identity (%)	Gap between Blocks (bp)	
	NBR1 or Ψ*BRCA1*	*NBR2* or *BRCA1*			*NBR1* or Ψ*BRCA1*	*NBR2* or *BRCA1*
1	92098–92594	45329–45863	542	83.9	3,794	0
2	87618–88303	44646–45328	689	91.0	306	0
3	86955–87311	44276–44645	388	81.2	1	316
4	86005–86953	43004–43958	974	89.6	0	1,853
5	84706–86004	39874–41150	1,321	86.4	343	0
6	83959–84362	39472–39873	405	87.7	61	0
7	83161–83897	38695–39471	781	88.9	315	0
8	82187–82845	38031–38684	667	92.1	7,457	0
9	74455–74729	37761–38030	281	76.9	796	443
10	72200–73658	35887–37317	1,473	87.4	4	623
11	70684–72195	33720–35263	1,555	90.3	194	0
12	69048–70489	32241–33719	1,482	91.5	62	0
13	68196–68985	31450–32240	794	91.2	494	0
14	67625–67701	31373–31449	77	87.0	…	…
Total			11,429			

BRCA1 isoform cannot substitute for the tumor-suppressor function of the wild-type protein in breast and ovarian tissues.

Screening for Rearrangements Involving the ΨBRCA1 *in Families with Breast and Ovarian Cancer*

The identification of two different recombination events involving homologous regions located respectively in the *BRCA1* gene and Ψ*BRCA1* leads to the possibility that these regions represent a strong hot spot for recombination. We therefore screened additional candidate families for the presence of such deletions, with two different combinations of forward primers located in *NBR1* or Ψ*BRCA1* and reverse primers located in *BRCA1*. The 13 tested families with breast and ovarian cancer were chosen on the basis of strict criteria and were negative in a mutation screening analysis of the *BRCA1* and *BRCA2* coding sequence and splice sites. The *BRCA1* promoter lies between exons 1A of *NBR2* and *BRCA1*. Therefore, rearrangements involving blocks 1–5 would disrupt the *NBR1* and the *NBR2* genes but would not disrupt any part of the *BRCA1* gene, including its promoter. We thus focused our analysis on screening for rearrangements within blocks 6–14. Because of difficulties, as mentioned above, in amplifying the region between *NBR1* and *BRCA1*, we split the screened interval into two segments: blocks 5–8 and blocks 10–14 (see fig. 1). No other large deletions resulting from recombination between *BRCA1* and its pseudogene were observed.

Discussion

We show here that the first exons of the *BRCA1* gene were replaced, in carriers from two families with breast and ovarian cancer, by those of the *BRCA1* pseudogene, Ψ*BRCA1* (fig. 3). Among the 14 blocks of homology identified between the *NBR1*/Ψ*BRCA1* and the *NBR2*/*BRCA1* regions, the breakpoint junctions of the 37-kb deletions identified in families F32 and F3514 were both found to be located in block 11 (fig. 1), at close but distinct sites within segments of 624 bp that are 98% identical. These segments contain the longest stretches of complete sequence identity found among any of the 14 blocks: 237 and 188 bp, compared with <115 bp. The mutant alleles harbor a chimeric gene that consists of Ψ*BRCA1* exons 1A, 1B, and 2 fused to *BRCA1* exons 3–24 (fig. 3). This chimeric gene lacks the *BRCA1* promoter and the *BRCA1* translation initiation codon. The concomitant absence of the *NBR2* gene in the mutant alleles does not appear to lead to a contiguous-gene-deletion syndrome, because carriers from families F32 and F3514 do not show any remarkable phenotypic trait apart from being predisposed to breast and ovarian cancers (10 cases of breast cancer and 2 cases of ovarian cancer in family F32; 2 cases of breast cancer, 1 case of ovarian cancer, and 1 case of bilateral breast cancer and ovarian cancer in family F3514). It should be noted that nothing is known, as yet, about the potential protein encoded by *NBR2*.

The screening of 13 additional highly selected families with breast and ovarian cancer, negative after intensive searches for *BRCA1* and *BRCA2* mutations, failed to identify any more large deletions resulting from homologous recombination between *BRCA1* and Ψ*BRCA1*. More such families should be tested with the primers we describe, to determine the frequency of these alterations in the *BRCA1* mutation spectrum. Notably, two families with breast and ovarian cancer that were reported in the literature to present an ab-

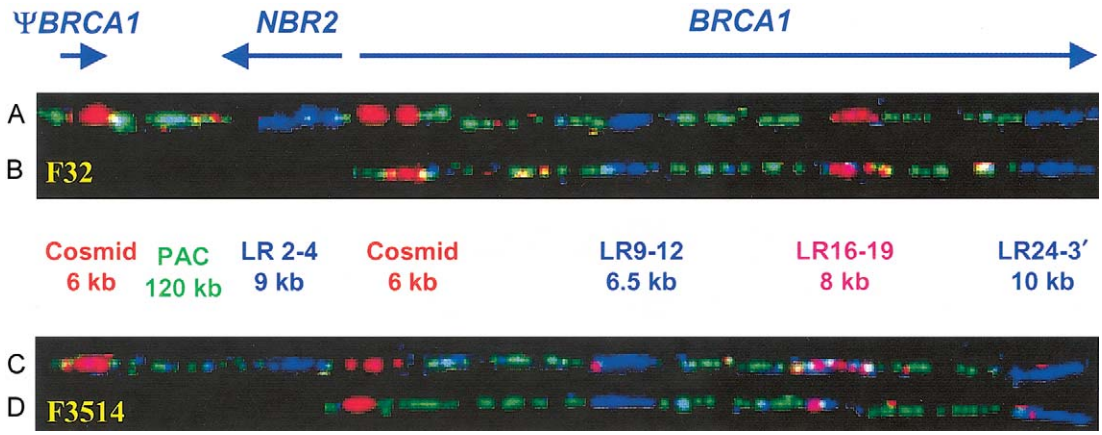

Figure 2 *BRCA1* color bar code of DNA samples from families F32 and F3514. Two allele populations could be visualized under a microscope when combed DNA samples were hybridized with fluorescent probes generated from PAC 103O14 (*green*) and cosmid D06121 (*red*) and from long-range PCR products *NBR2* LR2-4 (*blue*), *BRCA1* LR9-12 (*blue*), *BRCA1* LR16-19 (*deep pink*), and *BRCA1* LR24-3′ (*blue*). The sizes of the probes are indicated. The expected normal bar code for the wild-type allele was Ψ*BRCA1-NBR2-BRCA1* exons 1–2, 9–12, 16–19, and 24, as shown in *A* and *C*, for F32 and F3514, respectively. This pattern was found in 18 full signals (50%) for F32 and 26 full signals (58%) for F3514. An abnormal, shorter signal, without the 5′ part of the normal pattern, as shown in *B* and *D*, for F32 and F3514, respectively, was also found in 18 full signals (50%) for F32 and 19 full signals (42%) for F3514, corresponding to the mutant alleles with the deletion.

normal pattern in Southern blot experiments, with probes covering *BRCA1* exons 1 and 2 (Unger et al. 2000), might carry a recombination event involving Ψ*BRCA1*.

Duplication of genes, gene segments, and repeat-gene clusters generated by the evolution of the mammalian genome provides substrates for homologous recombination that results in DNA rearrangements. When rearrangements alter the genome by causing complete loss or gain of a gene(s) sensitive to a dosage effect or by disrupting the structural integrity of a gene, they are responsible for many diseases called "genomic disorders." These include notably α- and β-thalassemia, hemophilia A, growth hormone deficiency, Charcot-Marie-Tooth disease type 1A (CMT1A), hereditary neuropathy with liability to pressure palsies (HNPP) and Smith-Magenis syndrome (for review, see Lupski 1998). Several common features associated with the mechanisms that lead to DNA rearrangements were apparent when analyzing such diseases: (1) long regions of high similarity appear to be required for recombination, and (2) the greater the distance between repeats, the greater the repeat length required for efficient recombination. This is exemplified by unequal crossing-over events causing α-thalassemia, which involve repeated segments of ~4 kb located 3.7 or 4.2 kb apart within the α-globin locus; causing hemophilia A, which involve 98% identical repeats of 9.6 kb located ~500 kb apart within the factor VIII gene locus; and causing CMT1A and HNPP, which involve repeats (CMT1A-REPs) of 24 kb >98% identical located 1,500 kb apart. DNA sequence analysis of the junctions of the recombinations responsible for genomic disorders showed those crossing-over events took place in shorter regions within the repeats where the sequence identity is the strongest. Alternatively, specific elements that may stimulate recombination, such as MITE (χ- and *mariner*-like transposable element) and a χ-like sequence, were found in or close to the 1.4-kb CMT1-REP hot spot (Reiter et al. 1996) and the 2-kb NF1-REP hot spot (Lopez-Correa et al. 2001), respectively.

In the case of the *BRCA1* locus, the duplicated region is >10 kb and is located <50 kb upstream of *BRCA1*, which makes homologous recombination events likely to occur frequently. However, the requirement of an overall high similarity between the repeated segments is not met, because of numerous interruptions generated by insertions and/or deletions of repetitive sequences (see fig. 3). The fact that the duplicated copy and the original one diverged from one another to such an extent might be related to the unusually high content of repetitive elements in the *BRCA1* gene (Smith et al. 1996). In the case of the disorders mentioned above, genome rearrangements represent a high proportion of mutations. Contrary to this, it seems clear from the data obtained in our panel of ~100 families with breast and ovarian cancer that rearrangements between *BRCA1* and Ψ*BRCA1* are heavily outnumbered by point mutations, small insertions, small deletions, and splice-site mutations. This might be explained by the overall low similarity of the repeats (as mentioned above), associated with the large size of the *BRCA1* coding sequence, which provides more

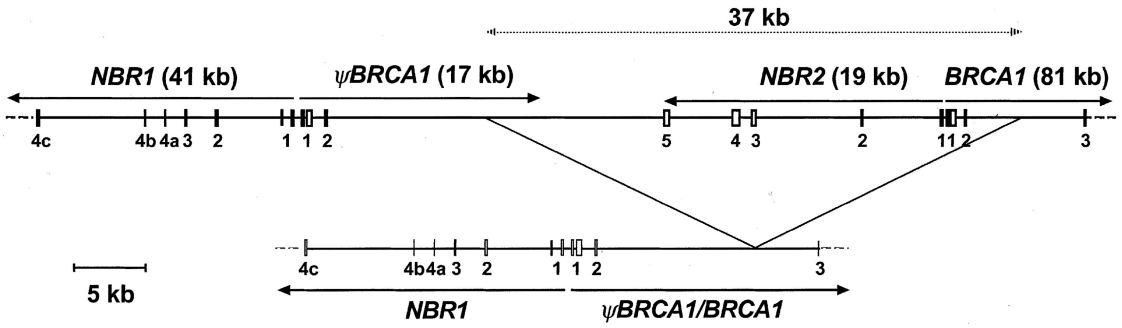

Figure 3 Representation of the 37-kb deletion in F32 and F3514. Drawn-to-scale schema of the normal and mutant alleles at the *NBR1* and *BRCA1* loci, showing the location of the 36,934-bp deletion. The size of the *NBR1,* Ψ*BRCA1, NBR2,* and *BRCA1* genes are given in parentheses in the normal allele schema.

opportunities for mutations during replication. Nevertheless, it is reasonable to expect more families with breast and/or ovarian cancer worldwide to carry large deletions, like the ones described here. Therefore, screening protocols should include PCR assays, such as those described in the present article, to detect this new type of *BRCA1* rearrangement in families with breast and ovarian cancer that have no mutation in *BRCA1/2* coding sequences and splice sites.

Acknowledgments

We thank the family members who collaborated in this study. We also thank H. T. Lynch, for a long-term collaboration in the study of these families, C. Bonnardel and C. Snyder, for their expert assistance, and M. Zody for communicating helpful information concerning the progress of the RP11-242D8 BAC sequencing by the Whitehead Institute/MIT Center for Genome Research. This work was supported by program grants from le Comité Départemental de l'Ain et du Rhône de la Ligue contre le Cancer and by the Institut Curie Programme Incitatif et Coopératif: Génétique et Biologie des Cancers du Sein. N.P., S.G., and L.P.-V. are supported by fellowships from the Ligue contre le Cancer de Haute-Savoie; the Ministère de l'Education National, de la Recherche et de la Technologie; and the Ligue Nationale contre le Cancer, respectively.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

BLAST search, http://www.ncbi.nlm.nih.gov/BLAST/
Breast Cancer Information Core, http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/
GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for BAC sequence [accession number AC060780] and *BRCA1* [accession number L78833])
Institut Curie, http://www.curie.net/genetique
LALIGN, http://www2.igh.cnrs.fr/bin/lalign-guess.cgi
NIX, http://www.hgmp.mrc.ac.uk/Registered/Webapp/nix/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *BRCA1* [MIM 113705])
RepeatMasker, http://ftp.genome.washington.edu/cgi-bin/RepeatMasker
SIM4, http://pbil.univ-lyon1.fr/sim4.html

References

Barker DF, Liu X, Almeida ER (1996) The *BRCA1* and *1A1.3B* promoters are parallel elements of a genomic duplication at 17q21. Genomics 38:215–222

Bensimon A, Simon A, Chiffaudel A, Croquette V, Heslot F, Bensimon D (1994) Alignment and sensitive detection of DNA by a moving interface. Science 265:2096–2098

BRCA1 Exon 13 Duplication Screening Group, The (2000) The exon 13 duplication in the *BRCA1* gene is a founder mutation present in geographically diverse populations. Am J Hum Genet 67:207–212

Brown MA, Jones KA, Nicolai H, Bonjardim M, Black D, McFarlane R, de Jong P, Quirk JP, Lehrach H, Solomon E (1995) Physical mapping, cloning, and identification of genes within a 500-kb region containing *BRCA1*. Proc Natl Acad Sci USA 92:4362–4366

Brown MA, Xu CF, Nicolai H, Griffiths B, Chambers JA, Black D, Solomon E (1996) The *5′* end of the *BRCA1* gene lies within a duplicated region of human chromosome 17q21. Oncogene 12:2507–2513

Gad S, Aurias A, Puget N, Mairal A, Schurra C, Montagna M, Pages S, Caux V, Mazoyer S, Bensimon A, Stoppa-Lyonnet D (2001*a*) Color bar coding the *BRCA1* gene on combed DNA: a useful strategy for detecting large gene rearrangements. Genes Chromosomes Cancer 31:75–84

Gad S, Scheuner MT, Pages-Berhouet S, Caux-Moncoutier V, Bensimon A, Aurias A, Pinto M, Stoppa-Lyonnet D (2001*b*) Identification of a large rearrangement of the *BRCA1* gene using colour bar code on combed DNA in an American breast/ovarian cancer family previously studied by direct sequencing. J Med Genet 38:388–392

Lahti-Domenici J, Rapakko K, Paakkonen K, Allinen M, Nevanlinna H, Kujala M, Huusko P, Winqvist R (2001) Exclusion of large deletions and other rearrangements in *BRCA1*

and *BRCA2* in Finnish breast and ovarian cancer families. Cancer Genet Cytogenet 129:120–123

Liu J, Prolla G, Rostagno A, Chiarle R, Feiner H, Inghirami G (2000) Initiation of translation from a downstream in-frame AUG codon on *BRCA1* can generate the novel isoform protein DeltaBRCA1(17aa). Oncogene 19:2767–2773

Lopez-Correa C, Dorschner M, Brems H, Lazaro C, Clementi M, Upadhyaya M, Dooijes D, Moog U, Kehrer-Sawatzki H, Rutkowski JL, Fryns JP, Marynen P, Stephens K, Legius E (2001) Recombination hotspot in *NF1* microdeletion patients. Hum Mol Genet 10:1387–1392

Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet 14:417–422

Mazoyer S, Dunning AM, Serova O, Dearden J, Puget N, Healey CS, Gayther SA, Mangion J, Stratton MR, Lynch HT, Goldgar DE, Ponder BA, Lenoir GM (1996) A polymorphic stop codon in *BRCA2*. Nat Genet 14:253–254

Michalet X, Ekong R, Fougerousse F, Rousseaux S, Schurra C, Hornigold N, van Slegtenhorst M, Wolfe J, Povey S, Beckmann JS, Bensimon A (1997) Dynamic molecular combing: stretching the whole human genome for high-resolution studies. Science 277:1518–1523

Montagna M, Santacatterina M, Torri A, Menin C, Zullato D, Chieco-Bianchi L, D'Andrea E (1999) Identification of a 3 kb *Alu*-mediated *BRCA1* gene rearrangement in two breast/ovarian cancer families. Oncogene 18:4160–4165

Payne SR, Newman B, King MC (2000) Complex germline rearrangement of *BRCA1* associated with breast and ovarian cancer. Genes Chromosomes Cancer 29:58–62

Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R, Drusedau M, Hogervorst FB, Hageman S, Arts PJ, Ligtenberg MJ, Meijers-Heijboer H, Klijn JG, Vasen HF, Cornelisse CJ, van't Veer LJ, Bakker E, van Ommen GJ, Devilee P (1997) *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. Nat Genet 17:341–345

Puget N, Sinilnikova OM, Stoppa-Lyonnet D, Audoynaud C, Pages S, Lynch HT, Goldgar D, Lenoir GM, Mazoyer S (1999*a*) An *Alu*-mediated 6-kb duplication in the *BRCA1* gene: a new founder mutation. Am J Hum Genet 64:300–302

Puget N, Stoppa-Lyonnet D, Sinilnikova OM, Pages S, Lynch HT, Lenoir GM, Mazoyer S (1999*b*) Screening for germline rearrangements and regulatory mutations in *BRCA1* led to the identification of four new deletions. Cancer Res 59: 455–461

Puget N, Torchard D, Serova-Sinilnikova OM, Lynch HT, Feunteun J, Lenoir GM, Mazoyer S (1997) A 1-kb *Alu*-mediated germ-line deletion removing *BRCA1* exon 17. Cancer Res 57:828–831

Reiter LT, Murakami T, Koeuth T, Pentao L, Muzny DM, Gibbs RA, Lupski JR (1996) A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. Nat Genet 12:288–297

Robinson MD, Chu CE, Turner G, Bishop DT, Taylor GR (2000) Exon deletions and duplications in *BRCA1* detected by semiquantitative PCR. Genet Test 4:49–54

Rohlfs EM, Chung CH, Yang Q, Skrzynia C, Grody WW, Graham ML, Silverman LM (2000*a*) In-frame deletions of *BRCA1* may define critical functional domains. Hum Genet 107:385–390

Rohlfs EM, Puget N, Graham ML, Weber BL, Garber JE, Skrzynia C, Halperin JL, Lenoir GM, Silverman LM, Mazoyer S (2000*b*) An *Alu*-mediated 7.1 kb deletion of *BRCA1* exons 8 and 9 in breast and ovarian cancer families that results in alternative splicing of exon 10. Genes Chromosomes Cancer 28:300–307

Serova OM, Mazoyer S, Puget N, Dubois V, Tonin P, Shugart YY, Goldgar D, Narod SA, Lynch HT, Lenoir GM (1997) Mutations in *BRCA1* and *BRCA2* in breast cancer families: are there more breast cancer–susceptibility genes? Am J Hum Genet 60:486–495

Serova O, Montagna M, Torchard D, Narod SA, Tonin P, Sylla B, Lynch HT, Feunteun J, Lenoir GM (1996) A high incidence of *BRCA1* mutations in 20 breast-ovarian cancer families. Am J Hum Genet 58:42–51

Smith TM, Lee MK, Szabo CI, Jerome N, McEuen M, Taylor M, Hood L, King MC (1996) Complete genomic sequence and analysis of 117 kb of human DNA containing the gene *BRCA1*. Genome Res 6:1029–1049

Swensen J, Hoffman M, Skolnick MH, Neuhausen SL (1997) Identification of a 14 kb deletion involving the promoter region of *BRCA1* in a breast cancer family. Hum Mol Genet 6:1513–1517

Unger MA, Nathanson KL, Calzone K, Antin-Ozerkis D, Shih HA, Martin AM, Lenoir GM, Mazoyer S, Weber BL (2000) Screening for genomic rearrangements in families with breast and ovarian cancer identifies *BRCA1* mutations previously missed by conformation-sensitive gel electrophoresis or sequencing. Am J Hum Genet 67:841–850

Xu CF, Brown MA, Nicolai H, Chambers JA, Griffiths BL, Solomon E (1997) Isolation and characterisation of the NBR2 gene which lies head to head with the human *BRCA1* gene. Hum Mol Genet 6:1057–1062