Germline mutations of the *BRCA1* gene in breast and ovarian cancer families provide evidence for a genotype–phenotype correlation

Simon A. Gayther[1]*, William Warren[2]*, Sylvie Mazoyer[1]*, Paul A. Russell[1], Patricia A. Harrington[1], Mathias Chiano[1,3], Sheila Seal[2], Rifat Hamoudi[2], Elizabeth J. van Rensburg[1,4], Alison M. Dunning[1], Richard Love[5], Gareth Evans[6], Doug Easton[7], David Clayton[3], Michael R. Stratton[2] & Bruce A.J. Ponder[1]

[1]*CRC Human Cancer Genetics Research Group, Box 238, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK*
[2]*Section of Molecular Carcinogenesis, Institute of Cancer Research, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK*
[3]*MRC Biostatistics Unit, Institute of Public Health, University Forvie Site, Robinson Way, Cambridge, CB2 2SR, UK*
[4]*Department of Human Genetics, University of Pretoria, P.O. Box 2034, Pretoria 0001, South Africa*
[5]*Cancer Prevention Programme, 1300 University Avenue, 7C, Madison, Wisconsin 53706, USA*
[6]*Regional Genetics Service, St. Mary's Hospital, Manchester, Whitworth Park, Manchester M13 0JH, UK*
[7]*Department of Community Medicine, Institute of Public Health, University Forvie Site, Robinson Way, Cambridge, CB2 2SR, UK*

**These authors contributed equally to the study.*

Correspondence should be addressed to S.A.G.

Mutations in the *BRCA1* gene, discovered in 1994, are associated with an 80–90% lifetime risk of breast cancer. We have analysed 60 families with a history of breast and/or ovarian cancer for germline mutations in *BRCA1*. Twenty-two different mutations were detected in 32 families (53%), of which 14 are previously unreported. We observed a significant correlation between the location of the mutation in the gene and the ratio of breast to ovarian cancer incidence within each family. Our data suggest a transition in risk such that mutations in the 3' third of the gene are associated with a lower proportion of ovarian cancer. Haplotype analysis supports previous data which suggest some *BRCA1* mutation carriers have common ancestors; however, we have found at least two examples where recurrent mutations appear to have arisen independently.

Germline mutations of the *BRCA1* gene are thought to be responsible for about half of all families with a dominant predisposition to breast cancer, and between 80 and 90% of families in which multiple cases of both breast and ovarian cancer occur[1]. These mutations are highly penetrant, conferring a risk of about 90% of either breast or ovarian cancer by the age of 70 years. The proportion of breast to ovarian cancer appears to vary between families. Epidemiological analysis suggests that the pattern of cancers is best described by a model in which the majority of families are strongly predisposed to breast cancer but have only a moderately increased risk of ovarian cancer, while a minority of families are equally strongly predisposed to both cancers[2]. So far, however, no clear correlation has been reported between the phenotype and the type of mutation or its location within the gene.

BRCA1 was predicted to act as a tumour suppressor gene because allele losses in the *BRCA1* region in tumours from *BRCA1* linked families exclusively affect the wild-type chromosome[3]. Consistent with this, the majority of the 53 distinct germline *BRCA1* mutations reported so far are frameshift or nonsense mutations predicted to lead to a truncated, and therefore possibly inactive, BRCA1 protein[4–11]. The mutations are scattered widely over the coding sequence of the gene, although in a largely North-American series of families, two of the reported mutations have recurred many times[9]. Haplotype analysis suggests that these are founder mutations[7,12].

To examine the spectrum of mutations in a set of families of predominantly British origin, and to determine whether there is any variation in the phenotypic risk associated with particular mutations as suggested by epidemiological studies, we have performed a systematic search for mutations in the coding region of the *BRCA1* gene in 60 families in which there are at least four cases of either breast cancer diagnosed before the age of 60 years, or epithelial ovarian cancer. Our results indicate that there is a connection between the location of a particular *BRCA1* mutation and the chances of developing either breast or ovarian cancer.

Germline mutations in *BRCA1*

BRCA1 mutation analysis was performed on genomic DNA using a combination of heteroduplex analysis (HA) and single strand conformation analysis (SSCA) to screen for DNA variants prior to direct sequence analysis to characterise the precise alteration. Individuals in whom these initial screening methods failed to detect any clearly causative DNA sequence alteration were subsequently re-analysed throughout the entire *BRCA1* coding region by sequence analysis. We found alterations consistent with causative mutations in 32 families (53%). In 21 families it was possible to analyse more than one affected individual, and in each case the germline alteration segregated with the disease.

There were 22 different mutations, 14 of which are described for the first time (Table 1). These include 11 frameshifts resulting from insertion or deletion of between 1 and 40 bp, and 5 nonsense mutations, all predicted to result in a truncated BRCA1 protein.

Five mutations occurred at the boundaries between exons and introns and are predicted to affect splicing. In two of these, cDNA was available from an affected individual, enabling sequence analysis across the relevant exon boundaries to confirm the effect of the mutations.

Table 1 Germline mutations in *BRCA1* in families with inherited breast and ovarian cancer

Family	Reported cases of cancer[a]		Description of mutation			
	Ovarian	Breast ≤ 60yrs¶	Exon	Codon	Alteration[b]	Predicted effect of alteration
B019	0	5	-	-	Inferred regulatory	Loss of transcript
OV010[c]	4 (1)	3 (2)	2	23	185delAG	Truncation codon 39
OV022	3	1	intron 5	-	331+2T>C	Splice aberration
OV047	4 (1)	2	8	169	Gln169ter	Truncation codon 169
B128	4 (1)	4 (1)	11	266	916delTT	Truncation codon 285
OV008	3	1	11	367	1220insC	Truncation codon 368
OV002	6 (1)	3 (1)	11	392	1294del40	Truncation codon 397
OV150	4	2	11	392	1294del40	Truncation codon 397
B138	1	5 (2)	11	392	1294del40	Truncation codon 397
B925	5 (4)	4 (3)	11	639	Leu639ter	Truncation codon 639
OV001	3	1	11	654	2073delA	Truncation codon 700
OV029	6	2	11	1001	3121delA	Truncation codon 1023
B208	1	7 (5)	11	1080	Leu1080ter	Truncation codon 1080
OV012	6	2 (1)	11	1111	3452del4	Truncation codon 1115
B146	1	3	11	1111	3452del4	Truncation codon 1115
OV007	4 (1)	0	11	1252	3875del4	Truncation codon 1262
OV025[c]	5	2	11	1289	3986delAA	Truncation codon 1293
OV 019	4	1	11	1355	4184del4	Truncation codon 1364
OV044[c]	10 (7)	(5)	11	1355	4184del4	Truncation codon 1364
B227[c]	2 (1)	5 (4)	11	1355	4184del4	Truncation codon 1364
B229[c]	2 (1)	2	11	1355	4184del4	Truncation codon 1364
OV16[c]	6	0	12	1395	4304G>A	Splice aberration
B202	1	5 (1)	12	1395	4304G>A	Splice aberration
B035	0	4 (1)	13	1443	Arg1443ter	Truncation codon1443
B088	(1)	10 (1)	13	1443	Arg1443ter	Truncation codon 1443
B155	2	5 (2)	13	1443	Arg1443ter	Truncation codon 1443
OV003	2	2	intron 13	-	4476+6T>C	Splice aberration
B217	0	4 (2)	intron 18	-	5271+1G>T	Splice aberration
B920	0	5	19	1727	Lys1727ter	Truncation codon 1727
B176	2	6	intron 19	-	5312+2del T	Splice aberration
B082[d]	1	10	20	1755	5382insC	Truncation codon 1829
B218	3 (1)	8 (2)	24	1837	5629delG	Truncation codon 1843

[a]The number of cases unconfirmed by either death certificate or pathology report is given in brackets; [b]Nomenclature according to Beaudet & Tsui[21]; [c]Family and mutation previously reported in Shattuk-Eidens *et al.*[9]; [d]Family and mutation are identical to the family and mutation reported by Black and Boyd in Shattuck-Eidens *et al.*[9]

The single base pair substitution 4479 + 6T > C, in family OV003, results in skipping of exon 13. The aberrant splicing creates a frameshift which results in truncation of the predicted protein. In family OV161, one of two families with a nucleotide substitution involving the last base of exon 12, analysis of cDNA from an affected individual revealed absence of the mutant *BRCA1* transcript. This individual was heterozygous for two polymorphisms in exons 11 and 16. Analysis of these polymorphisms in cDNA revealed retention of heterozygosity in exon 11, 5' of the mutation, but homozygosity for the exon 16 polymorphism 3' of the mutation. This indicates that the disease-associated alteration has caused a splicing defect resulting in skipping of *BRCA1* sequences beginning with exon 12 and possibly involving the remaining 3' portion of the gene. cDNA was not available from affected individuals of family B202 which carries the same mutation to confirm these observations; however, the mutation in this family showed complete segregation with the disease. cDNA was not available for affected individuals from families with the remaining three splice site alterations for a comparable analysis. However, all three mutations affect highly conserved splice donor consensus sequences within introns 5, 18 and 19 respectively. These mutations would be expected to lead to frameshift and premature truncation, as a result either of exon skipping or the activation of a cryptic splice site.

There are no obvious regions of mutation clustering within the gene: 17 of the 32 mutations (53%) are located within exon 11, which is consistent with the contribution of this exon to the entire *BRCA1* coding sequence. Five alterations were detected more than once: 4304 G > A and 3450del4 detected on two separate occasions, 1294del40 and Arg1443ter both detected three times and 4184del4 which was detected in four families. The two mutations most frequently reported in previous studies — 185delAG and 5382insC (refs 9, 11, 13) — were each detected only once in our study (see haplotype analysis).

Allele-specific expression assay

In 8 of the 28 families in which disease-associated alterations were not identified, and in which an affected individual was heterozygous for three tightly linked coding polymorphisms in exons 11 and 16, we analysed cDNA to search for unequal expression from the two *BRCA1* alleles, which might suggest a regulatory mutation. For each polymorphism, only one allele was detectable in the cDNA in one individual from family B019, indicating absence of transcript from the other *BRCA1* allele. cDNA was unavailable from other members of this family to confirm segregation of this alteration with the disease. No alterations suggesting absence of transcript were observed in the remaining seven informative families.

Haplotype analysis

Of the 32 mutations reported here, 9 distinct alterations in 16 families are recurrent either in this or in

Table 2 Haplotype analysis in families with recurrent *BRCA1* mutations

Mutation	Family[a]	Haplotype[b]			
		D17S855	*D17S1322*	*D17S1323*	*D17S1327*
185delAG	Canadian	**G**	**C**	**C**	**E**
	OV010	G	C	C	E
1294del40	Canadian	**F**	**E**	**F**	**M**
	B138	F	E	F	M
	OV002	F	E	F	M
	OV150	F/H	E	F	M
3121delA	Canadian	**E**	**E**	**F**	**M**
	OV029	F	E	F	M
3449del4	B146	E	D/E	F	M
	OV012	D	E	F	M
4184del4	Canadian	**A**	**F**	**F**	**M**
	B227	E	E	F	M
	B229	A	F	F	M
	OV019	A/D	E/F	F	M
	OV044	H	A	C	F
4304G>A	B202	E	E	F	M
	OV161	D/G	E	B/F	C/M
Arg1443ter	B035	G	D	B	F
	B088	D/E	E	F	M
	B155	E/H	E	F	M
5382insC	Canadian	**D**	**E**	**F**	**O**
	B082	D	E	F	N

[a]Canadian families refer to a similar study performed by Simard *et al.*[7] on a series of families ascertained in Canada; [b]Haplotype data derived for mutations and families from the current study (plain type) in comparison to haplotype data derived for the same mutation in the study of Simard *et al.* (bold type).

other studies. To assess whether these alterations occurred *de novo* or are founder mutations, we constructed haplotypes using four microsatellite markers located within or adjacent to *BRCA1*, which had been used in a comparable analysis of Canadian families[7,12] (Table 2). The mutation 185delAG, identified on four occasions by Simard *et al.*[7] and recently shown to be present in 1 in 100 Ashkenazi Jews[13], was detected only once but was on the same haplotype as the four families from the Canadian study. This haplotype was not found in any of the other 59 families. The haplotype associated with the mutation 5382insC detected in our family B082 and four times by Simard *et al.*, differs only in the allele associated with the marker *D17S1327*. Since both the alleles are extremely rare and differ by only 2 bp, it is likely that the mutation in B082 has the same ancestral origin as in the four Canadian families, and that the allele at *D17S1327* has undergone a mutational event since the origin of the *BRCA1* mutation.

Our three families with the mutation 1294del40 share the same haplotype as a Canadian family with the same mutation[7]. However, the alleles at the four markers are common, and so we cannot be completely certain that this alteration derives from a common ancestor. The mutation 4184del4 occurs on three distinct haplotypes, although two families are consistent with sharing the same rare disease haplotype as a Canadian family with the same mutation. Similarly, the mutation Arg1443ter, detected three times in this study, was associated with at least two different haplotypes. The three other recurrent mutations may also have distinct haplotypes. However, the haplotypes associated with these mutations involve very frequent alleles and we are therefore unable to say whether these mutations occurred *de novo* or derive from a common ancestor, with subsequent allele slippage and alteration of the original allele sizes.

Genotype–phenotype correlation

Because the families in which we identified mutations had widely differing proportions of breast and ovarian cancer, we were able to assess whether there was a correlation between the location of the mutation in the gene and the phenotype in respect of these cancers. We restricted the analysis to individuals in which there was confirmation of the diagnosis of breast or ovarian cancer from pathology reports or death certificates. In the 32 families in which mutations were identified, there were 86 confirmed cases of breast cancer under the age of 60 years and 76 cases of ovarian cancer. The locations of the *BRCA1* mutations, together with the proportion of breast and ovarian cancer within each family, are illustrated in Fig. 1.

A statistical test for trend of the ratio of breast to ovarian cancer cases with the location of the mutation in the gene (codon number) yielded a significant result ($P = 0.01$). A test of the null hypothesis (no genotype–phenotype correlation) against a 'change point' model in which the ratio of breast to ovarian cases differs when the mutations lie in different halves of the gene, but remains constant for mutations in the same half, was even more strongly significant ($P < 0.001$). Table 3 summarises a logistic regression analysis that attempts to discriminate between steady-trend and change point models. While introduction of a change-point into the trend significantly improves the fit of the regression model, the addition of a trend into the change-point has a negligible effect. These data are thus indicative of a difference in phenotypic expression for mutations in different halves of the gene rather than a more gradual change. The best estimate for the location of the change point occurs in exon 13, between codons 1435 and 1443, with a confidence region between codons 634 and 1731.

Discussion

Our study reports the identification of germline *BRCA1* mutations in 32 families with a history of breast and/or ovarian cancer, and correlates the location of the mutation along the length of the gene with high or low prevalence of ovarian cancer within families. The spectrum of mutations in our series is similar to that observed before now[9]. However, we found no missense mutations and a greater proportion of mutations predicted to affect splicing (19% compared with 3%). Furthermore, two putative hotspot mutations, which constitute 36% of all mutations reported previously, were detected only once apiece (6% of mutations). This may reflect differences in the geographical origin of our families compared to the predominantly North American origin of many previously published families[9,11]. In particular, the mutation 185delAG is now known to be a founder mutation common to Ashkenazi Jews, and is therefore likely to be relatively uncommon in the United Kingdom[12,13].

Our haplotyping results support evidence from previous studies that *BRCA1* families which carry mutations 185delAG and 5382insC have common ancestors[7,12]. However, it is also clear that other recurrent mutations, for example 4184del4 and Arg1443ter,

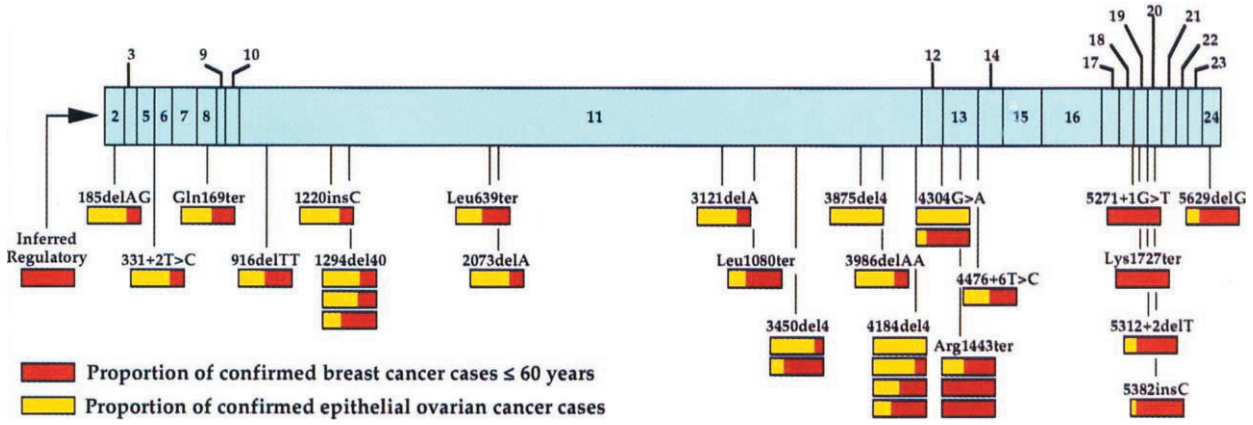

Fig. 1 The location of germline mutations in *BRCA1* in 32 families. Correlation with the phenotype in each family based on the ratio of confirmed cases of breast cancer under the age of 60 years to confirmed cases of ovarian cancer.

may sometimes have arisen independently. Although the 4184del4 mutation has occurred on three distinct haplotypes, there are no tandem or inverse repeat sequences in the region of the mutation which could be predicted to lead to a high frequency of deletion as a result of misalignment during DNA replication. By contrast, the majority of frameshift mutations we identified occurred at short repeat sequences, which would possibly be associated with slippage during replication. Two of the five nonsense mutations, including Arg1443ter, were alterations from cytosine to thymidine at CpG dinucleotides, perhaps caused by deamination of 5-methylcytosine[14].

Based on previously derived estimates for the proportion of breast and breast-ovarian cancer families due to *BRCA1* mutations, and the available linkage evidence in these families, the expected number of families due to *BRCA1* in this dataset was 49 (95% confidence interval [CI] 43.2–54.0%). Thus our mutation screen has only identified 66% (95% CI 50–84%) of the mutations estimated to be present. The clinical criteria we used to select families were the same as those used in previous Breast Cancer Linkage Consortium (BCLC) studies and on which the expected numbers are based, so differences due to selection of the families are unlikely[1]. Moreover, we could not identify coding abnormalities in some families with convincing evidence for linkage; of the four families with Lod scores of 0.9 or greater at *BRCA1* (B167, 1.40: OV025 1.25; OV034, 0.93; OV044, 0.93), only two had identified abnormalities. Subsequent haplotyping using markers tightly linked to the *BRCA2* locus on chromosome 13q12 indicated that of the two families in which no *BRCA1* mutation was found, OV034 was also consistent with linkage to *BRCA2* (Lod 0.46) whereas B167 was clearly inconsistent (Lod −1.90). Other possibilities are that coding mutations are present in these families but were missed, although the techniques we used for mutation analysis (an initial screen by SSCA/HA and re-analysis by direct sequencing) would be expected to be highly sensitive. Many of the families where no *BRCA1* mutation was identified could share a common haplotype indicating they may share the same *BRCA1* mutation, which we failed to detect. However, this haplotype is also frequently observed in the normal population. Mutant *BRCA1* alleles bearing duplications, translocations or large deletions are unlikely to have been detected using our PCR based assays, and we did not examine directly the regulatory regions of the gene. The allele specific expression assay indicates that regulatory mutations do occur, although our data and previous reports[4,9] suggest they are not very frequent. Finally, we were only able to analyse a single individual in some families, and it is possible that some of these may by chance have been phenocopies.

Table 3 A summary of the regression analysis

Model	Deviance	Degrees of freedom
No variation	65.59	26
Linear trend	50.15	25
Change-point	23.70	25
Linear trend and change point	22.80	24

Reduction of the deviance when the "change-point" is introduced into the linear trend model is 27.35 (50.15–22.80) on 1 df ($P < 0.001$) but only 0.90 (23.7–22.8) on 1 df (not significant) if the linear trend model is added to the change point.

Breast–ovarian cancer correlation

Our most striking finding is the correlation between the position of the mutation within the gene, and the ratio of breast to ovarian cancer incidence in the family. A suggestion of a similar correlation has been seen before, but it was below the level of significance[9]. The higher level of statistical significance here could be due to the higher proportion of ovarian cancer cases among our families, or perhaps the fact that our analysis was carried out on a series of families, in all of which mutation analysis of the whole gene had been completed. We cannot tell if this difference in ratio is due to a difference in penetrance of the mutation for breast cancer, ovarian cancer or both. Larger systematic studies will be needed to resolve this.

Several models can be invoked to explain our findings. For example, if the mutant BRCA1 peptides are present in the cell and retain partial wild-type BRCA1 function, differences in this residual function may be determined by the presence or absence of one or more specific domains in the truncated peptide. Under this model, one could postulate one or more domains N-terminal of the supposed change point,

the presence of which could supply important BRCA1 functions in ovarian epithelium but not in breast epithelium. Alternatively, if the mutant BRCA1 peptides are present in the cell but do not retain any wild-type BRCA1 function, a dominant–negative model would be necessary to explain the effects of different mutations. Variations and combinations of these models are possible; for example, a model in which truncated proteins resulting from mutations after the change point have partial wild-type activity and are protective against ovarian cancer, whereas mutations prior to the change point have no wild-type activity, but are able to act in a dominant-negative manner. This model is easier to reconcile with the observation in this and one other report[4] of inferred regulatory mutations which in each case are associated only with breast cancer in the family. A different possibility might be alternative splicing in the 3' end of the gene, such that mutations 3' of the supposed change point could be wholly or partially 'rescued' in ovarian but not breast epithelium by the presence of alternative transcripts which are not affected by the mutation. These models suggest experimental approaches, for example a search for 3' splicing differences in *BRCA1* between breast and ovarian epithelium, and a search for a functional domain, possibly involved in protein–protein interactions, in the region of the gene 5' to the apparent change point.

Methods

Patient material. 70 families were identified from our records which contained at least four cases of either female breast cancer diagnosed before the age of 60 years or epithelial ovarian cancer diagnosed at any age, for which DNA samples were available from at least one affected individual. Borderline ovarian cancer was not included. We excluded from our mutation screen three families which contained one or more cases of male breast cancer, since the majority of such families are likely to be due to *BRCA2* (ref. 15) (D. Ford & D.E., pers. comm.) and a further seven families with female breast cancer for which there was evidence against linkage to *BRCA1*, leaving 60 families to be analysed. 13 families contained breast cancer cases only, a single case of ovarian cancer was reported in 11 families and 4 families were site specific ovarian cancer. Where DNA was available from several affected individuals, analysis was initially performed on the youngest affected individual and positive results confirmed on other affected individuals. To calculate the expected proportion of *BRCA1* families in our dataset, we used previous estimates of the proportions of breast and breast-ovarian cancer families due to *BRCA1* derived from the Breast Cancer Linkage Consortium Studies[16] (D. Ford & D.E., pers. comm.) namely 92% (95% CI 76–100%) for families with ≥ 2 ovarian cancer cases, 81% (95% CI 58–98 %) for families with 1 ovarian cancer case and 27% (95% CI 14–53 %) for families with breast cancer only. Families with ovarian cancer only were given the same prior probability as for breast–ovarian families, since the proportion of linked families appears to be similar[17]. Multipoint Lod scores between markers flanking *BRCA1* and the disease were computed as described[16]. Posterior probabilities of linkage for each family were computed in the usual way from the prior probabilities and Lod scores, and then summed to give the estimated proportion of linked families. Likelihood based 95% confidence regions for the expected number of families with mutations, and the proportion of mutations detected, were derived based on values of the log-likelihood with 0.98 of the maximum value.

Mutation analyses of *BRCA1*. Exon 11 was analysed using a selective restriction enzyme based, non-radioactive heteroduplex analysis (RE-HA) specifically designed for the rapid screening of frameshift alterations. The remaining 21 coding exons were analysed using a combined, non-radioactive SSCA/HA. Sequence alterations causing single and double stranded DNA variation were identified using the Applied Biosystems fluorescent based semi-automated sequencing system. Sequence information for oligonucleotide primers used for PCR amplification of genomic DNA and cDNA prior to each type of analysis are available on request (e-mail-sg200@cus.cam.ac.uk). Other than the primers originally described by Miki *et al.*[4], new primers were designed and used to amplify exons 6, 7, 8, 11, 15, 16 and 23. For exon 11, 18 overlapping pairs of primers were designed to use in various combinations for SSCP, heteroduplex and sequence analyses. Nested primers were designed to use for direct sequence analysis from genomic DNA for exons 2, 3, 5–12, 14, 16, 17, 19, 20 and 22.

Restriction enzyme heteroduplex analysis. Exon 11 was amplified from genomic DNA in four fragments, each approximately 1 kb long. 50 µl PCR products were purified directly using the Wizard™ PCR prep system (Promega). In distinct reactions, each fragment was subsequently digested with two restriction enzymes designed to cleave the specific fragment once only. Restriction enzymes and digestion product sizes were as follows: Fragment 11:1-*Eco*RI to produce fragments of 408 and 590 bp and *Bgl*II to produce fragments of 521 and 477 bp; Fragment 11:2-*Bgl*I (594 and 270 bp) and *Hha*I (483 and 381 bp); Fragment 11:3-*Pst*I (644 and 391 bp) and *Hha*I (462 and 573 bp); Fragment 11:4-*Nsi*I (363 and 756 bp) and *Bst*EII (749 and 340 bp). Digestion products were denatured at 95 °C for 10 min and cooled to 37 °C over 2 h to induce heteroduplex formation. DNA fragments were subsequently electrophoresed through 20 cm × 20 cm × 0.1 cm, non-denaturing, 1 × MDE gels (J.T.Baker), using the Protean II™ vertical slab gel apparatus (Bio-Rad). The standard electrophoretic conditions used were 250 V for 12 h, at a constant 12 °C. DNA detection was by silver staining using the following method: 10 min incubation in a solution of 10% ethanol/0.5% acetic acid to fix the gels; 10 min incubation with a solution of 0.1% silver nitrate to stain DNA; two brief rinses with distilled water followed by a 20 min incubation in a solution of 1.5% sodium hydroxide/0.01% sodium borohydride/0.015% formaldehyde to develop the stain; 10 min incubation with 0.75% sodium carbonate to fix stain; 10 min incubation with 10% glycerol to sensitize gels prior to drying at 80 °C for 2 h.

SSCA/heteroduplex analysis. Coding exons of *BRCA1* were amplified from genomic DNA in 30 µl volumes. Gel electrophoresis was performed as for RE-HA except that products were electrophoresed through a 0.6 x MDE™ matrix and gels were run between 150–200 V. DNA detection and drying of the gels was also performed as described for RE-HA.

Sequence analysis. PCR products were prepared in 100 µl volumes. One oligonucleotide of each primer pair used for the amplification was biotinylated and HPLC purified. DNA fragments were immobilised onto streptavidin coated magnetic beads (Dynal) and denatured to produce single stranded template prior to fluorescent sequencing using an Applied Biosystems model 373A semi-automated sequencing system. The dideoxy termination method was used with the PRISM™ single stranded sequenase dye terminator DNA sequencing kit (Applied Biosystems). For the majority of *BRCA1* exons, a nested oligonucleotide primer was designed to perform the sequencing reaction. Where a suitable nested primer was not available PCR products were gel purified using the Wizard™ PCR prep system (Promega) prior to sequencing with the relevant primer used in the PCR reaction.

Allele-specific expression assay. Oligonucleotide primers were designed to amplify genomic DNA and cDNA for three regions

within exons 11 and 16 which encompass three previously characterised coding polymorphisms (C to T substitution at nt 2430, an A to G substitution at nt 3667 and an A to G substitution at nt 4956). cDNA as well as genomic DNA was available from a single individual in 14 of the 28 families in which a *BRCA1* mutation was not identified. Initially, genomic DNA fragments were subjected to SSCP/heteroduplex analysis to determine individuals heterozygous for the three polymorphisms. Subsequently, sequence analysis was performed from genomic DNA and cDNA for heterozygous individuals across the three polymorphic sites (see above) which enabled quantification of any reduction in copy number of each allelic transcript to be performed, in addition to characterising complete absence of transcript.

Haplotype analysis. Haplotype analysis was performed by typing family members with four microsatellite polymorphisms: *D17S855* (ref. 18), *D17S1322, D17S1323* (ref. 19), all located in *BRCA1*, and *D17S1327* (ref. 20), located 100 kb distally. Allele sizes were determined by direct comparison with DNA samples with alleles of known sizes provided by D. Goldgar and are comparable with those from the haplotype study of Canadian families[7]. In some cases, phase-known haplotypes could not be scored because only one individual was available or because the individuals analysed shared both haplotypes. *BRCA2* haplotype analysis was performed using *D13S260, D13S267, D13S219* and *D13S263* as described[15].

Statistical methods. We tested the association between the location of the mutations and disease phenotype using a permutation argument. The distribution of test statistics were evaluated by generating a large series of random permutations of the observed mutations between the families studied. The test we used for a general (trend) relationship was the standard chi-squared test for trend in the ratio of breast:ovary cancers along the gene. However, rather than assuming a chi-squared (one degree of freedom) distribution for this statistic (which assumes binomial variability of the breast:ovary split for any mutation), we evaluated its distribution in 50,000 random permutations between families. For the purpose of the exercise, we took the most conservative strategy and considered families B138, OV002 and OV150, which have the same mutation and share the same haplotype to be a single family with 10 ovarian cancers and 7 breast cancer cases. Similarly, families B229 and OV019 were considered as one family with 5 cases of ovarian cancer and 3 cases of breast cancer, and families B088 and B155 a single family with 2 cases of ovarian and 12 cases of breast cancer. Family B019 was excluded from the analysis as the mutation is an inferred regulatory mutation and so has no precise localisation. A similar approach was used for testing the hypothesis of no mutation/phenotype correlation against the change point alternatives. A chi-squared test can be computed for each 2*2 contingency table formed by counting breast and ovarian cancer cases with mutations before and after specified change-points. In the observed data, the maximum chi-squared value was obtained with a change-point on the exon 12–13 boundary. Again, the statistical significance of this value was assessed not by comparison with chi-squared tables but by generating random permutations of mutations between families. Since the exon 12–13 break point was not chosen *a priori*, the change point was re-chosen for each random permutation so as to maximise the chi-squared value. In this way the significance level computed allows for post-hoc choice of change point. Using a model for a change point between exons 12 and 13, the assumption of a binomial distribution of cases between breast and ovary was quite consistent with the data (Table 3). This assumption was therefore adopted in the remaining analyses: (1) to compare trend and change point models. (2) to set likelihood-based confidence intervals for the location of the change point in the change point model.

Acknowledgements

We thank C. Pye, J. Dearden, M. Ponder, D. Averill, D. Ford, F.Leonard, R. DiCioccio and S. Piver for help with family material, and the families for their co-operation. This work was supported by programme grants from the Cancer Research Campaign, and by Breakthrough Breast Cancer and the Jean Rook Fund. B.A.J.P. is a Gibb Fellow of the CRC; EJvR is the holder of the Lady Cade Memorial fellowship funded jointly by the CRC and the National Cancer Association of South Africa. S.M. is an EC Fellow.

Received 31 July; accepted 6 October 1995.

1. Easton, D.F. Bishop, D.T., Ford, D., Cockford, G.P. and the Breast Cancer Linkage Consortium. Genetic linkage analysis in familial breast and ovarian cancer. *Am. J. hum. Genet.* **52**, 718–722 (1993).
2. Easton, D.F., Ford, D., Bishop, D.T. and the Breast Cancer Linkage Consortium. Breast and ovarian cancer incidence in *BRCA1*-mutation carriers. *Am. J. hum. Genet.* **56**, 265–271 (1995).
3. Smith, S.A., Easton, D.F., Evans, D.G.R. & Ponder, B.A.J. Allele losses in the region 17q12–21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nature Genet.* **2**, 128–131 (1992).
4. Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* **266**, 66–71 (1994).
5. Futreal, P.A. *et al. BRCA1* mutations in primary breast and ovarian carcinomas. *Science* **266**, 120–122 (1994).
6. Castilla, L. *et al.* Mutations in the *BRCA1* in families with early onset breast and ovarian cancer. *Nature Genet.* **8**, 387–391 (1994).
7. Simard, J. *et al.* Common origin of *BRCA1* mutations in Canadian breast and ovarian cancer families. *Nature Genet.* **8**, 392–398 (1994).
8. Friedman, L.S. *et al.* Confirmation of *BRCA1* by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nature Genet.* **8**, 399–404 (1994).
9. Shattuck-Eidens, D. *et al.* A collaborative survey of 80 mutations in the *BRCA1* breast and ovarian cancer susceptibility gene. *J. Am. med. Assoc.* **273**, 535–541 (1995).
10. Hogervorst, F.B.L. *et al.* Rapid detection of *BRCA1* mutations by the protein truncation test. *Nature Genet.* **10**, 208–212 (1995).
11. Struewing, J.P. *et al.* Detection of eight *BRCA1* mutations in 10 breast/ovarian cancer families, including 1 family with male breast cancer. *Am. J. hum. Genet.* **57**, 1–7 (1995).
12. Tonin, P. *et al. BRCA1* mutations in Ashkenazi Jewish women. *Am. J. hum. Genet.* **57**, 189 (1995).
13. Streuwing, J.P. *et al.* The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nature Genet.* **11**, 198–200 (1995)
14. Shen, J.C., Rideout, W.M. III & Jones, P.A. High frequency mutagenesis by a DNA methyltransferase. *Cell* **71**, 1073–1080 (1992).
15. Wooster, R. *et al.* Localisation of a breast cancer susceptibility gene, *BRCA2*, to chromosome 13q12–13. *Science* **265**, 2088–2090 (1994).
16. Narod, S. *et al.* An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. *Am. J. hum. Genet.* **56**, 254–264 (1995).
17. Steichen-Gersdorf, E. *et al.* Familial site-specific ovarian cancer is linked to *BRCA1* on chromosome 17q12–21. *Am. J. hum. Genet.* **55**, 870–875 (1994).\
18. Weissenbach, J. *et al.* A second generation linkage map of the human genome. *Nature* **359**, 794–801 (1992).
19. Neuhausen, S.L. *et al.* A P1-based physical map of the *BRCA1* region from *D17S766* to *D17S78*. *Hum. molec. Genet.* **3**, 1919–1926 (1995).
20. Goldgar, D.E. *et al.* A large kindred with 17q-linked breast and ovarian cancer: genetic, phenotypic and genealogic analysis. *J. natnl. Cancer Inst.* **86**, 200–209 (1994).
21. Beaudet, A.L. & Tsui, L-C. A suggested nomenclature for designating mutations. *Hum. Mut.* **2**, 245–248 (1989).