A highly accurate, low cost test for BRCA1 mutations

Nathalie J van Orsouw, Rahul K Dhanda, Youssef Elhaji, Steven A Narod, Frederick P Li, Charis Eng, Jan Vijg

Abstract
The hereditary breast and ovarian cancer syndrome is associated with a high frequency of BRCA1 mutations. However, the widespread use of BRCA1 testing has been limited to date by three principal concerns: the fear of loss of health and life insurance, the uncertain clinical value of a positive test result, and the current lack of an inexpensive and sensitive screening test for BRCA1 mutations. We have developed an inexpensive system for gene mutational scanning, based on a combination of extensive multiplex PCR amplification and two dimensional electrophoresis. The efficiency of this system, as a screening test for BRCA1 mutations, was evaluated in a panel of 60 samples from high risk women, 14 of which contained a previously identified mutation in BRCA1. All 14 mutations were identified, as well as an additional five that had previously escaped detection. In addition to the 19 mutations, a total of 15 different polymorphic variants were scored, most of which were recurring. All were confirmed by nucleotide sequencing. The cost of screening per sample was calculated to be approximately US$70 for the manual technique used in this study, and may be reduced to approximately US$10 with the introduction of commercially available PCR robotics and fluorescent imaging. Implementation of this method of mutation screening in the research and clinical setting should permit rapid accrual of quantitative data on genotype-phenotype associations for the evaluation of diagnostic testing.

Keywords: genetic testing; two dimensional gene scanning (TDGS)

Breast cancer is a common disease with a life time risk in women of 10%. The risk increases substantially when one or more close relatives are affected. The first breast cancer susceptibility gene, BRCA1, was cloned in 1994.1 To date, well over 300 distinct “high penetrance” mutations in this gene have been described by the Breast Cancer Information Core.2 These confer a breast cancer risk of up to 90% by the age of 70.3 A second gene involved in hereditary breast cancer syndrome, BRCA2, was identified several months later,4 and at least one more gene remains to be found.5 Over 80% of families with two or more cases of premenopausal breast cancer and two or more cases of ovarian cancer are believed to carry a BRCA1 or BRCA2 mutation.6 In addition, other genes may play a role as risk modifiers.7

Comprehensive evaluation of breast cancer genes could be used to provide women with information concerning the risk of developing cancer. Much recent work has been directed towards screening women from breast cancer families and other high risk populations to estimate the contribution of BRCA1 mutations to the overall incidence of breast and ovarian cancers. On the basis of the results of these studies, attempts have been made to estimate the predictive value of a BRCA1 mutation, that is, whether different mutations have a different risk.8 Information is now emerging that knowledge of BRCA1 status is useful in preventive strategies. For example, oral contraceptive use reduces the risk of ovarian cancer by 50% in BRCA1 carriers.9

One of the central themes in cancer risk assessment is the efficiency and accuracy of genetic screening methods. Currently, the ability to conduct large scale, population based studies is constrained by the lack of an accurate and inexpensive method for mutation detection. Nucleotide sequence based screening, which is currently accepted as the gold standard, is offered at US$2400 for screening both BRCA1 and BRCA2 by Myriad Genetics (Salt Lake City, UT).

Here, we report on the evaluation of a practical test for BRCA1 mutation detection, which is based on a combination of extensive multiplex PCR amplification, followed by two dimensional electrophoresis (TDGS, two dimensional gene scanning).10 11 The results of this study indicate a sensitivity for detecting mutations equal to that of sequencing, but at a much lower cost.

Methods
DNA SAMPLES
DNA samples from patients with known mutations in BRCA1 were obtained from 60 patients with cancer and unaffected subjects. Mutations had originally been detected for 14 patients by sequence analysis of selected exons (exons 2 and 20) and by the protein truncation test of exon 11. No previous knowledge of mutation status of the panel was known when TDGS was applied.

PCR AMPLIFICATION
Multiplex PCR amplification was carried out in two steps. A 7-plex long distance PCR preamplification from genomic DNA provided the template for the amplification of the 37 genes.6

Institute for Drug Development, Cancer Therapy and Research Center, 8122 Datapoint Drive, Suite 700, San Antonio, TX 78229, USA
N J van Orsouw
J Vijg
Mosaic Technologies Inc, 1106 Commonwealth Avenue, Boston, MA 02215, USA
R K Dhanda
Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario, Canada MSG 2M9
Y Elhaji
S A Narod
Program in Population Sciences, Dana-Farber Cancer Institute and Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA
F P Li
Clinical Cancer Genetics and Human Cancer Genetics Programs, Ohio State University Comprehensive Cancer Center, 420 W 12th Avenue, 690c
MRF, Columbus, OH 43210, USA
C Eng
Correspondence to:
Dr van Oorsouw.
nvorsouw@saci.org

 Revised version received 2 June 1999
Accepted for publication 21 June 1999
DGGE optimised target fragments in four multiplex groups (fig 1). Primers for long and short PCR were obtained from Genosys (primer sequences not shown because of lack of space but will be made available upon request). Long distance PCR reactions were performed as a 7-plex PCR, using 100 ng of total genomic DNA, 2.5 mmol/l MgCl₂, 400 µmol/l of each dNTP, and 2.5 units of LA enzyme (LA PCR kit II, Takara) in a total volume of 25 µl in LA buffer. Primer concentrations were 0.125 µmol/l (exons 10-11, 12-13, 18-20), 0.25 µmol/l (exons 14-17, 21-24), and 0.5 µmol/l (exons 1-3, 5-9). Amplification was carried out in a PTC-100 thermocycler (MJ Research) for 60 seconds at 94°C, after which the enzyme was added, followed by 32 cycles at 94°C for 15 seconds, 63°C for 60 seconds, 68°C for 12 minutes, and a final extension at 68°C for 12 minutes.

Multiplex short PCR was performed in four multiplex groups. Each group required 0.5 µl of long distance PCR products, 3 mmol/l MgCl₂, 250 µmol/l of each dNTP, and 5 units of

---

**Figure 1** Schematic representation of TDGS as performed for BRCA1 mutation scanning. First, in a single multiplex long distance PCR, over-representation of the target sequences is obtained. Next, individual exons are amplified in four multiplex groups with exon 11 subdivided into 16 overlapping fragments. The products are combined and subjected to 2D electrophoresis, that is, fragments are first separated on the basis of size and subsequently further sorted according to base pair sequence in a denaturing gradient gel.
ampli
taq Gold Taq polymerase (Perkin Elmer) in a total volume of 25 µl in ampli
taq Gold PCR buffer. Primer concentrations were 0.125 µmol/l (exon 11.1), 0.25 µmol/l (exons 9, 11.2, 11.5, 11.6, 11.7, 11.8, 11.12, 11.15, 12, 13, 17, 18, 19, 21, 22, and 23), 0.375 µmol/l (exons 11.3, 11.9), 0.5 µmol/l (exons 3, 6, 8, 10, 11.14, 14, 16, and 24), 0.75 µmol/l (exons 2, 5, 7, 11.4, 11.10, 11.13, 15, and 20), 0.875 µmol/l (11.11), and 1.25 µmol/l (exon 11.16). Amplification was carried out in a PTC-100 thermo
cycler (MJ Research) at 95°C for 12 minutes, followed by 10 cycles at 94°C for 40 seconds, 43°C for one minute minus 0.5°C per cycle, 72°C for one minute 30 seconds, 20 cycles at 94°C for 40 seconds, 40°C for one minute, 72°C for one minute 30 seconds with an increase of one second per cycle, and one cycle at 72°C for 10 minutes. Subsequently, PCR fragments were subjected to one round of complete denaturation and renaturation, that is, 98°C for 10 minutes, 55°C for 30 minutes, and 37°C for 30 minutes, to create heterodu
plex molecules.

TWO DIMENSIONAL ELECTROPHORESIS

Five µl of each multiplex group were combined and mixed with loading buffer (0.25% xylene cyanol, 0.25% bromophenol blue, 15% ficoll, and 100 mmol/l Na2 EDTA). The mixture was directly loaded onto the top gel of an automatic 2D electrophoresis apparatus (CBS Scientific Co, Solana Beach, CA), containing 10% polyacrylamide (acrylamide:bisacrylamide 37.5:1) and electrophoresed for four hours at 200 V in 1 × TAE buffer at room temperature. Subsequently, fragments were further sepa
rated in the denaturing gradient part of the gel. The gradient used was 22-65% UF (100 UF=7 mol/l urea and 40% formamide) in 10% polyacrylamide and electrophoresis was for 16 hours at 100 V in 1 × TAE buffer at 60°C. The gel was stained with a mixture of SYBR green I and II (Molecular Probes) and photographed under UV light using the Imager gel documenta
tion system (Appligene).

NUCLEOTIDE SEQUENCING

Single fragments were amplified by PCR by using the same primers as used for TDGS, but without the GC clamps and with the exclusion of the heteroduplexing cycle. Twenty µl of PCR product were run on a 2% low melting agarose gel and DNA was purified from the excised bands using the SurePure DNA extraction kit (Embi Tec, San Diego, CA). DNA concentra
tions were estimated by UV spectrophotometry and by comparison to a size marker in an agar
ose gel. Sequence reactions contained 90 ng of template, 2 pmol of primer, and 8 µl of Termi
nator Ready Reaction Mix (Perkin Elmer). Cycling was carried out in a PTC-100 thermo
cycler for 25 cycles: 96°C for 10 seconds/50°C for five seconds/60°C for four minutes. Se
quencing products were ethanol precipitated and run in an ABI Prism 310 Genetic Analyzer (Perkin Elmer). Sequencing results were com
pared to the published sequence of Smith et al.13

Results

Two dimensional gene scanning (TDGS) is based on a combination of extensive multiplex PCR and two dimensional electrophoresis. The latter comprises a size separation step followed by denaturing gradient gel electrophoresis (DGGE), and allows single base changes to be distinguished among multiple DNA fragments in parallel.11 PCR amplification of the target sequences is carried out in two steps. First, all coding exons are amplified from genomic DNA by multiplex long distance PCR. Then,
the mixture of long fragments is used as the enriched template for the multiplex amplification of the relatively short (150 to 500 bp) DGGE optimised target fragments. PCR amplification of the BRCA1 gene required a multiplex of seven long fragments serving as template for a total of 37 short PCR fragments. These fragments covered all coding exons of the gene (the non-coding exons 1a, 1b, and the non-coding part of exon 24 were excluded); 16 overlapping fragments were needed to cover the large exon 11 (fig 1). Two dimensional separation of the mixture of all BRCA1 fragments simultaneously in a DGGE gel allowed detection of all mutational variants (fig 2). Heterozygous mutations or polymorphisms are detectable as four spots rather than one spot, representing the two homoduplex and the two heteroduplex variants. In some cases, the two heteroduplex variants are not well separated, resulting in the appearance of three instead of four spots. Examples of heterozygous mutations and polymorphisms are shown in fig 3.

To evaluate the sensitivity and specificity of this test, a panel of 60 DNA samples, previously screened for mutations in BRCA1 using the protein truncation test, alone or in combination with partial nucleotide sequencing, was rescreened by TDGS (table 1). In 14 of these samples a BRCA1 mutation had been identified. Twenty-seven samples were from family members of the mutation positive cases, but one were predicted to lead to a truncated protein. Included were seven deletions of one to 40 nucleotides, four insertions of one nucleotide, and three base substitutions. TDGS was applied to the panel in a blinded fashion and all variants were detected (detection rate of 100%, table 1). No false negatives were found.

Among the 19 samples previously tested negative for exon 11 and rescreened by TDGS, none contained a truncating mutation in exon 11, which was expected. However, five variants were identified by TDGS. One of these was an unclassified variant in exon 11, but this was a single T to G nucleotide substitution, which leads to a Asn to Lys amino acid substitution. Two additional unclassified variants were detected. The first was a G to T at nucleotide 546 in exon 7, which leads to a premature stop codon. The second was a single base pair substitution in intron 5, which leads to aberrant splicing and premature chain termination. These mutations have been described previously.

In addition to the mutations and variants described above, a number of polymorphic variants were detected in the 60 samples. Many of these were identical, as indicated by their characteristic spot configurations. This was confirmed by individual analysis of each variant by 1D DGGE (results not shown). On the basis of different spot configurations, a total of 18 TDGS variants in 12 TDGS fragments were distinguished and confirmed by direct nucleotide sequencing (table 2). None of the variations was false positive. Comparison to the BRCA1 sequence showed that these 18
different spot configurations comprised 15 different sequence variants (table 2).

Of the 15 variants confirmed by sequence analysis, 10 were known polymorphisms, three were unclassified variants, and two were novel variants not found in the BIC database (table 2). Two of the three unclassified variants occurred in subjects in whom a mutation in the gene had also been identified (samples 6 and 8, tables 1 and 2). The third unclassified variant was found in a patient with ovarian cancer for whom no BRCA1 mutation had been identified (sample 39 in table 1). One of the two novel variants (seen in one of 60 subjects), an A to T transversion at position 3107 resulting in a Lys to Asn amino acid change, was found in the same fragment (fragment 11.11), which already contained an unclassified variant. Neither nucleotide sequencing nor PTT had previously detected this variant in this patient with ovarian cancer (sample 39, tables 1 and 2). The second novel variant was also found in only one subject, in exon 23 (sample 36).

As mentioned above, the characteristic spot configuration of recurrent mutations and polymorphisms allows their identification once they have been sequenced. Fig 4 shows three different spot configurations for fragment 11.6 with some of the surrounding fragments as refer-
Table 2  Sequence analysis of polymorphic variants in BRCA1 detected in the panel of 60 subjects

<table>
<thead>
<tr>
<th>Exon</th>
<th>TDGS fragment</th>
<th>Nucleotide position*</th>
<th>Nucleotide change</th>
<th>Type of variant</th>
<th>Codon change</th>
<th>Type of polymorphism†</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>IVS7-34</td>
<td>G→T</td>
<td>Transition</td>
<td>—</td>
<td>—</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11</td>
<td>1186‡</td>
<td>A→G</td>
<td>Transition</td>
<td>356</td>
<td>Gln→Arg</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11</td>
<td>2201‡</td>
<td>C→T</td>
<td>Transition</td>
<td>694</td>
<td>Ser→Ser</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11</td>
<td>2201‡</td>
<td>G→A</td>
<td>Transition</td>
<td>693</td>
<td>Asp→Asn</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11</td>
<td>2434‡</td>
<td>T→C</td>
<td>Transition</td>
<td>771</td>
<td>Leu→Leu</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11.7</td>
<td>11.8</td>
<td>1186‡</td>
<td>Transition</td>
<td>772</td>
<td>Val→Ala</td>
<td>Unclassified variant</td>
</tr>
<tr>
<td>11.9</td>
<td>2731‡</td>
<td>C→T</td>
<td>Transition</td>
<td>871</td>
<td>Pro→Leu</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11.11</td>
<td>3232</td>
<td>A→G</td>
<td>Transition</td>
<td>1038</td>
<td>Gly→Gly</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>11.11</td>
<td>3107</td>
<td>A→T</td>
<td>Transversion</td>
<td>996</td>
<td>Lys→Arg</td>
<td>Not reported</td>
</tr>
<tr>
<td>11.13</td>
<td>3667‡</td>
<td>G→A</td>
<td>Transition</td>
<td>1040</td>
<td>Ser→Ser</td>
<td>Unclassified variant</td>
</tr>
<tr>
<td>13</td>
<td>4427‡</td>
<td>T→C</td>
<td>Transition</td>
<td>1436</td>
<td>Ser→Ser</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>16</td>
<td>4956‡</td>
<td>G→A</td>
<td>Transition</td>
<td>1613</td>
<td>Ser→Gly</td>
<td>Polymorphism</td>
</tr>
<tr>
<td>16</td>
<td>5075</td>
<td>G→A</td>
<td>Transition</td>
<td>1652</td>
<td>Met→His</td>
<td>Unclassified variant</td>
</tr>
<tr>
<td>23</td>
<td>IVS23+48</td>
<td>G→A</td>
<td>Transition</td>
<td>—</td>
<td>—</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

* Nucleotide numbering starts at the first transcribed base according to GenBank entry U14680.
† Named according to BIC as of 10.2.99.
‡ Common polymorphisms defining haplotypes (see also table 3).

Discussion

This study shows that BRCA1 mutation screening by TDGS is an accurate method of mutation detection. Previous evaluations of TDGS tests for other large human disease genes support the claim that DGGE based tests are as accurate as nucleotide sequencing.16–18 It may be that heterozygous mutations are even easier to detect on the basis of the DGGE principle (that is, three or four spots rather than one) than by sequencing (that is, two overlapping peaks). TDGS shows all heterozygous variations in the entire gene.

Table 3  Polymorphisms defining haplotypes

<table>
<thead>
<tr>
<th>Type</th>
<th>2196</th>
<th>2201</th>
<th>2430</th>
<th>2731</th>
<th>3232</th>
<th>3667</th>
<th>4427</th>
<th>4956</th>
<th>%*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A G C T C A A A T A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>B G T C T G G C G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>C A T C T G G C G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>D G C T T A A T A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

* Occurrence in 126 chromosomes.

Van Orsouw, Dhanda, Elhaji, et al

Table 3 Polymorphisms defining haplotypes

<table>
<thead>
<tr>
<th>Type</th>
<th>2196</th>
<th>2201</th>
<th>2430</th>
<th>2731</th>
<th>3232</th>
<th>3667</th>
<th>4427</th>
<th>4956</th>
<th>%*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A G C T C A A A T A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>63</td>
</tr>
<tr>
<td>B G T C T G G C G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>C A T C T G G C G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>D G C T T A A T A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

* Occurrence in 126 chromosomes.
In TDGS, recurrent polymorphisms can be recognised based on their unique spot distribution pattern, but new variants must be sequenced to determine the exact nature of the observed variation. Theoretically, 300 single base substitutions are possible along a 100 base pair DNA fragment. However, so far, an average of only 14 variations has been observed per BRCA1-TDGS fragment (information derived from the BIC). Thus, the possibility that a mutation in a given fragment will show the exact same spot configuration as a known polymorphism is small. However, it has become clear that for a gene containing many polymorphic sites, unequivocal identification of homozygously occurring alleles is only possible by careful examination of the variant in the context of the surrounding spots, which can also represent polymorphic fragments.

Increased standardisation and the application of automated image analysis systems containing information on the spot positions of all known variations should greatly facilitate spot identification for this purpose.

During the course of this study, attention was given to the actual costs of the test. It has previously been shown that the maximum throughput of TDGS for a typical large disease gene, for example, the retinoblastoma susceptibility gene RB1, is 60 tests per technician per week using manual procedures.\(^{19}\) The costs for BRCA1 are higher, because of the need for four instead of only one multiplex groups, and was calculated to be about US$70 per test. The most prominent cost components in the manual procedure involve labour and PCR enzymes. The costs of PCR primers were low since these were based on the total number of reactions possible with the lowest volume that could be ordered. It should be noted that had each fragment been separately amplified rather than amplified the costs associated with enzymes and labour would have been significantly higher.

The costs of the assay can be dramatically reduced by automation. PCR robotics reduces both labour and reagent costs (smaller PCR volumes), while fluorescent imaging using dye labelled primers obviates the need to stain the gels. The costs per sample are then mainly determined by the primers that have to be labelled with a fluorophore, while labour and enzyme costs become vanishingly small. According to our calculations, these proposed automation steps will drive the costs down to about US$10 per test, which is approximately one to two orders of magnitude lower than current sequencing protocols.\(^{19}\) This anticipated increase in cost-efficiency is based solely on commercially available instrumentation and not on technological improvement.

The TDGS test is ideal for large studies to establish the potential value of screening for mutations in BRCA1 as an aid to breast cancer management. However, the hereditary breast cancer syndrome is genetically heterogeneous and other susceptibility genes have been identified, including BRCA2, TP53, PTEN, and, arguably, ATM.\(^{20-24}\) In order to use genetic screening as an aid for risk assessment of breast cancer, a battery of genetic tests should be applied. TDGS is flexible since its design permits the rapid implementation of a test to include additional genes. TDGS tests, which screen multiple genes in a single assay, can and have been developed.\(^{10}\)

We are grateful to Shu-Hua Xu, David Rines, Stacey Caron, and Sean McGrath for their technical assistance. This research was supported by the Massachusetts Department of Public Health (3408790D4047, the Harvard Nathan Shock Center of Excellence in the Basic Biology of Aging (1P30AG13314-01), the Molecular Biological Applications Program of the Academy of Applied Science, and the Women’s Cancer Program, Dana-Farber Partners Cancer Center. CE is the Lawrence and Susan Marx Investigator in Human Cancer Genetics.