

ORIGINAL ARTICLE

# Detection of high frequency of mutations in a breast and/or ovarian cancer cohort: implications of embracing a multi-gene panel in molecular diagnosis in India

Ashraf U Mannan<sup>1</sup>, Jaya Singh<sup>1</sup>, Ravikiran Lakshmikeshava<sup>1</sup>, Nishita Thota<sup>1</sup>, Suhasini Singh<sup>1</sup>, TS Sowmya<sup>1</sup>, Avshesh Mishra<sup>1</sup>, Aditi Sinha<sup>1</sup>, Shivani Deshwal<sup>1</sup>, Megha R Soni<sup>1</sup>, Anbukayalvizhi Chandrasekar<sup>1</sup>, Bhargavi Ramesh<sup>1</sup>, Bharat Ramamurthy<sup>1</sup>, Shila Padhi<sup>1</sup>, Payal Manek<sup>1</sup>, Ravi Ramalingam<sup>1</sup>, Suman Kapoor<sup>1</sup>, Mithua Ghosh<sup>2</sup>, Satish Sankaran<sup>1</sup>, Arunabha Ghosh<sup>1</sup>, Vamsi Veeramachaneni<sup>1</sup>, Preveen Ramamoorthy<sup>1,3</sup>, Ramesh Hariharan<sup>1,4</sup> and Kalyanasundaram Subramanian<sup>1</sup>

Breast and/or ovarian cancer (BOC) are among the most frequently diagnosed forms of hereditary cancers and leading cause of death in India. This emphasizes on the need for a cost-effective method for early detection of these cancers. We sequenced 141 unrelated patients and families with BOC using the TruSight Cancer panel, which includes 13 genes strongly associated with risk of inherited BOC. Multi-gene sequencing was done on the Illumina MiSeq platform. Genetic variations were identified using the Strand NGS software and interpreted using the StrandOmics platform. We were able to detect pathogenic mutations in 51 (36.2%) cases, out of which 19 were novel mutations. When we considered familial breast cancer cases only, the detection rate increased to 52%. When cases were stratified based on age of diagnosis into three categories,  $\leq 40$  years, 40–50 years and  $> 50$  years, the detection rates were higher in the first two categories (44.4% and 53.4%, respectively) as compared with the third category, in which it was 26.9%. Our study suggests that next-generation sequencing-based multi-gene panels increase the sensitivity of mutation detection and help in identifying patients with a high risk of developing cancer as compared with sequential tests of individual genes.

*Journal of Human Genetics* (2016) 61, 515–522; doi:10.1038/jhg.2016.4; published online 25 February 2016

## INTRODUCTION

Breast cancer is the most frequently diagnosed cancer in the world and the leading cause of death among women. According to GLOBOCAN, 145 000 new cancer cases were diagnosed and 70 000 deaths occurred in India in 2012.<sup>1</sup> Considering the population size, the incidence of breast cancer in India appears to be lower when compared with developed nations such as the United States where it was 233 000 and European Union (EU-28) where it was 362 000 in 2012.<sup>1</sup> However, the ratio of number of deaths to new cases in India was 48.3% (70 000/145 000), which was much higher as compared with 18.9% (44 000/233 000) in the United States and 25.4% (92 000/362 000) in the European Union.<sup>2</sup> Hence, although the incidence seems to be lower in India than in the developed nations, mortality rates are higher, which can be attributed to lifestyle changes, delayed introduction of effective breast cancer screening programs and limited access to treatment.<sup>1,3</sup>

The burden of breast cancer on the Indian health-care system has been steadily increasing, thus stressing the need for cost-effective methods for early detection, screening, surveillance and follow-up programs. If detected at an early stage, breast cancer is eminently treatable. Women with early-onset breast cancer tend to have good prognosis provided they have access to multi-disciplinary treatment options. Early detection strategies, over the last 40–50 years in the developed nations, have resulted in consistent decrease in the breast cancer mortality.

Molecular diagnosis of hereditary breast and/or ovarian cancer (HBOC) is primarily based on the identification of germline mutations within the *BRCA1* and *BRCA2* (*BRCA1/2*) genes.<sup>4,5</sup> Mutations in these two tumor suppressor genes account for ~25% of hereditary breast cancers and ~5% of all breast cancers.<sup>6</sup> Besides *BRCA1/2*, mutations in several other genes causing cancer syndromes, such as *CDH1* (hereditary diffuse gastric cancer),<sup>7</sup> *PTEN* (Cowden syndrome),<sup>8</sup>

<sup>1</sup>Strand Center for Genomics and Personalized Medicine, Strand Life Sciences, Bangalore, India; <sup>2</sup>Triesta Sciences, Health Care Global Enterprises Limited, Bangalore, India; <sup>3</sup>Department of Medicine, Division of Bioinformatics and Personalized Medicine, University of Colorado School of Medicine, Aurora, CO, USA and <sup>4</sup>Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India

Correspondence: Dr AU Mannan, Strand Center for Genomics and Personalized Medicine, Strand Life Sciences, Kirloskar Business Park, Bellary Road, Hebbal, Bangalore 560024, India.

E-mail: ashraf@strandls.com

Received 13 October 2015; revised 17 December 2015; accepted 10 January 2016; published online 25 February 2016

*STK11* (Peutz–Jeghers syndrome)<sup>9</sup> and *TP53* (Li–Fraumeni syndrome),<sup>10</sup> confer an increased risk of breast cancer. Recent studies have identified germline mutations in additional genes including *ATM*, *BRIP1*, *CHEK2*, *NBN*, *PALB2*, *RAD51C* and *RAD51D*, which were shown to increase the risk of breast cancer.<sup>11–18</sup>

In India, only a few studies have been conducted to determine the prevalence of mutations in the *BRCA1/2* genes in various Indian cohorts. The mutation rates reported range from 2.9 to 28%.<sup>19–23</sup> In one of the first studies in India, Kumar *et al.*<sup>24</sup> screened 14 patients with family history of HBOC by using conformation sensitive gel electrophoresis followed by sequencing and identified *BRCA1* mutations in three patients. However, Saxena *et al.*<sup>25</sup> screened 204 North Indian patients affected with breast cancer, including 105 early-onset ( $\leq 40$  years), 65 late-onset ( $> 40$  years) cases and 34 familial cases by heteroduplex analysis combined with sequencing, the detection rate was  $\sim 2.9\%$  in the *BRCA1/2* genes both overall (6/204) as well as familial cases (1/34). In another study, Vaidyanathan *et al.*<sup>19</sup> screened 61 patients with family history of HBOC from Southern India by conformation sensitive gel electrophoresis combined with sequencing and reported a prevalence rate of 28% (17/61). Collectively, from these studies it is evident that there is a high variability in the reported prevalence of *BRCA1/2* mutations in various cohorts. In a recent study, Rajkumar *et al.*<sup>26</sup> screened 91 patients with family history of HBOC or early onset of cancer from Southern India by targeted re-sequencing of a multi-gene panel and reported a mutation rate of 26.4% (24/91).

In order to accurately determine the prevalence rate in India, more comprehensive studies examining a wider mutational spectrum in the high-risk genes in addition to *BRCA1/2* and in cohorts that include breast cancer patients from different regions of India are required. New advances in genomic technologies such as next-generation sequencing (NGS), allows the sequencing and analysis of multiple genes associated with a disease/cancer efficiently at a significantly lower cost as compared with the traditional methods. In the current study, we used an NGS-based approach to screen 13 genes associated with breast cancer in an Indian cohort (141 cases) and report a high prevalence rate (36.2%) of mutations.

## MATERIALS AND METHODS

### Families/cases and sample collection

Saliva or peripheral blood samples were obtained from 141 unrelated patients/families with indication of breast and/or ovarian cancer (BOC), referred to our laboratory between March 2014 and June 2015. Out of the 141 individuals referred, 62 had family history of BOC, 11 sporadic cases with no known family history, 38 had unknown inheritance pattern as family history details were not provided and 30 risk-prediction cases with strong family history. Informed consent was obtained from all subjects and this study was approved by Institutional Ethics Committee of Strand Life Sciences. Patient clinical data and details of family history of cancer(s) were obtained from the referring oncologist and whenever possible through pre-test genetic counseling by our genetic counselor. There were 111 patients or families with breast cancer (median age of patients 45 years), 11 families with ovarian cancer (median age of patients 52 years) and 19 families with breast and ovarian cancer (median age of patients 50 years). The age distribution of all patients ranged from 22 to 77 years with a median age of 47 years. Out of the 141 patients studied, there were 41 patients  $< 40$  years of age (29%), 48 patients in the age group of 40–50 years (34%) and 52 patients above 50 years (37%).

DNA was extracted from saliva using PrepIT-L2P kit (DNA Genotek, Ottawa, ON, Canada). For blood samples, QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany) was used for DNA isolation. The concentration of DNA was determined using the Qubit fluorimeter (Life Technologies, Carlsbad, CA, USA).

### Multi-gene panel selection

We sequenced the patient gDNA using the Trusight Cancer sequencing panel (Illumina, San Diego, CA, USA) that contains 94 genes suspected to have a role in cancer predisposition. An analytical validation study on this panel revealed that only 86 genes showed coverage of  $> 95\%$  ( $\geq 20$  reads). In these 86 genes, the mean read coverage was  $450\times$  with  $\sim 99\%$  of exonic and essential splice sites regions covered at  $\geq 20$  reads. We performed an extensive curation of these genes by reviewing published literature and online databases (HGMD, OMIM and Orphanet) and determined that 13 genes *ATM* (NM\_000051.3), *BRCA1* (NM\_007294.3), *BRCA2* (NM\_000059.3), *BRIP1* (NM\_032043.2), *CDHI* (NM\_004360.3), *CHEK2* (NM\_007194.3), *NBN* (NM\_002485.4), *PALB2* (NM\_024675.3), *PTEN* (NM\_000314.4), *RAD51C* (NM\_058216.1), *RAD51D* (NM\_002878.3), *STK11* (NM\_000455.4) and *TP53* (NM\_000546.4) were associated with HBOC predisposition.

### NGS—library preparation and sequencing

We used the Nextera DNA library preparation protocol (Illumina) to convert input genomic DNA (gDNA) into adapter-tagged indexed libraries. Approximately 50 ng of input gDNA was used in the tagmentation process, which involves simultaneous fragmentation and adapter tagging of gDNA followed by adapter ligation. This was followed by limited cycles of PCR (ABI 9700, Life Technologies) to allow the incorporation of sample-specific indices or multiplex identifier barcodes. The quality of the library was assessed using the BioAnalyzer (Agilent, Santa Clara, CA, USA). Next, 500 ng of individual libraries were pooled in batches of 9–12 samples and hybridized to biotin-labeled probes specific to the targeted regions. The pool was enriched for the target genomic regions by adding streptavidin beads that bind to the biotinylated probes. The biotinylated gDNA fragments bound to the streptavidin beads were magnetically pulled down from the solution. The partly enriched gDNA fragments were then eluted from the beads and subjected to a second round of hybridization. The tagged and amplified sample libraries were checked for quality and quantified using the BioAnalyzer (Agilent). Upto 6–10 pM of the pooled library was loaded and sequenced on the MiSeq platform (Illumina), according to the manufacturer's instructions.

### NGS—data analysis and interpretation

The trimmed FASTQ files were generated using MiSeq Reporter (Illumina). Reads were aligned against the whole-genome build: hg19 using Strand NGS v2.1.6. Strand NGS is an integrated platform that provides analysis, management and visualization tools for NGS data. It has a comprehensive DNA-Seq pipeline that includes alignment, read-quality assessment, filtering, small variant calling and copy number variation detection (<http://www.strand-ngs.com/>). In brief, the 150 bp paired-end reads were first aligned against the hg19 reference genome. Five base pairs from the 3' end of reads were trimmed, as were 3' end bases with base-quality below 10. Reads with length  $< 25$  bp after trimming were not considered for alignment. A maximum of five matches of alignment score at least 90% were computed. A gap size of 45% was allowed for detection of medium-sized indels (small insertion/deletion/duplication) up to 67 bp. Post alignment, reads were re-aligned using the local realignment functionality in Strand NGS v2.1.6. Following this, reads that failed vendor QC (quality control), reads with average quality  $< 20$ , reads with ambiguous characters and all duplicate reads were filtered out. The variant detection algorithm in Strand NGS v2.1.6 was then used to detect variants in the target regions covered by a minimum of 20 reads with at least two variants reads. Variants with a decibel score of at least 50 were reported and consecutive single base variants were merged to create multi-base variants in the final variant call format file. The variant call format file along with a low coverage ( $< 20$  reads) file generated using the filtered read list were uploaded into StrandOmics v3.0 (<http://www.strandl.com/strandomics/>; a proprietary clinical genomics interpretation and reporting platform from Strand Life Sciences) for all subsequent analysis and variant interpretation. The StrandOmics variant annotation platform integrates resources from various public databases and proprietary content ( $\sim 40\,000$  extra curated variant records). Public databases, such as: Uniprot and OMIM have been integrated into StrandOmics to extract information on gene association with the disease and protein domains for each gene. StrandOmics includes algorithms to assist in identifying variant impact from both public content

(HGMD, ClinVar, OMIM, ARUP, Uniprot, links to dbSNP, 1000 Genomes, Exome Variant Server, Exome Aggregation Consortium and proprietary content (~40 000 extra curated variant records) on genes, phenotypes and diseases. It is also integrated with *in silico* prediction tools, such as: SIFT, PolyPhen HVAR/HDIV, Mutation Taster, Mutation Assessor, FATHMM and LRT to assess the pathogenicity of the missense variants. Algorithms that evaluate sequence conservation at the variant position from the protein sequence assembly from 46 species (including primates, placental mammals and other vertebrates) are also built into StrandOmics.

To assess the pathogenicity of the variants in the essential splice sites and exon–intron boundaries, splice site prediction tools (NNSPLICE and ASSP) have been integrated to the NGS interpretation pipeline. The ‘interpretation interface’ in StrandOmics allows quick filtering and evaluation of variants identified in a sample.

### Variant classification

The identified variants in this study were classified according to the ACMG (American Society of Medical Genetics and Genomics) recommendations for standards for interpretation and reporting of sequence variations.<sup>27</sup> The variants were classified into five categories: (1) pathogenic, (2) likely pathogenic, (3) variant of uncertain significance (VUS), (4) likely benign and (5) benign.

### Confirmation of the detected variants by Sanger sequencing

To confirm pathogenic/likely pathogenic variants detected by NGS, we performed Sanger sequencing. Primers flanking each variant were designed and the genomic region encompassing the variant was amplified by PCR. Primer sequences and PCR conditions are available on request. The PCR products were purified using the GeneJet PCR Purification Kit (Thermo Fisher, Waltham, MA, USA) according to manufacturer's instructions. The purified PCR products were sequenced using both forward and reverse primers (which were used for the PCR amplification) using BigDye Terminator v3.1 kit (Life Technologies). The sequencing PCR products were purified and subsequently analyzed by the 3500DX Genetic Analyzer (Life Technologies).

### Copy number variation analysis for large deletion/duplication

In addition to single nucleotide variants and small indels, copy number analysis was performed to identify large deletions or insertions ranging from single exons to full genes. This was done by taking each non-overlapping target region in turn, of which there are 1736, and comparing normalized read coverage across 11 other samples from the same run. Normalized coverage-based copy number values (CNVs) and *Z*-scores<sup>28</sup> for each panel region were computed, using Strand NGS v2.1.6. For each sample, potential copy number variation changes in the genes of interest were identified for manual interpretation based on the following cutoffs: CNV > 3, *Z*-score > 2 for duplication and CNV < 1.2, *Z*-score < -2 for deletions.

### MLPA analysis

A large deletion identified in the *BRCA1* gene using the CNV algorithm above was confirmed by Multiplex ligation-dependent probe amplification (MLPA). MLPA was performed with 50 ng of gDNA according to manufacturer's instructions using SALSA MLPA P002 *BRCA1* probemix (MRC-Holland, Amsterdam, the Netherlands). Probe amplification products were run on the Genetic Analyzer 3500 DX (Life Technologies). MLPA peak plots were visualized, normalized and the dosage ratios were calculated by using the Coffalyser.Net software (MRC-Holland).

### Statistical analysis

BOC cases from our cohort with known age of onset (105 individuals) were considered for statistical analysis. The patients were grouped into three categories: <40 years, 40–50 years and >50 years and the number of pathogenic and likely pathogenic variations in each category were calculated. We tested the following null hypothesis: there is no difference in the proportion of subjects with pathogenic and likely pathogenic variations among these three age categories. For this, we used a Monte Carlo  $\chi^2$ -test. First, we computed the  $\chi^2$ -statistic on the 3 × 2 contingency table containing the actual observations.

We then used a Monte Carlo simulation procedure<sup>29</sup> to generate 10 000 contingency tables by permuting the 105 individuals randomly among the three age categories while retaining the size of each category; for each table, we computed the  $\chi^2$ -statistic. The fraction of random instances where this statistic exceeded the statistic computed on the actual observations was then calculated and reported as the *P*-value. All these calculations were performed in R (R project for statistical computing) v3.1.1.

## RESULTS

### Mutation spectrum in breast and ovarian cancer cohort

We screened a total of 141 DNA samples of unrelated BOC individuals/families for mutations in the 13 genes associated with HBOC by NGS. The clinical summary of all the subjects are provided in Supplementary table S1 (note: Supplementary Tables S1–S2 and Supplementary Figure S1–S2 are available as online Supplementary Information). A total of 50 pathogenic and likely pathogenic variants (hereby referred as mutations) were identified in 141 subjects, which account for 35.5% cases (Table 1). All 50 mutations identified by NGS were confirmed by Sanger sequencing. Next, we performed CNV analysis for detection of large deletions/duplications in 91 cases that were negative for pathogenic mutations through sequence analysis. We identified one large heterozygous deletion in the *BRCA1* gene, which accounts for 1.1% of the 91 BOC cases (Table 1). The combined mutation detection rate was 36.2% (51/141 cases) (Figure 1a). In our study, we detected 42 unique mutations of which 19 were novel (Table 1). None of the 19 identified novel mutations in our study was found in the 250 control chromosomes. All types of mutations, namely missense, nonsense, splice site, indel and large deletion were detected in our screen (Figure 1b). Mutations were detected in 8 of 13 tested genes: *ATM*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK2*, *PALB2*, *RAD51C* and *TP53*. Out of 51 mutations, 37 (72.5%) were detected in the *BRCA1/2* genes and 14 (27.5%) were detected in non-*BRCA* genes (Figure 1a).

### *BRCA1* and *BRCA2* mutations

All the 37 mutations, which were detected in the *BRCA1/2* genes, were truncating mutations; no missense mutation was identified in our screen (Figure 1b). Among the truncating mutations, we detected 21 indels, 1 large deletion, 9 nonsense and 6 splice site mutations (Figure 1b). Furthermore, majority (91.9%) of these mutations was detected in the *BRCA1* gene; only three mutations were detected in the *BRCA2* gene (Table 1). We also detected four recurrent mutations in the *BRCA1* gene: c.68\_69delAG, c.5074+1G>A, c.3352C>T and c.4837\_4838delinsGCC (Table 1). The c.68\_69delAG mutation was detected six times and it has been reported as an Ashkenazi Jewish founder mutation (popularly known as 185delAG).<sup>30</sup> In our screen, the c.5074+1G>A mutation was detected three times, c.3352C>T and c.4837\_4838delinsGCC mutations were each detected twice (Table 1). The clinical characteristics of the recurrent *BRCA1* mutation carriers show that majority of the recurrent mutations appeared to be associated with familial BOC cancer (Table 1).

### Mutations in the non-*BRCA* genes

In our cohort, 104 cases were negative for *BRCA1/2* mutations; we detected 14 additional mutations in the non-*BRCA* genes (*ATM*, *BRIP1*, *CHEK2*, *PALB2*, *RAD51C* and *TP53*) (Table 1), which accounts for 13.5% cases. Seven mutations were identified in the *TP53* gene including, two missense substitutions (p.Ser241Phe and p.Arg282Trp), two nonsense substitutions (p.Tyr163Ter and p.Arg342-Ter), two splice site variants (c.559+1G>A and c.994-1G>C) and one indel (c.294\_297delTTCC). In the *PALB2* gene, we identified three

**Table 1** List of pathogenic/likely pathogenic variations identified in breast and/or ovarian cancer cohort

No.	Gene	Exon/ intron	Case no.	Type of mutation	cDNA	Protein	Familial/ sporadic	Age at diagnosis (years)	Reference/novel
1	ATM	Ex 8	S-1044	Indel	c.1057_1058delTG	p.Cys353SerfsTer5	F	52	Novel
2	BRCA1	Ex 1–2	S-0837	LD	c.(?-119)_(80+1_81-1)del	Unknown	F	35	Novel
3	BRCA1	Ex 2	P-0194	Indel	c.68_69delAG	p.Glu23ValfsTer17	F	62	23788959
4	BRCA1	Ex 2	P-0005	Indel	c.68_69delAG	p.Glu23ValfsTer17	?	39	23788959
5	BRCA1	Ex 2	S-0395	Indel	c.68_69delAG	p.Glu23ValfsTer17	F	48	23788959
6	BRCA1	Ex 2	S-0746	Indel	c.68_69delAG	p.Glu23ValfsTer17	F	41	23788959
7	BRCA1	Ex 2	S-0827	Indel	c.68_69delAG	p.Glu23ValfsTer17	?	32	23788959
8	BRCA1	Ex 2	S-0900	Indel	c.68_69delAG	p.Glu23ValfsTer17	?	33	23788959
9	BRCA1	Ex 7	S-0063	Indel	c.512dupT	p.Gln172ThrfsTer10	F	57	Novel
10	BRCA1	Int 9	S-0392	SS	c.671-1G>T	Unknown	F	38	rs80358020 <sup>a</sup>
11	BRCA1	Ex 10	S-0569	Indel	c.779dupA	p.Tyr261ValfsTer4	F	40	Novel
12	BRCA1	Ex 10	S-0570	Indel	c.2214dupT	p.Lys739TerfsTer1	S	47	Novel
13	BRCA1	Ex 10	S-0567	Indel	c.3129_3138delTATTAATGAA	p.Asn1043LysfsTer2	F	32	Novel
14	BRCA1	Ex 10	N-0012	NS	c.3352C>T	p.Gln1118Ter	?	48	22762150
15	BRCA1	Ex 10	S-0726	NS	c.3352C>T	p.Gln1118Ter	?	34	22762150
16	BRCA1	Ex 10	S-0574	NS	c.3607C>T	p.Arg1203Ter	F	46	23961350
17	BRCA1	Ex 10	R-0279	Indel	c.4065_4068delTCAA	p.Asn1355LysfsTer10	?	42	14757871
18	BRCA1	Ex 10	S-0877	Indel	c.809delA	p.His270LeufsTer28	S	54	RCV000112793 <sup>b</sup>
19	BRCA1	Ex 10	S-1051	NS	c.1155G>A	p.Trp385Ter	F	<40	Novel
20	BRCA1	Ex 10	S-1133	Indel	c.1416delC	p.Asn473ThrfsTer2	F	33	Novel
21	BRCA1	Ex 11	V-0089	Indel	c.4120_4121delAG	p.Ser1374TerfsTer1	?	46	22006311
22	BRCA1	EX 11	S-0890	NS	c.4183C>T	p.Gln1395Ter	F	46	RCV000112239 <sup>b</sup>
23	BRCA1	Ex 11	R-0165	Indel	c.4165_4166delAG	p.Ser1389TerfsTer1	F	29	9760198
24	BRCA1	Ex 12	S-0186	NS	c.4327C>T	p.Arg1443Ter	F	NA	9792861
25	BRCA1	Ex 12	S-0571	NS	c.4349C>A	p.Ser1450Ter	F	52	Novel
26	BRCA1	Int 13	S-1175	SS	c.4484+1G>A	Unknown	F	49	rs80358063 <sup>a</sup>
27	BRCA1	Ex 15	S-0124	Indel	c.4837_4838delinsGCC	p.Ser1613AlafsTer9	F	37	RCV000157652 <sup>b</sup>
28	BRCA1	Ex 15	S-0485	Indel	c.4837_4838delinsGCC	p.Ser1613AlafsTer9	F	46	RCV000157652 <sup>b</sup>
29	BRCA1	Ex 16	S-0573	Indel	c.5035delC	p.Leu1679TerfsTer1	F	36	21156238
30	BRCA1	Int 16	P-0219	SS	c.5074+1G>A	Unknown	F	NA	17924331
31	BRCA1	Int 16	P-0008	SS	c.5074+1G>A	Unknown	F	45	17924331
32	BRCA1	Int 16	S-0589	SS	c.5074+1G>A	Unknown	?	49	17924331
33	BRCA1	Ex 17	N-0029	Indel	c.5133delA	p.Lys1711AsnfsTer2	?	50	Novel
34	BRCA1	Ex 22	P-0277	Indel	c.5440dupG	p. Ala1814GlyfsTer16	F	27	Novel
35	BRCA1	Int 22	S-1117	SS	c.5467+1G>A	Unknown	F	24	rs80358145 <sup>a</sup>
36	BRCA2	Ex 3	S-0595	NS	c.92G>A	p.Trp31Ter	F	35	17080309
37	BRCA2	Ex 11	S-1594	NS	c.3187C>T	p.Gln1063Ter	F	50	Novel
38	BRCA2	Ex 11	S-0590	Indel	c.3186_3189delTCAG	p.Ser1064LeufsTer12	F	42	Novel
39	BRIP1	Ex 7	S-1178	Indel	c.633delT	p.Gly212AlafsTer62	S	50	Exact <sup>c</sup>
40	CHEK2	EX 8	S-1456	NS	c.852C>A	p.Cys284Ter	F	51	Novel
41	PALB2	Ex 5	S-0790	NS	c.2012T>G	p.Leu671Ter	F	57	Novel
42	PALB2	Ex 7	S-0743	Indel	c.2716delT	p.Trp906GlyfsTer17	?	42	Novel
43	PALB2	Ex 8	P-0094	Indel	c.2762delA	p.Gln921ArgfsTer79	F	NA	Novel
44	RAD51C	Ex 3	S-0258	NS	c.409C>T	p.Gln137Ter	F	46	Novel
45	TP53	Ex 4	S-0764	Indel	c.294_297delTTCC	p.Ser99ArgfsTer23	F	40	Novel
46	TP53	Ex 5	S-0568	NS	c.489C>G	p.Tyr163Ter	F	42	8392033
47	TP53	Int 5	S-0407	SS	c.559+1G>A	Unknown	F	46	7903951
48	TP53	Ex 7	S-0313	MS	c.722C>T	p.Ser241Phe	F	60	24744791
49	TP53	Ex 8	P-0221	MS	c.844C>T	p.Arg282Trp	F	NA	22494262
50	TP53	Int 9	S-1347	SS	c.994-1G>C	Unknown	S	33	10980596
51	TP53	Ex 10	N-0025	NS	c.1024C>T	p.Arg342Ter	?	27	17567834

Abbreviations: F, familial; Indel, small insertion or deletion; LD, large deletion; MS, missense; NA, not applicable as risk-prediction cases; NS, nonsense; S, sporadic; SS, splice site; ?, unknown inheritance pattern.

<sup>a</sup>dbSNP database.

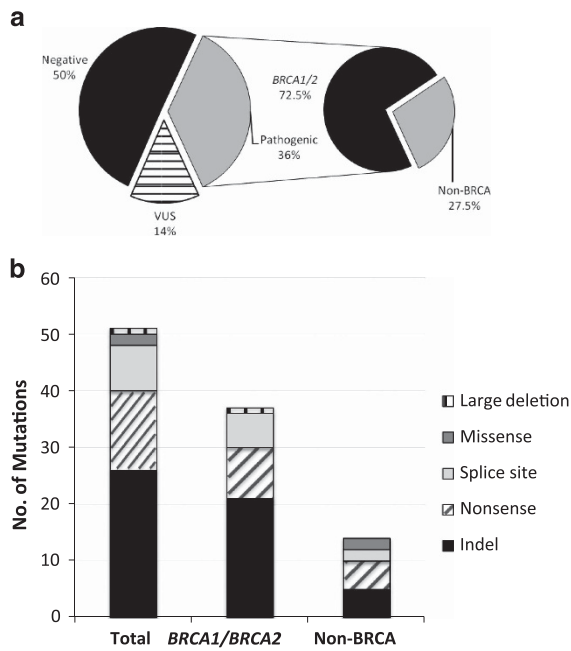
<sup>b</sup>ClinVar database.

<sup>c</sup>Exome Aggregation Consortium database.

mutations (c.2012T>G, c.2716delT and c.2762delA). One mutation was identified each in the ATM (c.1057\_1058delTG), BRIP1 (c.633delT), CHEK2 (c.852C>A) and RAD51C (c.409C>T) gene.

#### Identification of VUS

In our study, the variants that could not be classified as pathogenic/likely pathogenic or as benign/likely benign as per ACMG guidelines



**Figure 1** Mutation spectrum in breast and/or ovarian cancer cohort. (a) A pie chart depicting the total proportion of mutations (36%) identified in our study; of these, majority (72.5%) of mutations was detected in the *BRCA1/2* genes. The proportion of mutations in the non-*BRCA* genes was 27.5%. (b) In our study, the spectrum of mutation types detected was: indel, large deletion, missense, nonsense and splice site. Interestingly, all *BRCA1/2* mutations identified in our study were truncating mutations; missense mutations were only detected in the *TP53* gene.

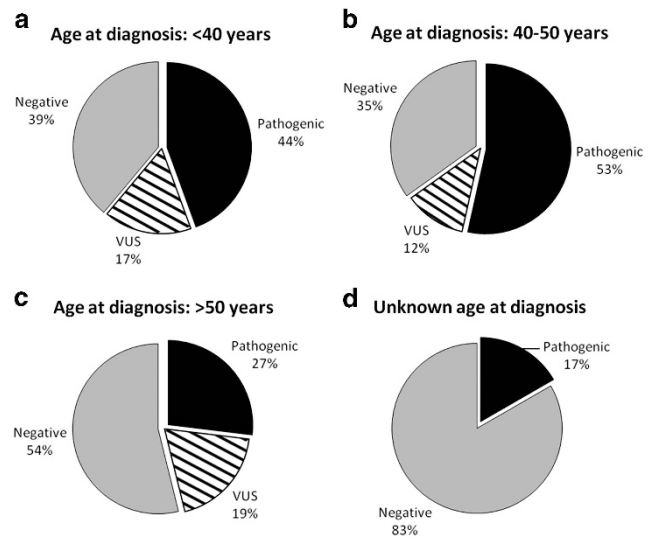
were labeled as VUS. We assessed the functional consequence of these identified VUSs on protein function based on evolutionary conservation at the variant amino-acid position (across 43 species spanning primates, placental mammals and other vertebrates), prediction from seven *in silico* analysis tools and information from literature and public databases (ClinVar, BIC and dbSNP) (Supplementary Table S2). We identified 22 VUSs and with the exception of one in-frame deletion (p.Lys1110del) in the *BRCA1* gene, all the identified VUSs were missense variants (Supplementary Table S2).

#### Detection of a large deletion in the *BRCA1* gene

By NGS CNV analysis, we detected a large heterozygous deletion in the *BRCA1* gene in one case of familial breast and ovarian cancer (Supplementary Figure S1). Our NGS panel enriches only the coding and flanking splice sites sequence of the target genes. In *BRCA1* (NM\_007294), the start codon is present in exon 2, therefore CNV analysis indicated deletion of exon 2 only (Supplementary Figure S1A); however, as the region upstream of the *BRCA1* gene was not sequenced, any deletions in this upstream region could not be assessed by the NGS data alone. So we performed MLPA analysis on this sample and were able to detect a heterozygous deletion encompassing exon 1–2 and some upstream region of the *BRCA1* gene (Supplementary Figure S1B).

#### Mutation detection rate based on cancer types and family history of cancer

In the analyzed cohort, the mutation detection rate in the 62 breast cancer cases with a family history of BOC was 52% (32/62), whereas in the 11 sporadic breast cancer cases, it was 36% (4/11). In the 30 risk-



**Figure 2** Correlation between age at diagnosis and mutation detection rate. We stratified our cohort into four age groups; namely <40 years, 40–50 years, >50 years and age at diagnosis-unknown; to determine whether mutation detection rate is correlated with age at diagnosis of cancer. The mutation rate in the age groups: <40 years, 40–50 years were much higher (>44%) as compared with age group >50 years (27%) (a–c). However, the difference was statistically significant only between the age groups: 40–50 years and >50 years ( $\chi^2$ -test,  $P=0.031$ ). The mutation rate in the age group-unknown was found to be 17% (d).

prediction cases involving unaffected individuals with a family history of BOC, it was 13.3% (4/30) and in the 38 cases with unknown family history, it was 29% (11/38) (Supplementary Figure S2). However, owing to small sample size, especially for the sporadic breast cancer group, the difference in mutation rate between familial and sporadic breast cancer groups was not statistically significant ( $\chi^2$ -test,  $P>0.05$ ). The mutation detection rate in TNBC (triple-negative breast cancer) cases was 23.5% (4/17) and in cases with only ovarian cancer, it was 36% (4/11) (Supplementary Figure S2).

#### Correlation between age at diagnosis and mutation detection rate

To determine whether the mutation detection rate in our screen was correlated with age at diagnosis of cancer, we stratified the patients into four age groups; namely <40 years, 40–50 years, >50 years and unknown. The mutation detection rate in the age groups were as following: 44.4% (16/36) in <40 years group, 53.4% (23/43) in 40–50 years group, 26.9% (7/26) in >50 years group and 16.7% (1/6) in unknown age group (Figure 2). The proportion of mutations in the <40 years and the 40–50 years categories were much higher (>44%) as compared with the proportion in the >50 years category (26.9%). The difference between 40–50 years group and >50 years group was statistically significant as determined by a Monte Carlo simulation assessing the  $\chi^2$ -statistic ( $P=0.031$ ). However, the difference in the proportion of mutations between <40 years group and >50 years categories was not statistically significant ( $P>0.05$ ).

#### DISCUSSION

HBOC shows significant genetic heterogeneity, and numerous studies have shown that besides *BRCA1/2* genes, mutations in several other genes have been also associated with HBOC. With the advent of NGS, genetic tests based on multi-gene panels are being offered by diagnostic laboratories worldwide. Recently, two large-scale studies

(> 1000 individuals) used NGS-based multi-gene panels and reported that in these cohorts, 9.0–9.3% individuals carried a *BRCA1/2* mutation and 3.9% individuals carried a mutation in non-*BRCA* genes associated with HBOC.<sup>31,32</sup> Several other large-scale screens also indicate a higher diagnostic yield in HBOC cohorts using multi-gene panels as compared with testing for *BRCA1/2* mutations alone.<sup>33–35</sup> Previous studies in Indian HBOC cohorts (>50 individuals) have reported variable prevalence rate of mutation ranging from 2.9 to 28%.<sup>19–23,26</sup>

In the current study, we used a NGS-based 13-gene panel to screen 141 unrelated individuals/families with an indication of BOC to determine the mutation detection rate in an Indian cohort. We detected 51 mutations in the cohort of which 19 were novel mutations. The overall mutation detection rate in our cohort was 36.2% (51/141). When, we considered familial breast cancer cases alone, the detection rate increased to 52% (32/62). Although, the mutation rate among sporadic breast cancer cases was also high (36.4% (4/11)), the sample size was very small. In India, previously, the highest mutation detection rate of 28% was reported in a familial BOC cohort (61 cases) from South India; by screening for *BRCA1/2* mutations by using conformation sensitive gel electrophoresis followed by sequencing.<sup>19</sup> In another study, Saxena *et al.*<sup>25</sup> reported a mutation detection rate of 2.9% in the *BRCA1/2* genes in 34 familial BOC cases from North India by heteroduplex analysis combined with sequencing. Hence, our finding suggests that NGS-based multi-gene panel screening considerably improves detection rate among familial breast cancer cases in India. Furthermore, in a recent study, Rajkumar *et al.*<sup>26</sup> reported a detection rate of 25.7% (19/74), also by using a NGS-based multi-gene panel in those BOC cases, which were negative for *BRCA1/2* mutations by an earlier PCR-dHPLC (PCR-denaturing high performance liquid chromatography)-based screening. This result further corroborates our finding that the NGS-based multi-gene panels increase the sensitivity of mutation detection in Indian BOC cohorts.

Of the 51 mutations, 37 were detected in the *BRCA1/2* genes and 14 in non-*BRCA* genes (*ATM*, *BRIP1*, *CHEK2*, *PALB2*, *RAD51C* and *TP53*). The proportion of non-*BRCA* gene mutations in our study was 27.5%, which is comparable to the detection rate reported in other NGS-based multi-gene panel studies.<sup>31,33–35</sup> With the exception of *TP53* (2 missense), all mutations detected in our study were protein truncating mutations, which will likely result in loss-of-function. In a recent study, it was shown that majority of patients who tested positive for non-*BRCA* mutations considered disease-specific screening or preventive measures and additional family testing, which demonstrate the actionability of the non-*BRCA* findings in changing the course of clinical management.<sup>36</sup>

Interestingly, among 37 mutations, which were detected in the *BRCA1/2* genes in our study, 34 mutations were detected in *BRCA1* and only 3 mutations were identified in *BRCA2*. The occurrence of *BRCA1* mutation in our cohort is 24.1% (34/141) and for *BRCA2*, it is 2.1% (3/141). This concurs well with most of the previous studies in Indian cohorts, which have also reported lower occurrence rates of *BRCA2* mutations as compared with *BRCA1*.<sup>19,21,23,25</sup> Most studies of *BRCA1/2* mutation frequencies in the Asian populations have reported a greater frequency of mutations in *BRCA2* compared with *BRCA1* with the notable exception of studies from India and Pakistan.<sup>37</sup>

The *BRCA1* variant c.68\_69delAG (mostly referred as 185delAG) has been reported in most Indian studies. This mutation, reported as a founder mutation in the Ashkenazi Jewish population, occurs at a very high frequency of 18.0% among BOC families of Ashkenazi Jews and at ~1% among Ashkenazi Jews in the general population.<sup>30,38</sup> In our

cohort, the c.68\_69delAG mutation was detected at a frequency of 4.2% (6/141), which is similar to the frequency range of 0.5–4.1% reported by majority of other Indian studies.<sup>22,23,25</sup> Remarkably, in two South Indian BOC cohorts, the c.68\_69delAG mutation was reported at a frequency of 16.4% (10/61) and 10.9% (10/91), which are similar to the frequency reported in Ashkenazi Jews.<sup>19,26</sup> To determine the frequency of c.68\_69delAG mutation in the patients from South India in our study, we stratified the cohort based upon geographic location. In our cohort, the frequency of c.68\_69delAG mutation in the South Indian patients was 4.8% (3/63). Interestingly, in the South Asian population, the c.68\_69delAG mutation has been reported with allele frequency of 0.006% (Exome Aggregation Consortium database), which is much lower (166 times) than frequency of ~1% reported in Ashkenazi Jewish population.<sup>30</sup> A previous study suggested that the origin of the c.68\_69delAG mutation in Indian population is independent of that of Ashkenazi Jews based on haplotype analysis.<sup>39</sup> Further studies will be helpful in determining the frequency of the c.68\_69delAG mutation and its origin in Indian population.

NGS data has proven useful for simultaneous detection of large deletions/duplications and single nucleotide variants/indels.<sup>31,40</sup> In our study, we detected one large *BRCA1* deletion (encompassing exon 1–2 and some upstream region as well), which accounts for 0.7% of the cohort. Previously, screening of 1781 individuals with indication of breast cancer by a NGS-based 25-gene panel led to detection of large genomic rearrangements in ~1% (18/1781) cases and in *BRCA1/2*, the detection rate was 0.3% (6/1781).<sup>31</sup> Moreover, the contribution of large genomic rearrangements to familial breast cancer in Asia has been considered to be minimal.<sup>37</sup> Our result was comparable to these findings. As per our knowledge, this is the first report of a large deletion in the *BRCA1* gene in an Indian cohort.

One of the key challenges of NGS-based multi-gene testing is the reporting of VUS. This makes recommendation for clinical management complex while also potentially creating anxiety or misunderstanding among patients. We identified 24 VUSs in 21 patients in our cohort; in 2 patients a mutation was also detected along with a VUS. All variants were missense substitution except for one in-frame deletion (p.Lys1110del) in the *BRCA1* gene. Among the 24 VUSs, 4 conservative missense substitutions were predicted damaging by multiple *in silico* prediction tools. However, clinical correlation data or functional evidence for these variants were unavailable; therefore, we labeled them as ‘VUS with probable damaging effect’. Further studies (clinical/functional) will be helpful in ascertaining the role (if any) of these variants in cancer predisposition.

In Indian population the proportion of patients with early onset of breast cancer is much higher when compared with Caucasian populations.<sup>25</sup> This is also true in other Asian populations where patients develop breast cancer at a younger age than their Caucasian counterparts.<sup>37</sup> In our study we stratified the cohort on the basis of age at diagnosis. The mutation detection rate in age groups, <40 years (44.4%) and 40–50 years (53.4%) were much higher as compared with age group >50 years (26.9%). Furthermore, the difference between 40–50 years group and >50 years group was statistically significant ( $P=0.031$ ). Our finding suggests that the patients with an earlier onset of breast cancer (<50 years) are likely have a mutation as compared with those patients with a late onset of the disease (>50 years).

It has been postulated that patients with an early onset of breast cancer are likely to have a *BRCA1/2* mutation.<sup>25</sup> Previously, Saxena *et al.* screened 121 patients’ from Northern India with age of onset  $\leq 40$  years for *BRCA1/2* mutations and detected mutation in 3.3% (4/121). In that study, the mutation rate in 83 patients with age of

onset >40 years was 2.4% (2/83) and overall detection rate was 2.9% (6/204). It should be noted that the study included only 16.7% patients with a positive family history of BOC. Saxena *et al.*<sup>25</sup> concluded that *BRCA1/2* mutations account for a lower proportion of breast cancer cases in Northern India compared with other ethnic populations. Interestingly, in a South Indian cohort screen, Vaidyanathan *et al.*<sup>19</sup> detected a mutation rate of 28% (17/61) in the *BRCA1/2* genes. Furthermore, they mentioned that it is not ideal to include families with early onset of breast cancer for *BRCA1/2* screening in India, as according to various population-based registries in India the average age of breast cancer patients have been reported to be 50–53 years. Hence, they argued that including patients with early-onset breast cancer for *BRCA1/2* screening may lead to an underestimation of the frequency of *BRCA1/2* mutations in Indian cohort.<sup>19</sup> However, in our study the frequency of *BRCA1/2* mutations in early-onset (<40 years) breast cancer was 44.4% (16/36), which was much higher than the detection rate reported in these two studies.<sup>19,25</sup> In contrast to previous studies, we did not observe any significant difference in the mutation detection rate between the patients from North and South India. In our cohort, the proportion of patients referred from North India was 39.7% (56/141), with a mutation detection rate of 35.7% (20/56) as compared with 44.6% (63/141) patients referred from South India, with a mutation detection rate of 28.6% (18/63). An increased mutation detection rate (35.7%) in the North Indian patients in our study as compared with previous report of mutation detection rate of 2.9%<sup>25</sup> could be due to increased sensitivity of mutation detection by NGS-based multi-gene panel. Also, the proportion of familial cases (47/56) among North Indian patients in our study was significantly higher (83.9%) as compared with Saxena *et al.*<sup>25</sup> study where familial cases contributed only 16.7% (34/204); this could also be attributed to higher mutation detection rate in our study. Interestingly, in an earlier study, Saxena *et al.* reported a mutation detection rate of 10% (2/20) albeit in a smaller cohort of 20 patients with familial BOC or early age of onset (<35 years) from North India.<sup>41</sup> In another study, a mutation detection rate of 12.5% (2/16) has been reported in 16 North Indian familial breast cancer cases.<sup>23</sup>

According to the latest NCRP (National Cancer Registry Programme) report, the annual percentage change in the incidence rate of breast cancer recorded in period: 2000–2009 in some regions of India has nearly doubled when compared with the rate recorded between 1982 and 1999.<sup>42</sup> Furthermore in India, the overall 5-year relative survival rate is 60.4%, which is relatively poorer compared with that reported in the developed nations (USA and European Union), where survival rates are >80%.<sup>43</sup> Consequently, there is an increased financial burden on the Indian health-care system and a pressing need for a cost-effective and comprehensive genetic testing method for diagnosis and early detection of BOC. Our study suggests that the NGS-based multi-gene testing increases the sensitivity of mutation detection and helps in identifying patients with a high risk of developing cancer as compared with sequential tests for single-genes performed in reflex-testing mode.

In the current study, we report a high prevalence rate (36.2%) of mutations in an Indian cohort. The rate is even higher (52%) when only familial cases were considered. Data from our study strongly suggests the incorporation of NGS-based multi-gene panels for routine genetic testing in India. It will potentially be helpful in developing effective strategies for early detection, prevention and better management of HBOC, which will have a positive impact on healthcare in India.

## CONFLICT OF INTEREST

All authors except MG are employees of Strand Life Sciences, a for-profit molecular diagnostic service provider. MG is an employee of Triesta Sciences, a for-profit diagnostic service provider.

## ACKNOWLEDGEMENTS

We thank the patients and families, who consented to participate in this study. We also thank all the physicians, who referred the patients to our center. We also thank the Strand Life Sciences laboratory, bioinformatics, interpretation and genomic counseling staffs for providing the infrastructure needed for this study.

- 1 Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J. & Jemal, A. Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87–108 (2015).
- 2 Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M. *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, E359–E386 (2015).
- 3 Jemal, A., Center, M. M., DeSantis, C. & Ward, E. M. Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol. Biomarkers Prev.* **19**, 1893–1907 (2010).
- 4 Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* **266**, 66–71 (1994).
- 5 Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J. *et al.* Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* **378**, 789–792 (1995).
- 6 Meindl, A., Ditsch, N., Kast, K., Rhiem, K. & Schmutzler, R. K. Hereditary breast and ovarian cancer: new genes, new treatments, new concepts. *Dtsch. Arztebl. Int.* **108**, 323–330 (2011).
- 7 Becker, K. F., Atkinson, M. J., Reich, U., Becker, I., Nekarda, H., Siewert, J. R. *et al.* E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer Res.* **54**, 3845–3852 (1994).
- 8 Liaw, D., Marsh, D. J., Li, J., Dahia, P. L., Wang, S. I., Zheng, Z. *et al.* Germline mutations of the *PTEN* gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat. Genet.* **16**, 64–67 (1997).
- 9 Hemminki, A., Markie, D., Tomlinson, I., Avizienyte, E., Roth, S., Loukola, A. *et al.* A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* **391**, 184–187 (1998).
- 10 Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Nelson, C. E., Kim, D. H. *et al.* Germ line *p53* mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**, 1233–1238 (1990).
- 11 Broeks, A., Urbanus, J. H., Floore, A. N., Dahler, E. C., Klijn, J. G., Rutgers, E. J. *et al.* ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. *Am. J. Hum. Genet.* **66**, 494–500 (2000).
- 12 Cantor, S. B., Bell, D. W., Ganesan, S., Kass, E. M., Drapkin, R., Grossman, S. *et al.* BACH1, a novel helicase-like protein, interacts directly with *BRCA1* and contributes to its DNA repair function. *Cell* **105**, 149–160 (2001).
- 13 Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., Wasielewski, M., de Snoo, A., Oldenburg, R. *et al.* Low-penetrance susceptibility to breast cancer due to *CHEK2*(\*) 1100delC in noncarriers of *BRCA1* or *BRCA2* mutations. *Nat. Genet.* **31**, 55–59 (2002).
- 14 Bogdanova, N., Feshchenko, S., Schürmann, P., Waltes, R., Wieland, B., Hillemanns, P. *et al.* Nijmegen Breakage Syndrome mutations and risk of breast cancer. *Int. J. Cancer* **122**, 802–806 (2008).
- 15 Heikkinen, K., Rapakko, K., Karppinen, S.-M., Erkkö, H., Knuutila, S., Lundán, T. *et al.* *RAD50* and *NBS1* are breast cancer susceptibility genes associated with genomic instability. *Carcinogenesis* **27**, 1593–1599 (2006).
- 16 Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A. *et al.* *PALB2*, which encodes a *BRCA2*-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* **39**, 165–167 (2007).
- 17 Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D. *et al.* Germline mutations in breast and ovarian cancer pedigrees establish *RAD51C* as a human cancer susceptibility gene. *Nat. Genet.* **42**, 410–414 (2010).
- 18 Rodríguez-López, R., Osorio, A., Ribas, G., Pollán, M., Sánchez-Pulido, L., de la Hoya, M. *et al.* The variant E233G of the *RAD51D* gene could be a low-penetrance allele in high-risk breast cancer families without *BRCA1/2* mutations. *Int. J. Cancer* **110**, 845–849 (2004).
- 19 Vaidyanathan, K., Lakhotia, S., Ravishankar, H. M., Tabassum, U., Mukherjee, G. & Somasundaram, K. *BRCA1* and *BRCA2* germline mutation analysis among Indian women from south India: Identification of four novel mutations and high-frequency occurrence of 185delAG mutation. *J. Biosci.* **34**, 415–422 (2009).
- 20 Valarmathi, M. T., Sawhney, M., Deo, S. S. V., Shukla, N. K. & Das, S. N. Novel germline mutations in the *BRCA1* and *BRCA2* genes in Indian breast and breast-ovarian cancer families. *Hum. Mutat.* **23**, 205 (2004).

- 21 Rajkumar, T., Soumitra, N., Nancy, N. K., Swaminathan, R., Sridevi, V. & Shanta, V. BRCA1, BRCA2 and CHEK2 (1100 del C) germline mutations in hereditary breast and ovarian cancer families in South India. *Asian Pac. J. Cancer Prev.* **4**, 203–208 (2003).
- 22 Hansa, J., Kannan, R. & Ghosh, S. K. Screening of 185DelAG, 1014DelGT and 3889DelAG BRCA1 mutations in breast cancer patients from North-East India. *Asian Pac. J. Cancer Prev.* **13**, 5871–5874 (2012).
- 23 Hedau, S., Jain, N., Husain, S. A., Mandal, A. K., Ray, G., Shahid, M. *et al.* Novel germline mutations in breast cancer susceptibility genes BRCA1, BRCA2 and p53 gene in breast cancer patients from India. *Breast Cancer Res. Treat.* **88**, 177–186 (2004).
- 24 Kumar, B. V., Lakhotia, S., Ankathil, R., Madhavan, J., Jayaprakash, P. G., Nair, M. K. *et al.* Germline BRCA1 mutation analysis in Indian breast/ovarian cancer families. *Cancer Biol. Ther.* **1**, 18–21 (2002).
- 25 Saxena, S., Chakraborty, A., Kaushal, M., Kotwal, S., Bhatnager, D., Mohil, R. S. *et al.* Contribution of germline BRCA1 and BRCA2 sequence alterations to breast cancer in Northern India. *BMC Med. Genet.* **7**, 75 (2006).
- 26 Rajkumar, T., Meenakumari, B., Mani, S., Sridevi, V. & Sundersingh, S. Targeted Resequencing of 30 Genes Improves the Detection of Deleterious Mutations in South Indian Women with Breast and/or Ovarian Cancers. *Asian Pac. J. Cancer Prev.* **16**, 5211–5217 (2015).
- 27 Richards, C. S., Bale, S., Bellissimo, D. B., Das, S., Grody, W. W., Hegde, M. R. *et al.* ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. *Genet. Med.* **10**, 294–300 (2008).
- 28 Chan, K. C. A., Jiang, P., Chan, C. W. M., Sun, K., Wong, J., Hui, E. P. *et al.* Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc. Natl Acad. Sci. USA* **110**, 18761–18768 (2013).
- 29 Ihaka, R. & Gentleman, R. R. A language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299–314 (1996).
- 30 Struewing, J. P., Abeliovich, D., Peretz, T., Avishai, N., Kaback, M. M., Collins, F. S. *et al.* The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nat. Genet.* **11**, 198–200 (1995).
- 31 Tung, N., Battelli, C., Allen, B., Kaldete, R., Bhatnagar, S., Bowles, K. *et al.* Frequency of mutations in individuals with breast cancer referred for BRCA1 and BRCA2 testing using next-generation sequencing with a 25-gene panel. *Cancer* **121**, 25–33 (2015).
- 32 Lincoln, S. E., Kobayashi, Y., Anderson, M. J., Yang, S., Desmond, A. J., Mills, M. A. *et al.* A systematic comparison of traditional and multigene panel testing for hereditary breast and ovarian cancer genes in more than 1000 Patients. *J. Mol. Diagn.* **17**, 533–544 (2015).
- 33 Castéra, L., Krieger, S., Rousselin, A., Legros, A., Baumann, J.-J., Bruet, O. *et al.* Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. *Eur. J. Hum. Genet.* **22**, 1305–1313 (2014).
- 34 Minion, L. E., Dolinsky, J. S., Chase, D. M., Dunlop, C. L., Chao, E. C. & Monk, B. J. Hereditary predisposition to ovarian cancer, looking beyond BRCA1/BRCA2. *Gynecol. Oncol.* **137**, 86–92 (2015).
- 35 Kurian, A. W., Hare, E. E., Mills, M. A., Kingham, K. E., McPherson, L., Whittemore, A. S. *et al.* Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. *J. Clin. Oncol.* **32**, 2001–2009 (2014).
- 36 Desmond, A., Kurian, A. W., Gabree, M., Mills, M. A., Anderson, M. J., Kobayashi, Y. *et al.* Clinical actionability of multigene panel testing for hereditary breast and ovarian cancer risk assessment. *JAMA Oncol.* **02114**, 1–9 (2015).
- 37 Kim, H. & Choi, D. H. Distribution of BRCA1 and BRCA2 mutations in asian patients with breast cancer. *J. Breast Cancer* **16**, 357–365 (2013).
- 38 Karami, F. & Mehdipour, P. A comprehensive focus on global spectrum of BRCA1 and BRCA2 mutations in breast cancer. *Biomed Res. Int.* **2013**, 928562 (2013).
- 39 Kadalmani, K., Deepa, S., Bagavathi, S., Anishetty, S., Thangaraj, K. & Gajalakshmi, P. Independent origin of 185delAG BRCA1 mutation in an Indian family. *Neoplasma* **54**, 51–56 (2007).
- 40 Feng, Y., Chen, D., Wang, G.-L., Zhang, V. W. & Wong, L.-J. C. Improved molecular diagnosis by the detection of exonic deletions with target gene capture and deep sequencing. *Genet. Med.* **17**, 99–107 (2015).
- 41 Saxena, S., Szabo, C. I., Chopin, S., Barjhoux, L., Sinilnikova, O., Lenoir, G. *et al.* BRCA1 and BRCA2 in Indian breast cancer patients. *Hum. Mutat.* **20**, 473–474 (2002).
- 42 NCRP (National Cancer Registry Programme). *Time Trends in Cancer Incidence Rates 1982–2010*. (Indian Council of Medical Research, Bangalore, 2013).
- 43 Allemani, C., Weir, H. K., Carreira, H., Harewood, R., Spika, D., Wang, X.-S. *et al.* Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet* **385**, 977–1010 (2014).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)