Supplementary Material

A) Test Methodology:

Test Methodology

Targeted gene sequencing: Selective capture and sequencing of the protein coding regions of the ?genome/?genes are performed. Mutations identified in the exonic regions are generally actionable compared with variations that occur in noncoding regions. Targeted sequencing represents a cost-effective approach to detect variants present in ?multiple/?large genes in an individual.

DNA extracted from blood was used to perform targeted gene capture using a custom capture kit. The libraries were sequenced to mean >80–100X coverage on Illumina sequencing platform.

The classification of the variations is done based on American College of Medical Genetics as described below.[1]

Pathogenic

(i) 1 Very strong (PVS1) AND (a) ≥1 Strong (PS1–PS4) OR (b) ≥2 Moderate (PM1–PM6) OR (c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) OR (d) ≥2 Supporting (PP1–PP5)
(ii) ≥2 Strong (PS1–PS4) OR
(iii) 1 Strong (PS1–PS4) AND (a) ≥ 3 Moderate (PM1–PM6) OR (b)2 Moderate (PM1–PM6) AND ≥2 Supporting (PP1–PP5) OR (c)1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)

Likely Pathogenic

1. (i) 1Very strong (PVS1) AND 1 moderate (PM1– PM6) OR
2. (ii) 1 Strong (PS1–PS4) AND 1–2 moderate (PM1–PM6) OR
3. (iii) 1 Strong (PS1–PS4) AND ≥2 supporting (PP1–PP5) OR
4. (iv) ≥3 Moderate (PM1–PM6) OR
5. (v) 2 Moderate (PM1–PM6) AND ≥2 supporting (PP1–PP5) OR
6. (vi) 1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)

Benign

1. (i) 1 Stand-alone (BA1) OR
2. (ii) ≥2 Strong (BS1–BS4)

Likely Benign

1. (i) 1 Strong (BS1–BS4) and 1 supporting (BP1– BP7) OR
2. (ii) ≥2 Supporting (BP1–BP7)

Uncertain Significance

1. (i) Other criteria shown above are not met OR
2. (ii) The criteria for benign and pathogenic are contradictory

We followed the GATK best practices framework for the identification of variants in the sample using Sentieon (v201808.07).[2] The sequences obtained are aligned to human reference genome (GRCh38.p13) using Sentieon aligner[2,3] and analyzed using Sentieon for removing duplicates, recalibration, and realignment of indels.[2] Sentieon haplotype caller has been used to identify variants that are relevant to the clinical indication. Gene annotation of the variants is performed using VEP program[4] against the Ensembl release 99 human gene model.[5] In addition to SNVs and small Indels, copy number variants (CNVs) are detected from targeted sequence data using the ExomeDepth (v1.1.10) method.[6] This algorithm detects rare CNVs based on comparison of the read-depths of the test data with the matched aggregate reference dataset.

Clinically relevant mutations were annotated using published variants in literature and a set of diseases databases—ClinVar, OMIM (updated on 11th May?2020), GWAS, HGMD (v2020.2), and SwissVar.[7–11] Common variants are filtered based on allele frequency in 1000Genome Phase 3, gnomAD (v2.1), EVS, dbSNP (v151), 1000 Japanese Genome, and our internal Indian population database.[12–16] Nonsynonymous variants effect is calculated using multiple algorithms such as PolyPhen-2, SIFT, MutationTaster2, and LRT. Only nonsynonymous and splice site variants found in the hereditary cancer panel-focused genes were used for clinical interpretation. Silent variations that do not result in any change in amino acid in the coding region are not reported.

1. #The transcript used for clinical reporting generally represents the canonical transcript, which is usually the longest coding transcript with ?strong/?multiple supporting evidence. However, clinically relevant variants annotated in alternate complete coding transcripts could also be reported.
2. #The in-silico predictions are based on Variant Effect Predictor, Ensembl release 99 (SIFT version—5.2.2; PolyPhen—2.2.2), dbNSFPv4.0 (LRT version—December?5, 2019) and MutationTaster2 (MT2). MutationTaster2 predictions are based on ?NCBI/?Ensembl 66 build (GRCh38 genomic coordinates are converted to hg19 using UCSC LiftOver and mapped to MT2).

Appendix Coverage of hereditary cancer panel -Focus ED genes

	Percentage of coding region covered	Gene	Percentage of coding region covered	Gene	Percentage of coding region covered
AIP	100.00	*ALK*	100.00	*APC*	100.00
APC	100.00	*AR*	100.00	*AR*	100.00
ATM	100.00	*BAP1*	100.00	*BARD1*	100.00
BLM	100.00	*BMPR1A*	100.00	*BRCA1*	100.00
BRCA2	100.00	*BRCA2*	100.00	*BRIP1*	100.00
BUB1B	100.00	*CD82*	100.00	*CDC73*	100.00
CDH1	100.00	*CDK4*	100.00	*CDKN1C*	100.00
CDKN2A	100.00	*CEBPA*	100.00	*CEP57*	100.00
CHEK2	100.00	*CYLD*	100.00	*CYLD*	100.00
DDB2	100.00	*DICER1*	100.00	*DIS3L2*	100.00
EGFR	100.00	*ELAC2*	100.00	*ENG*	100.00
Gene	**Percentage of coding region covered**	**Gene**	**Percentage of coding region covered**	**Gene**	**Percentage of coding region covered**
EPCAM	100.00	*ERCC2*	100.00	*ERCC3*	100.00
ERCC4	100.00	*ERCC5*	100.00	*EXT1*	100.00
EXT2	100.00	*EZH2*	100.00	*FANCA*	98.65
FANCB	100.00	*FANCC*	100.00	*FANCC*	100.00
FANCD2	99.91	*FANCD2*	99.91	*FANCE*	100.00
FANCE	100.00	*FANCF*	100.00	*FANCG*	100.00
FANCI	100.00	*FANCL*	100.00	*FANCM*	100.00
FH	100.00	*FLCN*	100.00	*FLCN*	100.00
GATA2	100.00	*GPC3*	100.00	*HRAS*	100.00
KIT	100.00	*MAG*	100.00	*MAX*	100.00
MAX	100.00	*MEN1*	100.00	*MEN1*	100.00
MET	100.00	*MLH1*	99.96	*MLH1*	99.96
MLH3	100.00	*MRE11*	100.00	*MSH2*	100.00
MSH3	100.00	*MSH3*	100.00	*MSH6*	100.00
MSR1	100.00	*MUTYH*	99.82	*MUTYH*	99.82
MXI1	100.00	*NBN*	100.00	*NF1*	100.00
NF2	94.30	*NF2*	94.30	*NSD1*	100.00
PALB2	100.00	*PHOX2B*	100.00	*PMS1*	100.00
PMS2	100.00	*PRF1*	100.00	*PRKAR1A*	100.00
PTCH1	100.00	*PTEN*	100.00	*PTEN*	100.00
RAD50	100.00	*RAD51C*	100.00	*RAD51D*	100.00
RAD54L	100.00	*RAD54L*	100.00	*RB1*	100.00
RECQL4	100.00	*RET*	100.00	*RHBDF2*	100.00
RNASEL	100.00	*RUNX1*	100.00	*SBDS*	100.00
SDHAF2	100.00	*SDHB*	100.00	*SDHC*	94.13
SDHD	82.44	*SDHD*	82.44	*SLX4*	100.00
SMAD4	100.00	*SMARCB1*	100.00	*STK11*	100.00
SUFU	100.00	*TGFBR2*	100.00	*TMEM127*	100.00
TP53	95.34	*TSC1*	100.00	*TSC2*	100.00
TSC2	100.00	*VHL*	100.00	*VHL*	100.00
WRN	100.00	*WT1*	100.00	*XPA*	100.00
XPA	100.00	*XPC*	100.00	*XPC*	100.00

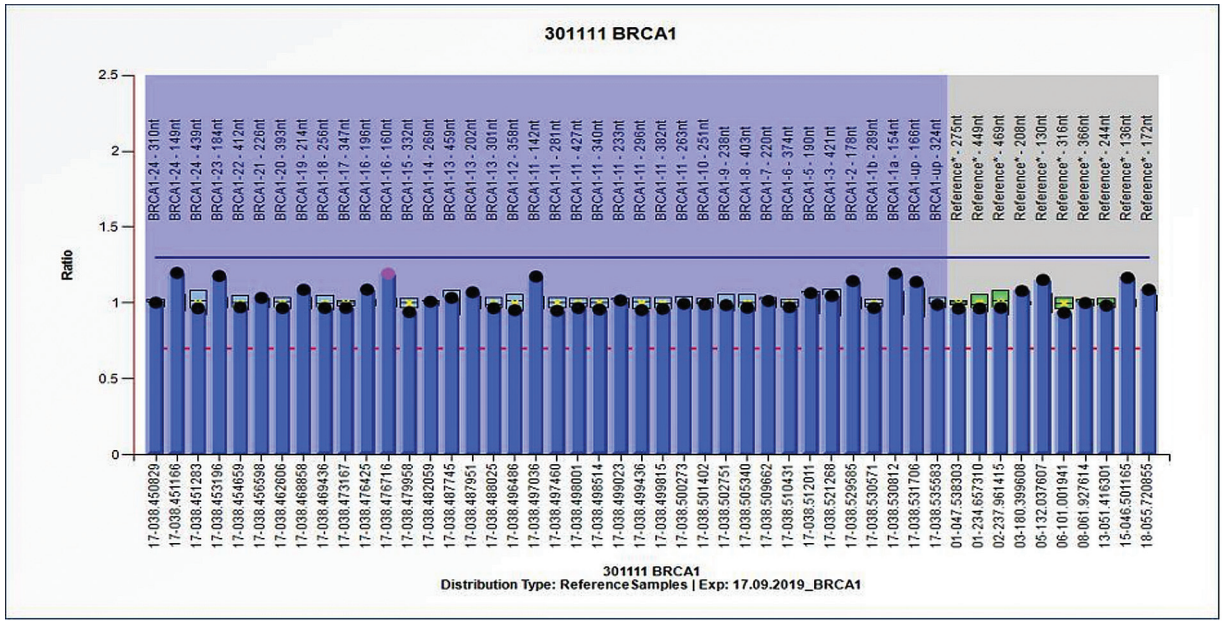
MLPA Methodology

Copy number changes in 24 exons of the BRCA1 and 27 exons of BRCA2 gene were identified by hybridizing with multiplex ligation-dependent probe amplification (MLPA)-based assay. Each MLPA probe consists of two hemiprobes that bind to adjacent sites on the target sequence. Upon ligation and subsequent polymerase chain reaction (PCR) amplification, each distinct MLPA probe (specific to distinct target regions) generates an amplicon with a unique length that is separated and quantified by capillary electrophoresis. Heterozygous deletions within target sequences will prevent efficient probe binding and give a 35 to 50% reduced relative peak area of the amplification product specific to that probe set. Copy number differences of various exons between test and control DNA samples can be detected by analyzing the MLPA peak patterns. The MLPA test will not detect the point mutations in the BRCA1 and BRCA2 genes, which accounts for most of genetic defects. It is, therefore, recommended to use MLPA in combination with sequence analysis.

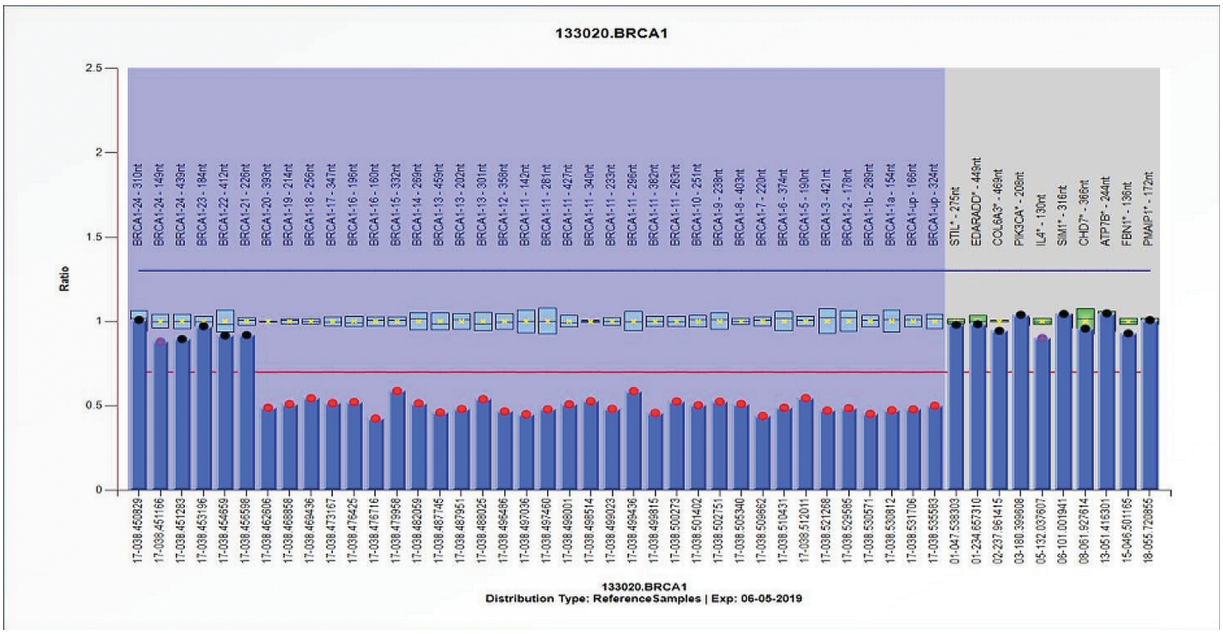
- Sequence changes (e.?g., SNPs, point mutations, small indels) in the target sequence detected by a probe can cause false positive results. ?Mutations/?SNPs (even when >20 nt from the probe ligation site) can reduce the probe signal by preventing ligation of the probe oligonucleotides or by destabilizing the binding of a probe oligonucleotide to the sample DNA. Therefore, single exon deletions detected by MLPA should always be confirmed by other methods like multiplex PCR or sequencing.
- MLPA cannot detect any changes that lie outside the target sequence of the probes and will not detect most inversions or translocations. Even when MLPA did not detect any aberrations, the possibility remains that biological changes in that gene or chromosomal region do exist but remain undetected.

MLPA BRCA 1 Ratio Chart

Reference BRCA 1 Gene



Deletion in BRCA 1 Gene



B) VUS Distribution in the Study

VUS identified	Number of patients	Mutation type
BRCA 2	02	Missense
ATM	03	Missense
FANCI	01	Missense
MLH 1	01	Missense
CHECK2	01	Missense
FANCM	01	Missense
ERCC2	01	Missense

Funding
Not applicable.

References

1 Richards S, Aziz N, Bale S, et?al;ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17(05): 405–424
2 Freed D, et?al. The Sentieon Genomics Tools—a fast and accurate solution to variant calling from next-generation sequence data. bioRxiv115717.
3 Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010;26(05):589–595
4 McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 2010;26(16):2069–2070
5 Zerbino DR, Achuthan P, Akanni W, et?al. Ensembl 2018. Nucleic Acids Res 2018;46(D1)D754–D761
6 Plagnol V, Curtis J, Epstein M, et?al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 2012;28(21):2747–2754
7 Landrum MJ, Lee JM, Benson M, et?al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 2016;44(D1):D862–D868
8 McKusick VA. Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders. Baltimore: Johns Hopkins University Press (12th edition), 1998
9 Welter D, MacArthur J, Morales J, et?al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42Database issueD1001–D1006
10 Stenson PD, Mort M, Ball EV, et?al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet 2017;136(06): 665–677
11 Mottaz A, David FP, Veuthey AL, Yip YL. Easy retrieval of single amino-acid polymorphisms and phenotype information using SwissVar. Bioinformatics 2010;26(06):851–852
12 Auton A, Brooks LD, Durbin RM, et?al;1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526(7571):68–74
13 Karczewski KJ, Francioli LC, Tiao G, et?al;Genome Aggregation Database Consortium. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581 (7809):434–443
14 Exome Variant Server. NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.?edu/?EVS/
15 Nagasaki M, Yasuda J, Katsuoka F, et?al;ToMMo Japanese Reference Panel Project. Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. Nat Commun 2015; 6:8018
16 Sherry ST, Ward MH, Kholodov M, et?al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001;29(01):308–311

