SureSeq

Simultaneous Detection of Genetic and Copy-Number Variations in *BRCA1/2* Genes

Lyudmila Georgieva, Stephanie Carpenter, Natalie Milner, Venu Pullabhatla, James Reid, Kay Wiebrands, Aysel Heckel, Graham Speight Oxford Gene Technology (OGT).

A Sysmex Group Company

What binds us, makes us.

Introduction

Loss-of-function mutations in tumour suppressor genes *BRCA1* and *BRCA2* have been implicated in an increased risk for breast and ovarian cancer[1,2]. Screening for germline mutations in these genes allows research into familial risk of developing breast and ovarian cancer. In addition, assessment of somatic mutations in tumour samples can help research into tumour development, drug response and the development of new therapies.

A wide range of genetic variations are associated with breast and ovarian cancer, including single nucleotide variants (SNVs), small insertions/deletions (indels) and copy-number variations (CNVs). For more than a decade, the gold standard for mutational screening has been Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA), imposing significant time and cost burden.

Advances in next-generation sequencing (NGS) now allows for the reliable detection of CNVs in addition to SNVs/indels in a single assay.

In this study, we tested the capability of the SureSeq™ Breast Cancer + CNV Panel to overcome the challenges currently experienced and provide a possible future single assay to be developed for breast and ovarian cancer.

Methods

The SureSeq hybridisation-based approach was used throughout this study; the workflow of this is outlined in Figure 1.

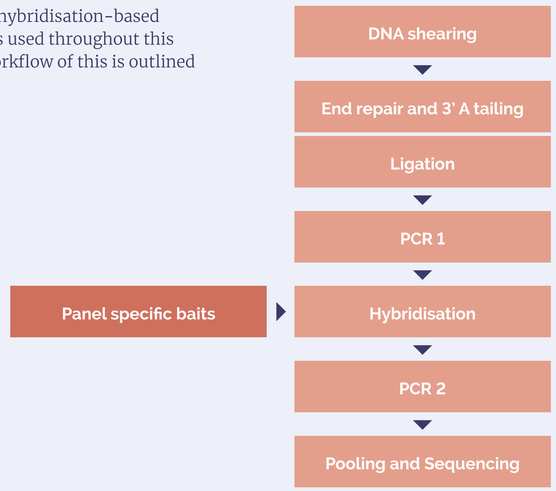


Figure 1: OGT SureSeq workflow. The SureSeq workflow allows users to go from extracted DNA to sequencer in 1.5 days with minimal handling time.

The SureSeq Breast Cancer + CNV Panel can be used for detection of SNV/indels and CNVs in 7 genes.

ATM	*BRCA1*	*BRCA2*	*TP53*	*CHEK2*	*PALB2*	*TP53*

A number of different sample types were profiled:

1. Set 1 includes 16 whole blood extracted samples with known germline CNVs. These samples were also used in spike-in experiments to generated 'mosaic' samples and to test our workflow and software.
2. Set 2 includes 20 FFPE derived tumour samples.
3. Set 3 includes 24 samples (Coriell, Camden, NJ) with no known CNVs in the regions of interest and isused as reference.

The resulting libraries were sequenced using a 2x150 bp read length (v2) protocol on an Illumina MiSeq®.

CNV detection concordance was assessed by comparing NGS calls to events reported by an orthogonal technology, MLPA for set 1 samples and CytoSure® Cancer +SNP Arrays and CytoSure Interpret Software (Oxford Gene Technology) for set 2 samples.

Bioinformatics Analysis

Data sequencing analysis including CNV detection was performed using Interpret, OGT's complimentary NGS analysis software.

Results I

Confident detection of SNV and indels in 7 genes

Facilitated by OGT's expert bait design, the hybridisation-based SureSeq Breast Cancer + CNV Panel delivers excellent coverage uniformity, allowing consistent detection of SNVs and indels in germline samples, as well as somatic analysis on FPPE tissues down to 1% VAF (Figures 2 and 3).

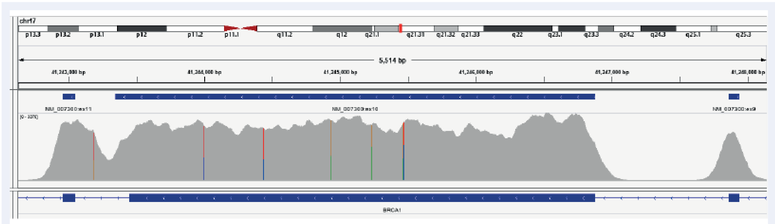


Figure 2: Illustration of the excellent coverage uniformity of *BRCA1* exons 9, 10 and 11 in FFPE samples. Depth of coverage per base (grey). Gene coding region as defined by RefSeq (bottom track).

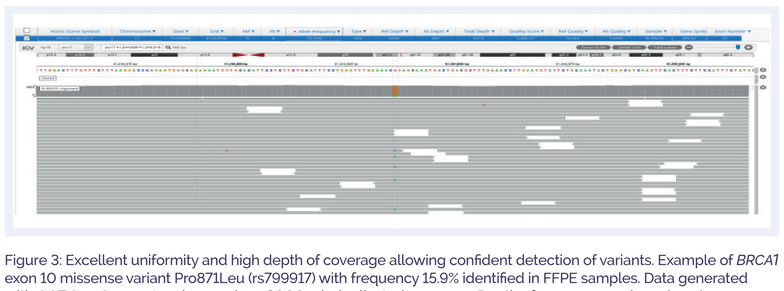


Figure 3: Excellent uniformity and high depth of coverage allowing confident detection of variants. Example of *BRCA1* exon 10 missense variant Pro871Leu (rs799917) with frequency 15.9% identified in FFPE samples. Data generated with OGT SureSeq protocol averaging ~2000x deduplicated coverage. Depth of coverage per base (grey).

Results II

Confident detection of Copy Number Variations

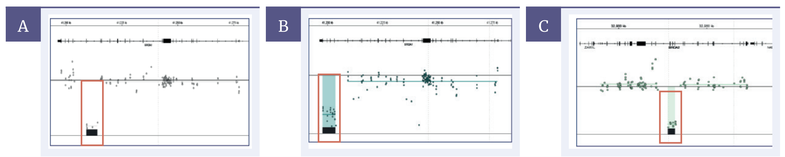
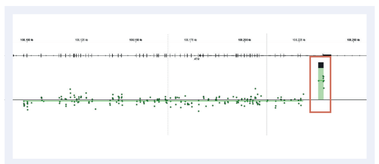


Figure 4: Detection of germline CNVs in *BRCA1* and *BRCA2*. A ~Deletion of ex20 *BRCA1* (4.99kb), B deletion of ex21-23 *BRCA1* (5.93kb). C Deletion of ex14-17 *BRCA2* (4.21kb).

NGS



Array

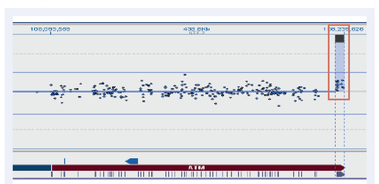
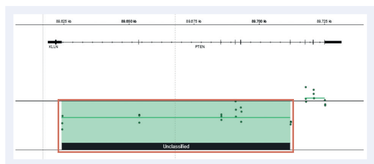


Figure 5: Somatic *ATM* duplication of exon 62-63, fully concordant with array data.

NGS



Array

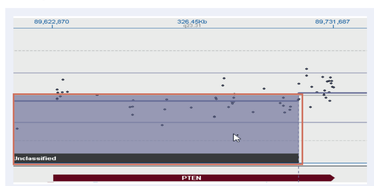


Figure 6: Somatic *PTEN* deletion of exon 1-5, fully concordant with array data.

Results III

Validation of the SureSeq Breast Cancer + CNV Panel and enhanced CNV detection software with 16 samples with germline CNVs

- CNVs of various sizes are detected in genes *BRCA1*, *BRCA2* and *TP53* (Table 1, Figure 4).
- We observe very high concordance between NGS and MLPA.
- Some minor discrepancies in exon numbering were observed, especially for *BRCA1* where multiple transcripts were present. These discrepancies were resolved by analysis with the different transcripts.
- In the dilution samples generated from the sample carriers of germline CNV we detected CNVs as mosaics down to 30% frequency.

Sample	Expected MPLA			NGS								
	Gene	Type	Region	Chr	Start	End	Length	Gene	Type	Region	Log2	# Markers
1	*BRCA1*	Deletion	ex 21-23	17	41,197,681	41,203,615	5.93Kb	*BRCA1*	Deletion	ex 21-24	-1.02	18
2	*TP53*	Duplication	ex 2-6	17	7,578,157	7,581,148	2.99Kb	*TP53*	Duplication	ex 2-6	0.54	17
3	*BRCA2*	Deletion	ex 1-2	13	32,889,610	32,890,705	1.1Kb	*BRCA2*	Deletion	ex 1-2	-0.83	4
4	*BRCA1*	Duplication	ex 5-7	17	41,256,133	41,258,586	2.45Kb	*BRCA1*	Duplication	ex 4-6	0.58	9
5	*BRCA2*	Deletion	ex 14-17	13	32,927,868	32,932,076	4.21Kb	*BRCA2*	Deletion	ex 14-16	-1.00	12
6	*BRCA2*	Deletion	ex 25-27	13	32,965,284	32,972,916	7.63Kb	*BRCA2*	Deletion	ex 25-27	-0.96	14
7	*BRCA1*	Deletion	ex 14-20	17	41,209,035	41,228,642	19.61Kb	*BRCA1*	Deletion	ex 14-20	-0.96	25
8	*BRCA1*	Deletion	ex 16-17	17	41,217,848	41,223,264	5.42Kb	*BRCA1*	Deletion	ex 16-17	-0.99	8
9	*BRCA1*	Deletion	ex 3	17	41,267,694	41,267,844	150b	*BRCA1*	Deletion	ex 3	-0.99	3
10	*BRCA1*	Deletion	ex 2-24	17	41,197,681	41,276,372	78.69Kb	*BRCA1*	Deletion	ex 2-24	-0.88	136
11	*BRCA1*	Deletion	ex 13-15	17	41,226,337	41,235,823	9.49Kb	*BRCA1*	Deletion	ex 12-15	-0.89	13
12	*BRCA1*	Deletion	ex 20	17	41,209,035	41,214,020	4.99Kb	*BRCA1*	Deletion	ex 20	-0.97	5
13	*BRCA1*	Deletion	ex 1-7	17	41,256,133	41,277,275	21.14Kb	*BRCA1*	Deletion	ex 2-6	-0.98	19
14	*BRCA1*	Deletion	ex 1-2	17	41,275,998	41,277,275	1.28Kb	*BRCA1*	Deletion	ex 1-2	-1.08	7
15	*BRCA2*	Deletion	ex 1-27	13	32,889,610	32,972,916	83.31Kb	*BRCA2*	Deletion	ex 1-27	-0.98	120
16	*BRCA1*	Deletion	ex 5	17	41,257,365	41,267,377	10.0kb	*BRCA1*	Deletion	ex 4	-0.98	3

Table 1: Data generated with the SureSeq Breast Cancer + CNV Panel using a combination of OGT workflow and Interpret software showed very high concordance with alternative technology (MLPA). NGS data annotations are is based on NM_007300 (*BRCA1*), NM_000546 (*TP53*) and NM_000059 (*BRCA2*)

Validation of the SureSeq Breast Cancer + CNV Panel and enhanced CNV detection software with 20 FFPE samples

- Data presented here are from 20 FFPE samples that were processed using the OGT workflow in combination with Interpret, OGT's gene variant and CNV detection software.
- We have detected CNVs in 7 samples (Table 2, Figures 5 and 6). The range of CNVs detected include mostly whole gene events.
- CNV events were reported at predicted tumour content as low as 30%.

Sample	Expected NGS								Array					
	Chr	Start	End	Length	Type	Gene	Region	Mean Log2	Chr	Start	End	Type	Length	Mean Log2
1	11	108,226,829	108,236,245	9.42 kb	Duplication	*ATM*	ex 62-63	0.19	11	108,235,454	108,239,116	Duplication	3.66 kb	0.19
2	22	29,083,854	29,130,725	46.87Kb	Duplication	*CHEK2*	whole gene	0.33	22	29,083,811	29,106,617	Duplication	22.80kb	0.20
3	17	41,197,731	41,258,486	60.76Kb	Duplication	*BRCA1*	ex 5-24	0.23	17	38,545,009	41,251,888	Duplication	2.70 Mb*	0.23
3	22	29,083,854	29,130,725	46.87Kb	Deletion	*CHEK2*	whole gene	-0.53	22	16,132,087	51,244,549	Deletion	35.11 Mb*	-0.48
3	13	32,889,610	32,972,916	83.31Kb	Deletion	*BRCA2*	whole gene	-0.45	13	19,029,882	94,612,344	Deletion	75.58 Mb*	-0.39
3	16	23,614,769	23,652,529	37.76Kb	Duplication	*PALB2*	ex 3-13	0.19	16	15,829,063	27,440,802	Duplication	11.61 Mb*	0.20
4	13	32,889,610	32,972,916	83.31Kb	Duplication	*BRCA2*	whole gene	0.33	13	19,029,882	32,970,045	Duplication	13.94 Mb*	0.24
5	13	32,889,610	32,972,916	83.31Kb	Duplication	*BRCA2*	whole gene	0.26	13	20,532,949	100,211,297	Duplication	79.67 Mb*	0.17
6	17	7,572,892	7,583,037	10.15Kb	Deletion	*TP53*	whole gene	-0.22	17	5,287,037	7,996,894	Deletion	2.70 Mb*	-0.26
6	10	89,624,190	89,712,020	87.83Kb	Deletion	*PTEN*	ex 1-5	-0.20	10	76,790,493	89,717,633	Deletion	12.92 Mb*	-0.18
6	11	108,098,312	108,114,855	16.54Kb	Duplication	*ATM*	whole gene	0.22	11	108,093,227	108,116,088	Duplication	22.86 kb*	0.33
7	16	23,614,769	23,652,529	37.76Kb	Duplication	*PALB2*	ex 3-13	0.24	16	15,820,851	33,841,508	Duplication	18.02 Mb*	0.29
8	11	108,098,312	108,236,245	137.93Kb	Deletion	*ATM*	whole gene	-0.45	11	108,094,339	111,298,222	Deletion	3.20 Mb*	-0.45
8	22	29,117,528	29,130,725	13.2Kb	Deletion	*CHEK2*	ex 2-5	-0.27	22	29,117,566	29,123,982	Deletion	6.41 kb	-0.31
8	13	32,889,610	32,972,916	83.31Kb	Duplication	*BRCA2*	whole gene	0.35	13	27,832,703	33,697,771	Duplication	5.86 Mb*	0.29

*Part of a big event overlapping with the gene.

Table 2: Data generated with the SureSeq Breast Cancer + CNV Panel using a combination of OGT workflow and enhanced CNV detection software shows very high concordant with alternative technology (OGT array).

Conclusions

- Superior uniformity of coverage from a hybridisation-based enrichment using the SureSeq Breast Cancer + CNV panel allowed simultaneous detection of SNVs, indels as well as larger structural alterations in a single assay.
- We have demonstrated the capability of a SureSeq Breast Cancer + CNV panel in combination with Interpret software to detect germline and mosaic CNVs, ranging from a single exon to whole gene, with frequency as low as 30%.
- Our approach allows for the simultaneous evaluation of multiple gene-specific aberrations using a single assay.

References

1. King MC *et al.*, Science. 2003;302:643–646.
2. Antoniou, A. *et al.*, Am J Hum Genet. 2003;72:1117–1130.

Further information **contact@ogt.com | ogt.com | Oxford Gene Technology Ltd., Begbroke Science Park, Woodstock Road, Begbroke, Oxfordshire, OX5 1PF, UK. | SureSeq**: For Research Use Only; Not for Diagnostic Procedures.

990320 01/21