The Analysis of FFPE Samples by Next-Generation Sequencing (NGS) of Key Genes for Research into Breast and Ovarian Cancer

Jacqueline Chan, Sabine Eckert, Lyudmila Georgieva and Graham Speight

Oxford Gene Technology (OGT), Begbroke Science Park, Begbroke Hill, Woodstock Road, Begbroke, Oxford, UK

Introduction

One of the challenges in cancer research is the high level of genetic complexity and tumour heterogeneity.

Research that generates detailed information about the genetic profile of each individual tumour will further our understanding and may be used in the future to guide treatment strategies[1].

NGS has enabled the simultaneous study of multiple mutations in high-penetrance cancer predisposition genes. However, tissue biopsies are typically archived as formalin-fixed, paraffin embedded (FFPE) blocks which can significantly compromise the quality and amount of nucleic acids available for genomics research.

To overcome these issues, we have used the SureSeq™ FFPE DNA Repair Mix, in combination with a hybridisation-based NGS custom enrichment panel, the SureSeq Ovarian Cancer Panel (Table 1) to identify somatic variation in key DNA repair genes associated with ovarian cancer.

BRCA1	*BRCA2*	*ATM*	*TP53*	*ATR*	*NF1*	*PTEN*

Table 1: Key ovarian cancer-related genes in the SureSeq Ovarian Cancer Panel

To evaluate the application of a hybridisation-based approach we:

- Compared the uniformity of coverage between a PCR-based and a hybridisation-based enrichment approach for the analysis of *BRCA1* and *BRCA2* in solid tumour samples[a].
- Identified potentially important variants in *TP53* and *BRCA1* genes from DNA extracted from FFPE blocks of type II epithelial ovarian cancer (EOC) samples[b].

SureSeq hybridisation workflow

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

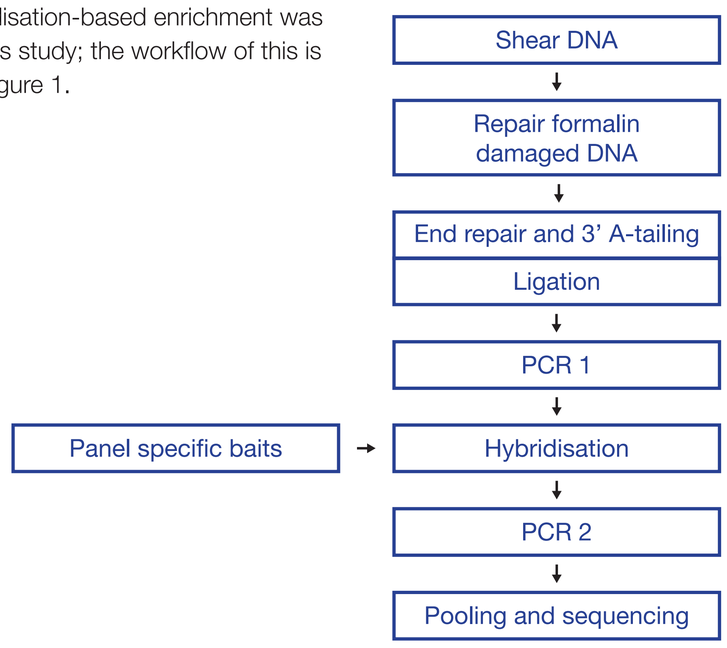

Figure 1: OGT SureSeq workflow. The SureSeq workflow allows users to go from extracted DNA to sequencer in 1.5 days with minimal handling time.

Formalin damage in DNA can be reduced through use of FFPE DNA repair mix

We tested a range of FFPE-derived DNA and formalin-compromised DNA (Horizon Diagnostics - HD803 and HD799) and found pre-treatment with the SureSeq FFPE DNA Repair Mix significantly improves the number of on-target reads, thereby increasing the flexibility of the assay (Figure 2A). Use of the Repair mix also enables a reduced DNA input down to 50 ng to be used (if necessary) whilst maintaining a good depth of coverage (Figure 2B).

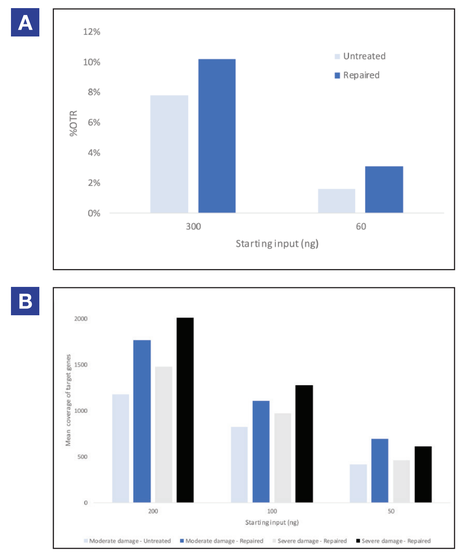

Figure 2: Example data obtained using FFPE DNA extracted from colon cancer samples and Horizon Diagnostics Quantitative Multiplex Formalin Compromised Reference Standards. Panel A shows that the SureSeq FFPE DNA Repair Mix improves on-target rate; Panel B demonstrates the Repair mix permits the use of lower DNA inputs whilst maintaining depth of coverage.

Hybridisation-based enrichment generates highly uniform coverage of key targets

To confidently call low frequency variants, NGS reads need to be evenly distributed across all regions of interest. Uniformity of coverage is a useful value with which to compare this distribution and can be expressed as the percentage of target bases that have >20% of the mean coverage.

As reported extensively in the literature[1-3], we found the uniformity of coverage from hybridisation-based capture approaches, such as SureSeq, consistently outperform those enriched using amplicon-based methods (Figure 3). The uniformity of coverage for most samples is >99% of bases covered at >20% of the mean, ensuring that all bases within the panel can be assessed confidently. In addition, the use of hybridisation-based capture instead of amplification-based enrichment allows the removal of PCR duplicates which can obscure the minor alleles present within a sample.

Uniformity of coverage using amplicon and hybridisation-based approaches

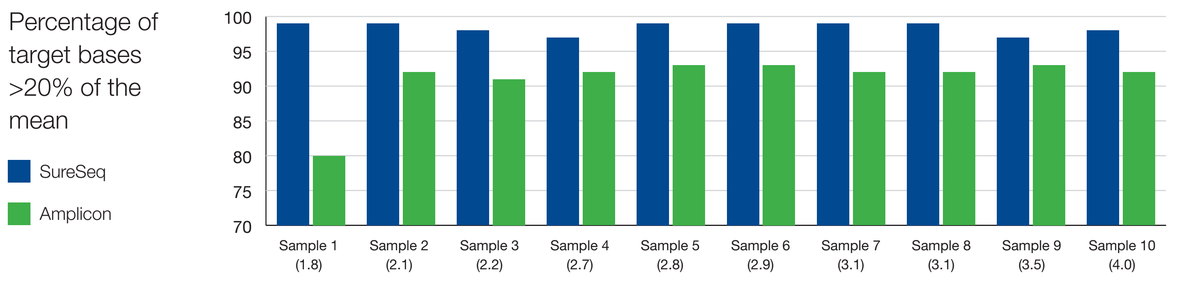

Figure 3: Assessment of the uniformity of sequencing coverage from FFPE-derived DNA using an amplicon and the SureSeq hybridisation-based capture approaches. Enrichment by SureSeq (dark blue bars) demonstrates better uniformity than that of an amplicon-based approach (green bars). Samples are ordered by increasing DIN determined by Agilent 2200 TapeStation – value in brackets.

Confident detection of germline and somatic variants in key cancer-related genes

We tested over 100 EOC samples determined by pathology to contain ≥40% tumour cells and identified one or more deleterious *TP53* variant(s) with the minor allele frequencies (MAF) ranging from 1 to 87%. In addition to the mutations in *TP53*, some samples were found to have variants in *BRCA1* (Figure 5). Figures 4 – 5 were visualised using Integrated Genomics Viewer[4]; the grey vertical bars denote the depth of coverage per base, green horizontal bars the targeted region, and the red heatmap - the GC content.

Figure 4: This sample (DIN score 3.0) was found to have a single base insertion - Cys124TrpfsTer25, at 2.8% in exon 10 of *TP53* (transcript NM_ 000546) which results in a premature termination. It also contains a germline SNP (rs1042522).

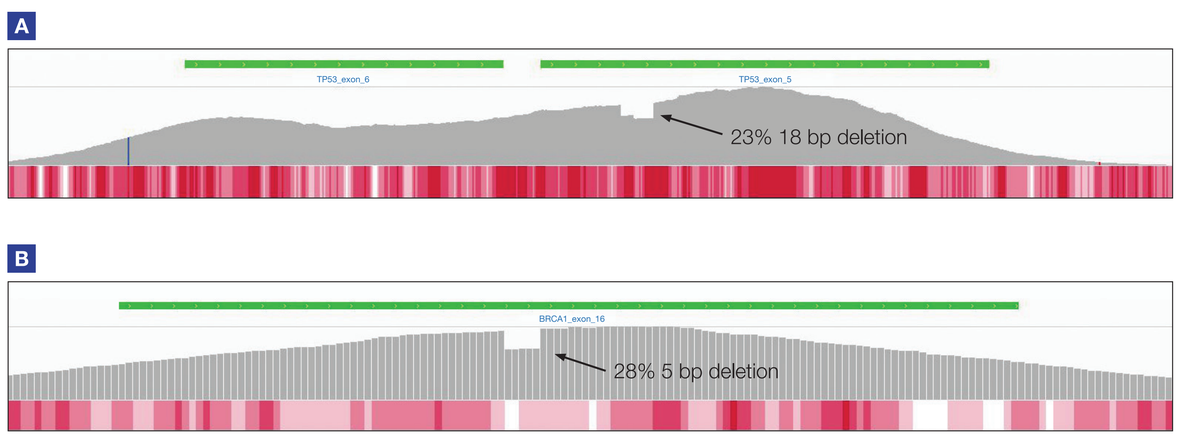

Figure 5: *TP53* exon 6 (panel A) and *BRCA1* exon 16 (panel B). This sample (DIN score 2.6) contains an 18 bp deletion in *TP53* at 23% and a 5 bp deletion of 28% allele frequency in *BRCA1.*

Conclusions

- It is possible to obtain important sequence information from as little as 50 ng of formalin-compromised DNA.
- Superior uniformity of coverage was demonstrated using a hybridisation-based enrichment approach.
- High levels of uniformity were maintained across a range of starting DNA input amounts in formalin-compromised DNA.
- The SureSeq hybridisation-based approach is a robust method for the confident identification of germline and somatic mutations in *TP53* and *BRCA1* from type II EOC tissue samples.
- The utility of this approach and panel permits the analysis of somatic variation in these key DNA repair genes associated with ovarian and breast cancers but can also be used for research into many other cancers including prostate, pancreatic and melanoma.

References

1. Ross, J.S. and Cronin, M., 2011. Whole cancer genome sequencing by next-generation methods. *American journal of clinical pathology*, 136(4), pp.527-539.
2. Kurman, R.J. and Ie-Ming, S., 2010. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *The American journal of surgical pathology*, 34(3), pp.433-443.
3. Samorodnitsky, E., Jewell, B.M., Hagopian, R., Miya, J., Wing, M.R., Lyon, E., Damodaran, S., Bhatt, D., Reeser, J.W., Datta, J. and Roychowdhury, S., 2015. Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Human mutation*, 36(9), pp.903-914.
4. Thorvaldsdóttir, H., Robinson, J.T. and Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*, 14(2), pp.178-192.

Acknowledgements

Samples kindly provided by –

[a] Prof. Charlie Gourley (Cancer Research UK Edinburgh Centre)

[b] Prof. Rober Zeillinger and Dr. Nicole Concin (Medical University of Vienna and Medical University, Dept. of Gynecology and Obstetrics, Vienna, Austria)

Oxford Gene Technology, Begbroke Science Park, Begbroke, Oxfordshire OX5 1PF, United Kingdom
For further information, email: contact@ogt.com or visit: www.ogt.com