SureSeq

Development of a targeted hybridisation-based NGS workflow for use with cell-free DNA

Graham Speight, Carolin Müller, James Reid, Lynne Edhouse, Jolyon Holdstock, Lyudmila Georgieva Oxford Gene Technology (OGT).

A Sysmex Group Company

What binds us, makes us.

Introduction

The capture and molecular analysis of cell-free DNA (cfDNA) released from cells under certain pathological conditions, including cancer, provides a non-invasive alternative to currently used practices via use of liquid biopsy samples.

The rapid development of next-generation sequencing (NGS) technologies in recent years, has led to a significant reduction in sequencing cost with improved accuracy.

In liquid biopsy research, NGS can be applied to sequence circulating tumour DNA (ctDNA; DNA fragments released by tumour cells). As such, NGS could provide molecular characterisation of tumour-specific genomic alterations in ctDNA, allowing for non-invasive and real-time monitoring of disease development.

The aim of this study is to evaluate a modified version of the OGT™ Universal NGS Workflow Solution in conjunction with a custom SureSeq™ targeted panel for use with cfDNA.

Methods

A modified version (not commercially available) of the OGT Universal NGS Workflow Solution was used throughout this study (Fig 1). The approach offers a streamlined NGS library preparation protocol with Unique Molecular Identifiers (UMIs) and unique dual indexing (UDI), followed by hybridisation-based target enrichment.

For this capability study, we used a 64 kb SureSeq myPanel™ Custom panel comprising 213 target exons in 40 genes.

Sequencing was conducted using either MiSeq® V2 300 or NextSeq® 500/550 Mid output chemistry (Illumina).

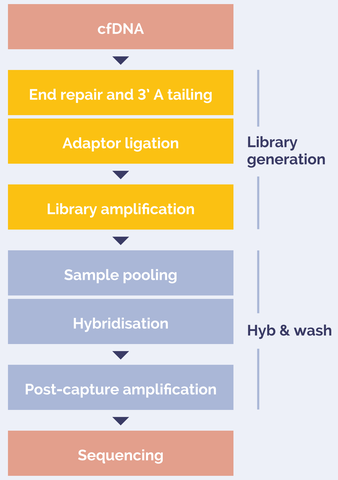

Figure 1: Modified OGT SureSeq workflow for cfDNA. The workflow allows users to go from extracted cfDNA to sequencer in 1.5 days with minimal handling time.

Samples

We tested a commercially available cfDNA standard (Seraseq® ctDNA Complete™ Reference Material AF1%, SeraCare) with 10–50 ng DNA input. To mimic different variant allele frequencies (VAF), we diluted the reference DNA to create samples with a frequency range of 0.5 - 1%.

Bioinformatic Analysis

Sequencing data analysis was performed using OGT's proprietary Interpret NGS Analysis Software. Specifically, the software was used for read alignments, UMI deduplications, coverage calculations and variant calling. The data can also be visualised within the software using an IGV (Integrated Genomic Viewer).

Results I

Generation of high-quality sequencing libraries from low input cfDNA

We have successfully generated high-quality sequencing libraries using 10 – 50 ng cfDNA.

Hybridisation-based enrichment of these libraries has achieved a high unique coverage of ≥1,000x from as little as 10 ng cfDNA input (minimum 11 M reads).

We have achieved very high mean coverages (1,500x - 2,000x) from 25 ng and 50 ng DNA inputs (minimum 12 M reads).

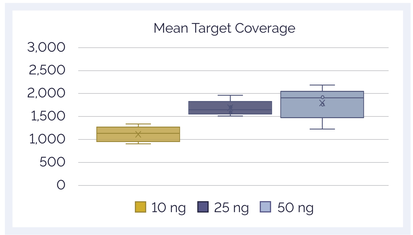

Figure 2: Mean Target Coverage for libraries generated from 10, 25 or 50 ng cfDNA. Error bars reflect standard deviation of replicates (n=5).

Results II

Highly uniform coverage allowing confident SNVs detection in cfDNA samples

The hybridisation-based OGT Universal workflow coupled with OGT's expert bait design deliver excellent depth of coverage and coverage uniformity for 10 – 50 ng cfDNA (Fig 3), allowing consistent detection of SNVs (Fig 4).

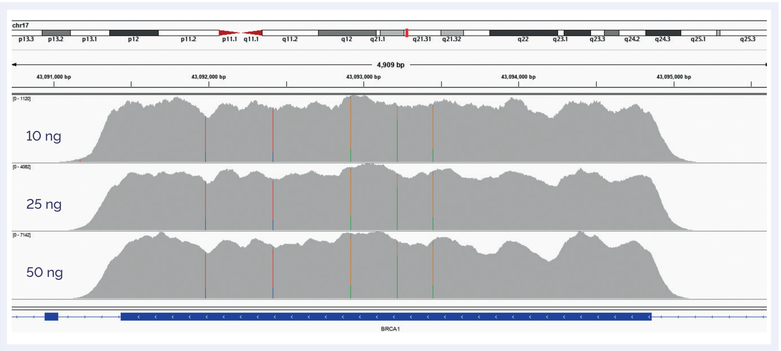

Figure 3: Illustration of the excellent coverage uniformity of *BRCA1* exon 10 in cfDNA samples with 10, 25 or 50 ng input amount. Depth of coverage per base (grey). Gene coding region as defined by RefSeq (bottom track)

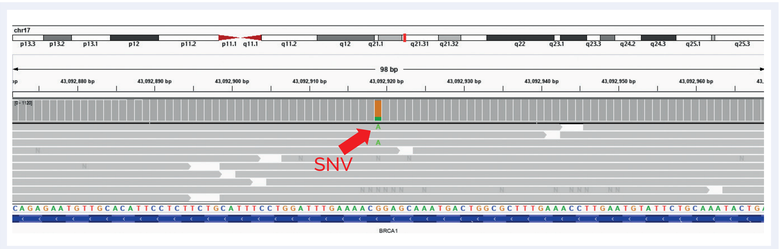

Figure 4: Example of *BRCA1* exon 10 missense variant Pro871Leu (rs799917) with frequency of 20% identified in 10 ng cfDNA. Shown are the genomic location (top), depth of coverage per base (grey bars), aligned example reads (grey lines) and the reference sequence.

Results III

For proof of principle experiments, we tested 0.5% VAF for inputs of 25 and 50 ng, and 1% VAF for 10 ng of SeraCare cfDNA reference material. These VAF were chosen to match a theoretical detection limit of ≥10 supporting reads for variants based on the Mean Target Coverage we achieved for each input in Results I.

Confident detection of SNVs and indels with 0.5% VAF using 25 ng or 50 ng cfDNA

	Gene	Variant	25 ng 0.5% VAF		50 ng 0.5% VAF	
			Total reads	Variant reads	Total reads	Variant reads
SNV	*AKT1*	c.49G>A	3,998	48	7,892	67
	APC	c.4348C>T	6,989	48	10,497	52
	BRAF	c.1799T>A	4,785	33	7,096	60
	CTNNB	c.121A>G	7,342	80	10,225	101
	FGFR3	c.746C>G	4,643	51	8,710	71
	KIT	c.2447A>T	4,087	28	5,618	57
	KRAS	c.35G>A	4,809	32	6,613	82
	NRAS	c.182A>G	5,033	42	6,800	40
	PDGFRA	c.2525A>T	4,919	30	7,426	85
	PIK3CA	c.3140A>G	4,940	22	7,147	35
	RET	c.2753T>C	4,074	39	5,768	39
	TP53	c.524G>A	7,234	69	10,790	105
	TP53	c.743G>A	7,032	50	10,821	97
	TP53	c.818G>A	4,821	40	6,498	59
Indels	*PDGFRA*	c.1694_1695insA	4,747	61	6,882	43
	PIK3CA	c.3204_3205insA	3,522	19	5,222	28
	APC	c.4666dup 6A>7A	3,042	7	4,240	17
	TP53	c.723del	5,321	36	8,298	65
	TP53	c.267del	3,246	27	4,707	57

A modified version of OGT's Universal workflow together with a SureSeq myPanel Custom panel successfully detected all anticipated SNVs (n=14) and Indels (n=5) with 0.5% VAF using 25 or 50 ng SeraCare cfDNA (Table 1).

The data also shows that with deeper sequencing, Mean Target Coverage of 5,400x and 7,700x can be achieved for 25 ng (29 M reads) and 50 ng (41 M reads) inputs, respectively. This suggests a VAF detection limit below 0.5% might be possible.

Table 1: Detection of SNVs and Indels in 25 and 50 ng SeraCare cfDNA reference standards with 0.5% VAF. Variant reads are filtered for unique UMIs only, and not for UMI family sizes.

Confident detection of SNVs and indels with 1% VAF using 10 ng cfDNA

	Gene	Variant	10 ng 1% VAF	
			Total reads	Variant reads
SNV	*AKT1*	c.49G>A	3,909	34
	APC	c.4348C>T	6,219	30
	BRAF	c.1799T>A	4,543	37
	CTNNB	c.121A>G	6,127	55
	FGFR3	c.746C>G	2,404	21
	KIT	c.2447A>T	3,601	27
	KRAS	c.35G>A	3,654	46
	NRAS	c.182A>G	4,455	36
	PDGFRA	c.2525A>T	4,609	54
	PIK3CA	c.3140A>G	4,382	45
	RET	c.2753T>C	3,837	34
	TP53	c.524G>A	5,196	74
	TP53	c.743G>A	6,064	60
	TP53	c.818G>A	4,249	67
Indels	*PDGFRA*	c.1694_1695insA	4,210	51
	PIK3CA	c.3204_3205insA	3,330	27
	APC	c.4666dup 6A>7A	2,986	30
	TP53	c.723del	5,117	55
	TP53	c.267del	2,085	16

We have achieved detection of all anticipated SNVs (n=14) and Indels (n=5) with 1% VAF using 10 ng SeraCare cfDNA (Fig 5, Table 2). A Mean Target Coverage of 4,200x was achieved with 43 M reads.

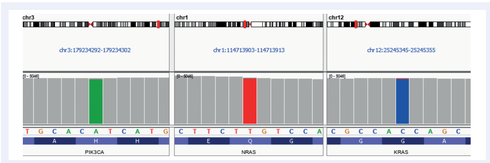

Figure 5: Three examples of detected SNVs with 1% VAF in 10 ng – *PIK3CA* exon 21 His1047Arg (COSM775), *NRAS* exon 3 Gln61Arg (COSM584) and *KRAS* exon 2 Gly12Asp (COSM521). Shown are the genomic location (top), depth of coverage per base (bars) and the reference sequence.

Table 2: Detection of SNVs and Indels in 10 ng SeraCare cfDNA reference standards with 1% VAF. Variant reads are filtered for unique UMIs only, and not for UMI family sizes.

Conclusions

- We have demonstrated that a modified version (not commercially available) of OGT Universal NGS Workflow Solution is suitable for very low inputs of cell-free DNA.
- High coverage uniformity after hybridisation-based enrichment using a custom panel allowed simultaneous detection of SNVs and indels in hundreds of target exons.
- We have shown that our approach reliably detected somatic mutations down to 0.5% VAF.
- These results confirm the capability of utilising this type of approach for cell-free DNA research applications.

Further information **contact@ogt.com | ogt.com | Oxford Gene Technology Ltd., Begbroke Science Park, Woodstock Road, Begbroke, Oxfordshire, OX5 1PF, UK.**

990348 10/22