SureSeq

Development of a target-capture NGS assay for use in molecular-based research of myeloid measurable residual disease (MRD)

Authors: Natalie Milner, Giulia Poloni, James Reid, Venu Pullabhatla, Lyudmila Georgieva, Graham Speight Oxford Gene Technology (OGT).

A Sysmex Group Company

What binds us, makes us.

Introduction

Molecular methods are increasingly supporting traditional immunophenotypic multiparameter flow cytometry (MFC) in detection of measurable residual disease (MRD), including acute myeloid leukaemia (AML).

Real-time qPCR (RQ-PCR) and digital droplet PCR (ddPCR) are highly sensitive technologies but are limited by the number of targets that can be detected in one assay.

Next-generation sequencing (NGS) offers the opportunity to evaluate many genes in a single assay. Improved accuracy together with falling costs are facilitating the use of NGS in MRD.

We have developed a target-capture NGS approach to support researchers in studies of molecular-based MRD monitoring in myeloid malignancies.

This method provides the opportunity to evaluate many genes and variant types in a single assay.

Materials and Methods

Workflow

Libraries were generated using OGT's Universal NGS Workflow (Fig. 1). The workflow is ideally suited to low frequency variant detection through the inclusion of unique dual indexing (UDIs) and Unique Molecular Identifiers (UMIs).

Samples

Myeloid Reference DNA Standard (Horizon Discovery), with 6 SNVs, 2 indels and a 300bp *FLT3* ITD diluted with normal DNA to generate a frequency ranges between 0.02%-0.1% variant allele frequency (VAF).

Panel

SureSeq™ Myeloid MRD panel targeting 43 exons in 13 genes (8 kb target; 11 kb baited).

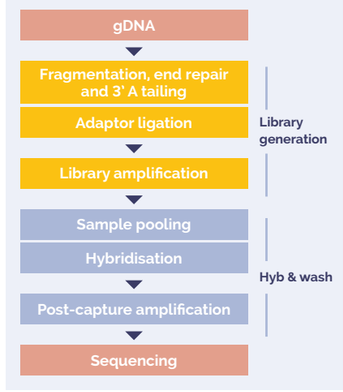

Figure 1: DNA to sequencer in 1.5 days with minimal handling time.

Sequencing

2 x 150 PE reads, Illumina NextSeq® 500.

Bioinformatic Analysis

Sequencing data analysis was performed using OGT's proprietary Interpret software, including read mapping, error correction, coverage calculation and variant calling.

Assay development

A variant at 0.1% VAF requires a depth of 10,000x to be detected with 10 supporting reads.

A variant at 0.05% VAF requires a depth of 20,000x to be detected with 10 supporting reads.

We increased reads/sample and monitored the depths through:

- Depth of coverage distribution.
- Uniformity of coverage across targets.

We used the UMIs for error correction and determined the effect of increased reads on:

- Distribution of UMI family size (FS).
- Percentage of data removed in FS2 filtering.

Results

UMI coverage supports detection @ 0.05%

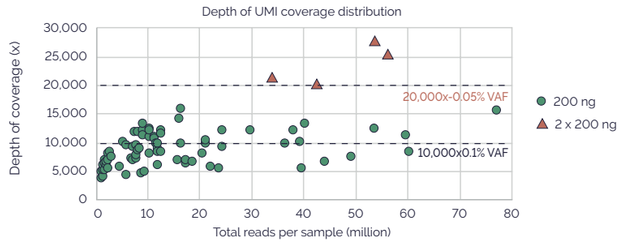

Figure 2: Relationship between sequencing and unique depth of coverage

- Average depth of coverage increases proportionally up to 10 million total reads (green circles).
- At higher total reads unique coverage is increased by increasing the input DNA (brown triangles).
- >20,000x coverage is achieved with 40 million reads/sample.

Average UMI family size increases with total reads

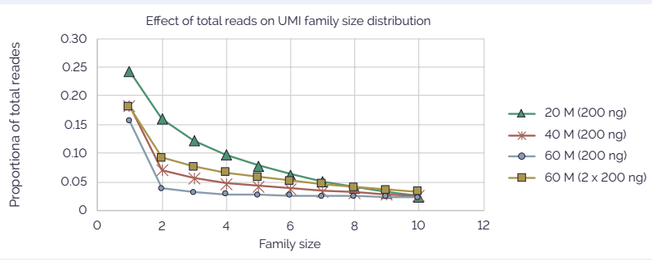

Figure 4: UMI family size (FS) distribution showing increase with increasing reads/sample

Highly uniform UMI coverage across target regions

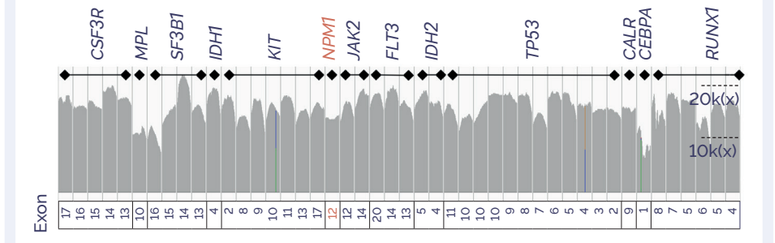

Figure 3: IGV plot showing coverage profile of target regions in the panel

- Uniform coverage is essential to maintain equal sensitivity across all targets.
- High uniformity is demonstrated (Fig. 3)-including difficult target: *NPM1* exon 12.

Increasing total reads reduces error correction data loss

- Removing reads with a family size of 1 reduces background for SNVs and increases the accuracy of the call.
- Increasing the reads per sample increases the family size (Fig. 4).
- Increasing the DNA input caused a small decrease in the family size with 60 M reads/sample (gold squares).
- An increase in family size corresponds to a reduction in FS1 reads removed in error correction (Table 2).

Reads/sample (million)	Mean UMI coverage	Mean Reads in FS1
10 M-20 M	8,930x	24%
30 M-40 M	9,193x	18%
50 M-60 M	11,460x	16%
35 M-60 M2x200ng*	23,224x	18%

Table 2: Percentage of reads removed in error correction reduces as depth increases.

100% detection of variants @ ≥0.04% VAF (40m reads/sample)

	Gene	Read depth	0.04% VAF		0.05% VAF		Neg. ctrl	
		Variant	Total	Variant	Total	Variant	Total	Variant
SNV	*SF3B1*	c.2219G>A	15,289	10	14,312	4	15,086	0
	JAK2	c.1849G>T	24,520	5	22,375	18	22,017	0
	FLT3	c.2503G>T	26,556	10	24,127	9	23,991	0
	IDH2	c.515G>A	18,950	4	16,565	6	18,404	0
	TP53	c.722C>T	14,199	10	12,618	4	13,838	0
Indel	*NPM1*	c.860_863dup	16,536	8	11,536	10	4,805	0
	JAK2	c.1611_1616del	26,267	8	25,683	7	22,017	0
ITD	*FLT3*	ITD300	13,119	70	12,208	57	21,686	0

Table 1: Detection of SNVs, Indels and an ITD, with expected frequency ranges of 0.04%-0.05%. SNVs are filtered to remove FS1 reads.

- OGT's Universal NGS Workflow combined with the SureSeq Myeloid MRD panel confidently detected all anticipated variants ≥0.04% including *NPM1* insertion and a 300 bp *FLT3* ITD.
- No reads are observed in the negative control (excluding *IDH1* with a known high level of background).

Conclusions

- The OGT Universal NGS Workflow is suitable for use in MRD monitoring providing high uniformity across all target regions including *NPM1* exon 12.
- Increasing input DNA and reads per sample improves variant detection and accuracy through greater coverage and improved error correction.
- Our approach achieves coverage depth required to detect somatic mutations ≥ 0.04% VAF.
- 100% detection of variants ≥ 0.04% VAF: SNVs, indels (including *NPM1*) and a 300 bp *FLT3* ITD.
- This assay provides researchers with the capability to use capture-based NGS technology to simultaneously detect a number of variants in MRD monitoring.

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

990352 06/23 v2