Optimised, 1-day hybridisation-based NGS protocol yields 1% variant detection in MPN samples, as quickly and cost-effectively as multiplex PCR enrichment

Graham Speight[1], Ephrem Chin[1], Lyudmila Georgieva[1], Nick Cross[2] and David Cook[1]

[1]Oxford Gene Technology, Oxford, UK and [2]National Genetics Reference Laboratory - Wessex, UK

ogt

Introduction

- Myeloproliferative neoplasms (MPNs) are a group of diseases that affect blood cell production in the bone marrow resulting in the overproduction of one or more cell types. The key MPN driver mutations involve *JAK2* (V617F [occurrence of 50-98% depending on the MPN subtype]), *JAK2* exon 12, *MPL* W515K/L and *CALR* exon 9 insertion and deletions.

- Designed for research into the diagnosis, aetiology and prognosis of MPNs, the SureSeq™ Core MPN Panel has been developed by Oxford Gene Technology (OGT) in collaboration with recognised cancer experts to deliver accurate detection (down to a 1% variant allele fraction [VAF]) of somatic variants of these key MPN driver mutations.

- The aim of this study is to evaluate the SureSeq™ Core MPN Panel in conjunction with a new streamlined 1-day hybridisation-based Next Generation Sequencing (NGS) library preparation kit (LPK).

Methods

Preparation of purified DNA to sequencer-ready libraries in 7 hours 45 minutes

- An enhanced version of the SureSeq™ Library Preparation Kit (OGT) was utilised which incorporated an enzymatic DNA fragmentation in combination with a rapid hybridisation of just 30 minutes. This enhanced protocol reduces the overall processing time by 6 hours, and now offers a streamlined, 1-day workflow from purified DNA sample to sequencer.
- This protocol offers a similar turn-around time to amplicon-based enrichment protocols, without the associated disadvantages, such as PCR bias, allelic bias (indels) and drop-outs, as well as poor uniformity of coverage.

Figure 1: Workflow of SureSeq NGS library preparation, from DNA to sequencer.

Panel content
- The SureSeq Core MPN Panel covers 3 genes important in MPN research: *JAK2* (V617F and exon 12), *MPL* (W515K/L/R/A and S505N) and *CALR* (exon 9).

Study design
- The SureSeq Core MPN Panel was technically validated using the *JAK2* V617F Genotyping Sensitivity Panel provided by the National Institute for Biological Standards and Control (NIBSC) in order to confirm the lower levels of analytical allele detection and confidence.
- The SureSeq Core MPN Panel was also used to confirm a broader set of variants in 14 research samples (provided by the National Genetics Reference Laboratory - Wessex, UK) containing variants for each of the targeted regions. Sequencing was conducted on a MiSeq® using a V2 300 bp cartridge (Illumina).

Results

Accurate and reproducible variant detection even in heterogeneous samples
- Heterogeneous cancer samples pose significant challenges as alleles are likely to be present at a lower fraction than would be expected for standard germline variants. Samples typically contain a mixture of cancer and normal cells; moreover, cancer can consist of several molecularly distinct clones. In order to detect alleles that contribute only a small percentage to the reads at any locus, a highly uniform and sensitive enrichment is required.
- The SureSeq Core MPN Panel has been technically validated with samples from the NIBSC and has been shown to accurately detect alleles down to a 1% VAF in *JAK2* (V617F) at a read depth of >1000x (Table 1).

NIBSC JAK2 V617F Sample	% VAF	Wild type reads	Variant reads
10%	13.1	1038	157
5%	4.6	1623	78
1%	1.0	2172	23
0.1%	0.4	2414	9
0%	0.0	2555	1

Table 1: Data generated from a 48 sample run on an Illumina MiSeq. The SureSeq Core MPN Panel permitted the detection of alleles at 1% VAF with high confidence.

Results

The Core MPN Panel in combination with the enhanced workflow is able to reliably detect single nucleotide variants (SNVs) as well as insertions (5 bp insertion in *JAK2* exon 12 and *CALR* exon 9) and deletions (5 bp deletion exon 12 *JAK2* and 52 bp deletion *CALR* exon 9) (see below).

Wild-type

52 bp deletion

Figure 2: Detection of a 52 bp deletion (exon 9 *CALR*). Wild-type sample (top panel) is compared to a 52 bp somatic deletion (bottom panel).

Wild-type

5 bp deletion

Figure 3: Detection of a 5 bp deletion (exon 12 *JAK2*). Wild-type sample (top panel) is compared to a 5 bp somatic deletion (bottom panel).

Results

Technical validation of the Core MPN Panel and enhanced library preparation approach with 14 research samples
- Data presented here are from 14 research samples (including 2 wild-type controls) that were processed using the enhanced LPK in combination with the Core MPN Panel.
- See below (Table 2) for the range of variants and VAFs detected from 14 research samples. These include SNVs as well as 5 bp insertions in *JAK2* (exon 12) and deletions of up to 52 bp in *CALR* (exon 9) (Table 2). No variants were identified in the control samples.

#	Gene	Variant	Mean target coverage	% VAF
1	*CALR*	K385NfsTer47 (5 bp insertion in exon 9 *CALR*)	3640	37.3
2	*JAK2*	V617F	1263	10.7
3	*JAK2*	V617F	1352	9.8
4		Control - no variant detected	2655	
5	*MPL*	W515R	2750	18.1
6	*JAK2*	V617F	2971	2.7
7	*JAK2*	V536F537insFHKIRNEDLIL (33 bp insertion in exon 12 *JAK2*)	1574	3.1
8	*CALR*	K385NfsTer47	2321	42.8
9	*JAK2*	N542E543del	1749	29.9
10	*CALR*	L367WfsTer46 (52 bp deletion exon 9 *CALR*)	947	16.3
11	*MPL*	W515L	1421	60.7
12	*MPL*	W515L	903	3.8
13	*CALR*	K385NfsTer47 (5 bp ins in exon 9 *CALR*)	1265	14.0
14		Control - no variant detected	3234	

Table 2: Data generated using the SureSeq Core MPN Panel in combination with the enhanced LPK was 100% concordant with independent findings (National Genetics Reference Laboratory – Wessex, UK).

Conclusions

- We have successfully utilised the OGT 1-day hybridisation-based SureSeq LPK protocol in combination with the SureSeq Core MPN Panel to reliably and routinely detect somatic SNVs by NGS down to a 1% VAF.
- The uniformity of coverage of this approach permitted the detection of *CALR* and *JAK2* indels (including 52 bp deletions and 5 bp insertions).
- This enhanced protocol incorporates an enzymatic fragmentation which permits the high-throughput preparation of 48 samples from genomic DNA to sequencer in a 1-day workflow.
- To achieve >1000x de-duplicated depth (required for confident detection of 1% VAF), 48 samples can be reliably sequenced in a single MiSeq (V2 300 bp) run. This allows the generation of high quality data in a cost effective and timely manner.