Beyond FISH, SNVs and indels – improved resolution of translocation detection using next generation sequencing (NGS)

Lyudmila Georgieva, Ezam Uddin, Aysel Heckel, Jacqueline Chan, James Reid, Jolyon Holdstock, David Cook and Graham Speight

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

Introduction

Karyotyping, FISH, RT-PCR and microarrays are currently considered to be the gold standard techniques for structural variant discovery and detection. However, there is a desire to combine analysis of large structural alterations such as translocations alongside smaller mutations such as SNVs and indels. With development of newer technologies such as NGS for DNA- and RNA-sequencing, simultaneous discovery of multiple mutation types is now possible, enabling development of more comprehensive assays.

Chronic myeloid leukaemia (CML) is a myeloproliferative neoplasm with incidence of 1 to 2 per 100,000 and constitutes 15-20 % of adult leukaemias[1]. CML is characterised by the Philadelphia chromosome (Ph), resulting from the t(9;22)(q34;q11) balanced reciprocal translocation. The translocation generates the *BCR-ABL1* fusion gene encoding the BCR-ABL1 protein with constitutive kinase and oncogenic activity[1,2]. The breakpoints in the *ABL1* gene lie in 90 kb long intron 1, upstream of the *ABL1* tyrosine kinase domains encoded in exons 2 to 11. The breakpoints within *BCR* are mapped to a 3.1 kb area spanning exons 13 to 15, the major breakpoint cluster region (M-bcr), found in 90% of CML and 20 to 30% of cases with Ph-positive B-cell acute lymphoblastic leukaemia (Ph+ B-ALL)[3].

In this study, we tested the capability of a SureSeq myPanel™ NGS Custom Cancer Panel to detect known t(9;22)(q34;q11) translocations.

Methods

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

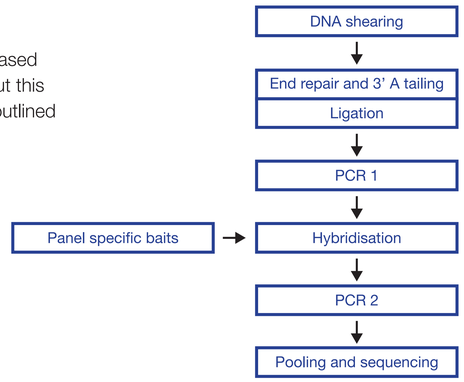


Figure 1: OGT SureSeq workflow. The SureSeq workflow allows users to go from extracted DNA to sequencer in 1.5 days with minimal handling time.

Study design

We utilised a SureSeq myPanel NGS Custom panel and associated library preparation kit* to determine whether this approach can be used for detection of the t(9;22)(q34;q11) translocation.

We used a hybridisation-based enrichment approach to test characterised DNA samples, CML cell lines K562, KU-812, MOLM-1 and JK-1, and 3 research samples** with breakpoints occurring across multiple positions in the major breakpoint area of *BCR* (exons 13-15) (Figure 2).

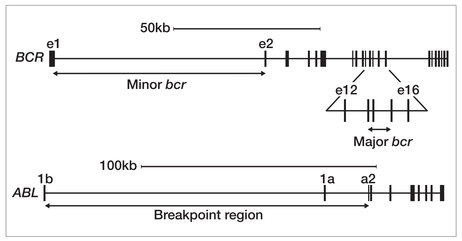


Figure 2: Translocation breakpoint clusters in *BCR* and *ABL1* genes.

To mimic the *BCR-ABL1* translocation with different frequencies we made a serial dilution of cell line K562 in order to create translocation carriers with frequency range 0 - 100%.

Sequencing was conducted using a V2 300 bp cartridge (Illumina).

Data was analysed using OGT's proprietary translocation detection software.

Results

We have achieved high depth (>1000x) and uniformity of coverage across the targeted regions which enabled successful detection of all translocation events.

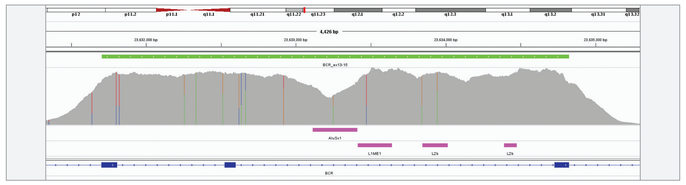


Figure 3: Excellent uniformity of coverage of the *BCR*, exon 13-15 region, generated with OGT SureSeq protocol averaging ~1500x deduplicated coverage. Depth of coverage per base (grey). Target region (green). Repeat regions (pink).

Prediction and confirmation of translocation breakpoints in CML cell lines

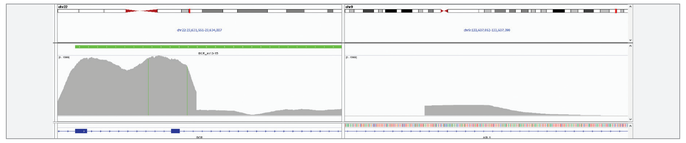
Using the OGT workflow we were able to reliably detect *BCR-ABL1* translocations in 4 CML cell lines. For all cell lines, predicted breakpoints were found to match the coordinates reported in the literature[4] (Table 1, Figure 4).

Sample Number	Sample Name	*BCR* read value	*ABL1* read value	*BCR* breakpoint location (chr22)	*ABL1* breakpoint location (chr9)	P value
1	K562	1274	4041	23, 632, 742	133, 607, 145	$<10^{-150}$
2	MOLM-1	1320	115	23, 631, 909	133, 643, 072	$5.26x10^{-2}$
3	KU-812	727	186	23, 632, 850/23, 632, 613	133, 643, 198/133, 597, 361	$3.00x10^{-76}$
4	JK-1	2104	265	23, 632, 390/23, 632, 389	133, 628, 935/133, 628, 932	$5.16x10^{-19}$

Table 1: Accurate detection of *BCR-ABL1* translocations by early stage software (in development). Read values are a formulation of read depth, not true read depth.

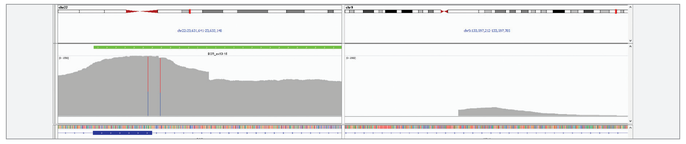
Figure 4: *BCR-ABL1* translocation detection in K562 **A**, MOLM-1 **B**, KU-812 **C** and JK-1 **D** cell lines.

A



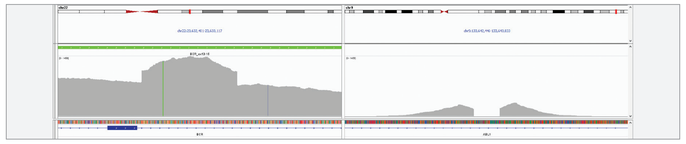
Breakpoint location *BCR* chr22 : 23,632,742, *ABL1* breakpoint location chr9 : 133,607,145.

B



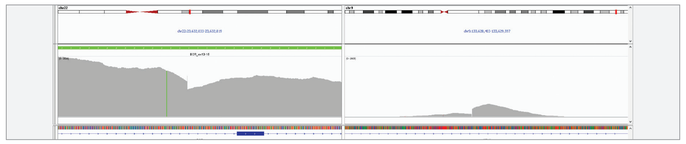
Breakpoint location *BCR* chr22 : 23,631,909, *ABL1* breakpoint location chr9 : 133,597,361.

C



Both *BCR-ABL1* and *ABL1-BCR* translocations were detected. *BCR/ABL1* : chr22 : 23,632,850 / chr9 : 133,643,198; ABL1/*BCR*: chr9 : 133,643,072 / chr22 : 23,632,613.

D



Both *BCR-ABL1* and *ABL1-BCR* translocations were detected. *BCR/ABL1* : chr22 : 23,632,390 / chr9 : 133,628,935; *ABL1/BCR*: chr9 :133,628,932 / chr22 : 23,632,389.

Confident detection of a *BCR-ABL1* fusion event was reported at frequencies as low as 5%.

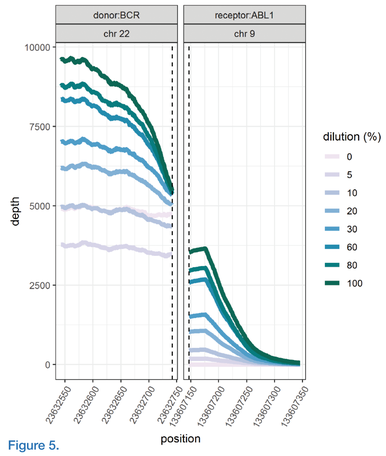


Figure 5.

This data (Figure 5, Table 2) is based on a serial dilution of DNA from K562 cell line in order to create *BCR-ABL1* translocation carriers with frequencies of 0 – 100%.

Expected frequency (%)	*BCR* read value	*ABL1* read value	P value
100	4041	1274	$<10^{-150}$
80	3260	1033	$<10^{-150}$
60	2888	920	$<10^{-150}$
30	1768	545	$<10^{-150}$
20	1319	380	$2.4x10^{-133}$
10	678	166	$9.6x10^{-52}$
5	462	74	$4.9x10^{-8}$
0	378	1	0.99

Table 2.

Figure 5, Table 2: Accurate and reproducible detection of *BCR-ABL1* translocation in K562 cell line (*BCR* breakpoint location chr22:23632742, *ABL1* breakpoint location chr9:133607147) with a range of frequencies. Read values are a formulation of read depth, not true read depth.

Evaluation of the *BCR-ABL1* fusion detection panel in research samples

Data presented here (Table 3) are from 3 research samples from carriers of a *BCR-ABL1* translocation generated using the OGT workflow. For all samples the predicted breakpoints were found to be 100% concordant with independent findings (West Midlands Regional Genetics Laboratory, Birmingham – UK).

Sample Number	*BCR* read value	*ABL1* read value	*BCR* breakpoint location (chr22)	*ABL1* breakpoint location (chr9)	P value
1	577	100	23, 634, 130	133, 607, 145	$3.15x10^{-19}$
2	511	103	23, 633, 828	133, 717, 546/	$2.18x10^{-27}$
3	435	95	23, 634, 596	133, 730, 109	$9.95x10^{-30}$

Table 3: Accurate detection of *BCR-ABL1* translocations. Using the OGT workflow we were able to reliably detect *BCR-ABL1* translocations with different breakpoint locations and frequency in all samples by early stage software (in development). Read values are a formulation of read depth, not true read depth.

Conclusions

- We have successfully utilised the OGT hybridisation-based SureSeq protocol in combination with a SureSeq myPanel NGS Custom panel to reliably detect somatic *BCR-ABL1* translocations.
- The uniformity of coverage of this approach permitted the detection of known *BCR-ABL1* translocations with different breakpoint locations down to a 5% frequency.
- Alongside suitable software, this approach can be used for detection of multiple types of mutations in a single assay, including a combination of SNVs, indels and translocations; and could be adapted for a number of different applications in future.

References

1. Sessions J: Chronic myeloid leukemia in 2007. Am J Health Syst Pharm 2007, 64:S4-9.
2. Quintas-Cardama A, Cortes J: Molecular biology of bcr-abl1-positive chronic myeloid leukemia. Blood 2009, 113:1619-1630.
3. Wong S, Witte ON: The BCR-ABL story: bench to bedside and back. Annu Rev Immunol 2004, 22:247-306.
4. Ross DM, Schafranek L, Hughes TP, Nicola M, Branford S, Score J: Genomic translocation breakpoint sequences are conserved in BCR-ABL1 cell lines despite the presence of amplification. Cancer Genet Cytogenet 2009, 189:138-139.

Acknowledgements

**Samples kindly provided by West Midlands Regional Genetic Laboratory - Birmingham, UK.