SureSeq

Oxford Gene Technology
A Sysmex Group Company

Application note

NGS advances: Efficient fusion event detection in haematological malignancy clinical research

The SureSeq™ Myeloid Fusion Panel, crafted in collaboration with leading experts in myeloid cancer research to align seamlessly with the latest clinical guidelines, was used in this study. Harnessing RNA-based partner gene agnostic technology, our panel allows for simultaneous interrogation of multiple baited target fusions enhancing sample classification, including identification of novel and rare fusions.

Introduction

Acute myeloid leukaemia (AML) and acute lymphoblastic leukaemia (ALL) are both heterogeneous groups of haematological malignancies characterised by both phenotypic and genetic diversity, for which fusion genes are particularly prevalent[1]. For example in AML, well-documented fusion events include *RUNX1::RUNX1T1*, *CBFB::MYH11* and *PML::RARA*, as well as rearrangements involving *KMT2A* and *NUP98*[2]. An estimated 30% of AML patients carry a fusion gene(s)[3], highlighting the significance of fusion gene research into disease classification, prognosis and treatment approaches[4]. Additionally, structural rearrangements of chromosome 3 at the *MECOM* locus [inv(3)(q21q26.2)/t(3;3)(q21;q26.2)] cause *MECOM* overexpression and are observed in 1–2.5% of AML patients[2].

Traditionally, techniques such as fluorescence *in situ* hybridisation (FISH) and reverse transcription polymerase chain reaction (RT-PCR) have been employed for fusion gene detection[5,6]. However, with the advent of next-generation sequencing (NGS), there has been a paradigm shift in fusion gene identification and characterisation.

NGS-based fusion gene panels offer the advantage of simultaneous detection of multiple fusion genes in a single assay, including multiple partners for the same genes[3]. Additionally, interrogation of gene fusions can facilitate the discovery of novel fusion partners, thereby enhancing the depth and breadth of sample characterisation compared to conventional methods.

In this study, conducted at OGT's state-of-the-art NGS research facility in Oxford, the SureSeq Myeloid Fusion Panel was used to corroborate previously characterised fusion-positive and fusion-negative samples, and typify fusions, if present, in a cohort of research biobank samples from Birmingham Women's and Children's NHS Foundation Trust.

Methodology

Samples

50 RNA samples extracted from whole blood or bone marrow were obtained from Birmingham Women's and Children's NHS Foundation Trust*. This 50-sample cohort included 35 samples that were initially characterised as fusion-positive by either FISH, qPCR or G-band and 15 fusion-negative controls. Fusion-negative controls comprised of 2 samples containing fusions not targeted by the SureSeq Myeloid Fusion Panel (untargeted fusions) and 13 AML/ALL samples that were previously characterised as fusion-negative using FISH. Commercially obtained control RNAs included the Human Lymphocyte RNA cell line (AMSBio, UK) as a fusion-negative control and the Universal Human Reference RNA (Agilent, UK) as a fusion-positive control. All samples and controls were assessed using RNA TapeStation Analysis (Agilent, UK) to confirm that RNA was intact. 100–500 ng RNA was used for cDNA synthesis and subsequent OGT Universal NGS Workflow.

*3 initial samples were excluded from analysis following insufficient RNA extraction for further analysis.

Workflow

The SureSeq Universal NGS Workflow Solution was used throughout this study, with the addition of a cDNA synthesis step (Figure 1). The approach offers a streamlined NGS library preparation protocol with Unique Molecular Identifiers (UMIs) and Unique Dual Indexes (UDIs), followed by hybridisation-based target enrichment. For this research study, we used the 61 kb SureSeq Myeloid Fusion Panel comprising 18 target genes (predominantly drivers) allowing for the agnostic detection of baited gene fusions. Furthermore, 3 housekeeping genes used in the panel are used to allow relative expression detection in *MECOM*. Sequencing was conducted using the MiSeq® V2 300 (Illumina, UK).

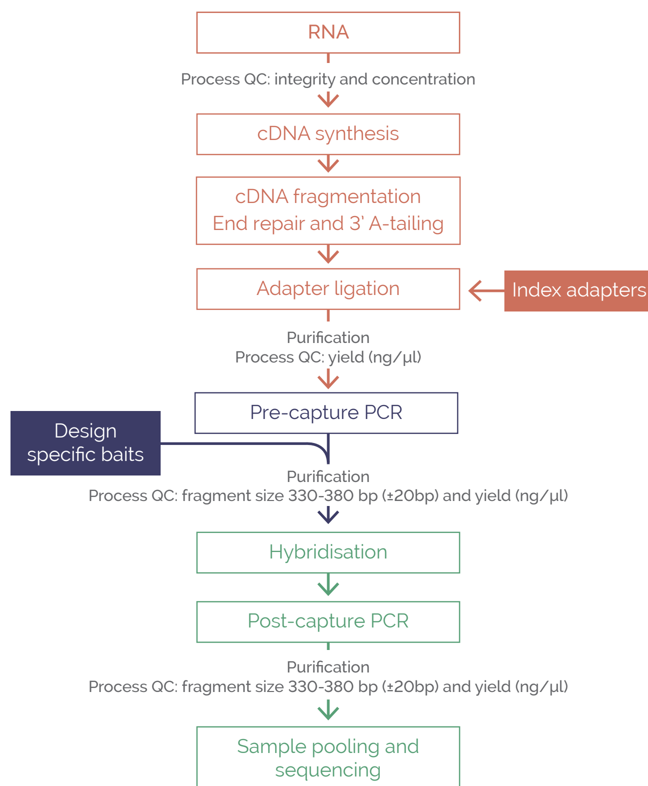


Figure 1. Workflow of cDNA synthesis and sample library preparation

Bioinformatic analysis

Sequencing data analysis was performed using OGT's proprietary Interpret NGS Analysis Software. This software was used to easily visualise fusion genes detected, focus in on breakpoints, the number of reads spanning each breakpoint and alignment of sequence reads at nucleotide resolution. Additionally, the software was used to report normalised gene expression levels related to *MECOM* rearrangements.

Results

Ensuring quality control and concordance

With 100% of the samples displaying cDNA synthesis efficiency above the acceptance threshold and >99% uniquely mapped reads (Figure 2), all samples passed quality control and were further analysed for fusion presence.

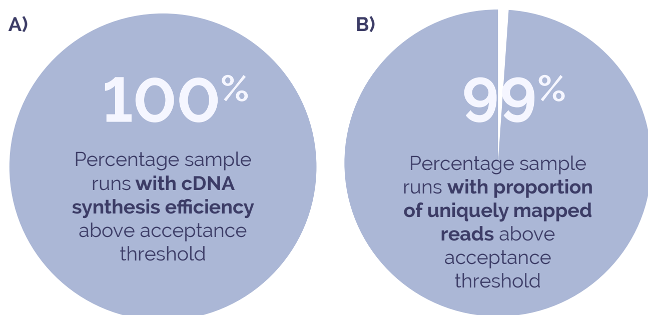


Figure 2. Quality metrics for processed samples
Note, due to a sample processing error, higher than expected levels of DNA (>500 ng) were pooled for some samples prior to sequencing leading to reduced UMR (%), consequently a pass rate of 100% is not observed in Figure 2B.

Table 1 lists the fusion content of the 50-sample cohort (35 fusion-positive samples and 15 fusion-negative controls) tested in this study. The study examined 5 commonly occurring fusions in either AML or ALL using 30 samples: *ETV6::RUNX1*; *RUNX1::RUNX1T1*; *BCR::ABL1*; *PML::RARA* and *CBFB::MYH11*. We also successfully detected *MECOM* inv(3);t(3;3) events that result in *MECOM* overexpression. Furthermore, we tested the partner-agnostic fusion detection capability of SureSeq Myeloid Fusion workflow using 4 samples containing 3 of the most-common *KMT2A* rearrangements in AML: *KMT2A::MLLT3*, *KMT2A::MLLT10* and *KMT2A::AFF1*. 15 negative controls (2 untargeted fusions and 13 fusion-negative samples) were also assessed to confirm assay specificity. We observed 100% concordant detection in all 50 samples, which had previously been characterised by either FISH, qPCR or G-band testing. Furthermore, OGT's SureSeq Interpret software allowed detection of reciprocal fusions for *ETV6::RUNX1*, *PML::RARA* and *BCR::ABL1* that are known to occur frequently in AML or ALL patients which are not typically tested for by FISH or qPCR.

Target gene	Fusion	Feature	Proportion of fusion+ samples (%)	Concordance rate* (%)	Fusion transcript(s) detected
RUNX1	*ETV6::RUNX1*	Commonly occurring	25.7	100	*ETV6::RUNX1; RUNX1::ETV6*
RUNX1	*RUNX1::RUNX1T1*	Commonly occurring	14.3	100	*RUNX1::RUNX1T1*
PML	*PML::RARA*	Commonly occurring	22.9	100	*PML::RARA; RARA::PML*
BCR	*BCR::ABL1*	Commonly occurring	11.4	100	*BCR::ABL1; ABL1::BCR*
CBFB	*CBFB::MYH11*	Commonly occurring	11.4	100	*CBFB::MYH11*
MECOM [EVI1]	-	Leads to *MECOM* overexpression	2.9	100	-
KMT2A	*KMT2A::MLLT3*	Promiscuous target gene	5.7	100	*KMT2A::MLLT3*
KMT2A	*KMT2A::MLLT10*	Promiscuous target gene	2.9	100	*KMT2A::MLLT10*
KMT2A	*KMT2A::AFF1*	Promiscuous target gene	2.9	100	*KMT2A::AFF1*

Negative controls	Proportion of control samples (%)	Concordance rate* (%)
Untargeted fusions	13.3%	100
Fusion-negative samples	86.6%	100

Table 1. Concordance of SureSeq Myeloid Fusion with orthogonal methods
*with FISH/qPCR/G-band orthogonal testing

MECOM rearrangements involving inv(3)/t(3;3), which are included in the WHO classification for AML, are rare recurrent cytogenetic abnormalities found in a subset of AML cases and are associated with poor prognosis. The inv(3)/t(3;3) *MECOM* rearrangements are unique as they do not form fusion genes, rather they 'hijack' the *GATA2* enhancer resulting in *MECOM* overexpression. Our assay successfully detected a *MECOM* inv(3)/t(3;3) rearrangement in our sample cohort (Table 1) for which overexpression was characterised relative to housekeeping genes while negative controls showed no *MECOM* expression above our detection threshold (Figure 3).

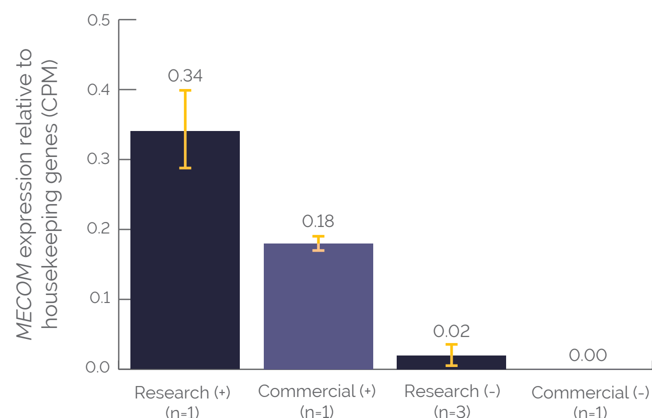


Figure 3. Confident detection of *MECOM* Expression with SureSeq

Figure 3 shows the consistent and confident detection of *MECOM* overexpression in research and commercial samples (including positive and negative controls). *MECOM* expression is normalised to the expression of housekeeping genes and expression values are calculated as counts per million (CPM). 'Research (+)' refers to the research sample containing *GATA2::MECOM* (inv(3)/t(3;3)) noted in Table 1 (n=1). 'Commercial (+)' refers to Universal Human Reference RNA (UHRR) used as positive control (n=1). 'Research (-)' refers to blood extracted RNA with no *GATA2::MECOM* (inv(3)/t(3;3)) (n=3). 'Commercial (-)' refers to normal human lymphocyte RNA used as negative control (n=1). Error bars represent standard deviation.

The SureSeq Myeloid Fusion Panel facilitates partner-agnostic fusion detection as shown by multiple *KMT2A* rearrangements tested in this study (Figure 3). We tested 3 of the most recurrent *KMT2A* rearrangements in AML: *KMT2A::MLLT3*; *KMT2A::AFF1*; *KMT2A::MLLT10* and successfully detected all three fusions. While the *KMT2A* breakpoint was located consistently in exon 8 or 9; the breakpoints in partner genes are more variable (Figure 4). Our assay identified exact location of breakpoints which is especially important for establishing target(s) for monitoring of measurable residual disease (MRD), identifying changes in gene expression and altered splicing and/or polyadenylated events.[7,8]

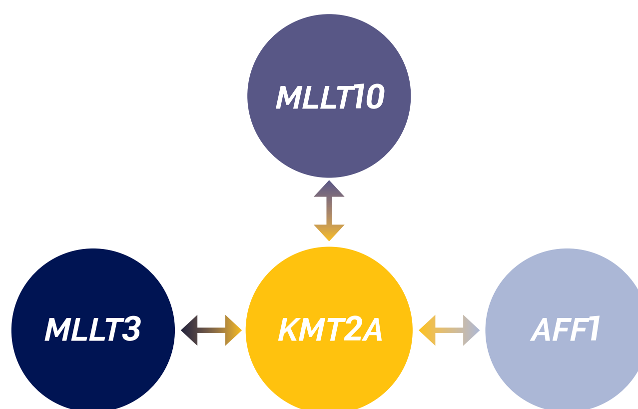
Figure 4 highlights the partner-agnostic fusion detection capability of the SureSeq Myeloid Fusion Panel with breakpoints identified in 3 *KMT2A* rearrangements: *KMT2A::MLLT3*; *KMT2A::AFF1*; *KMT2A::MLLT10*.

Our ability to detect multiple breakpoints and multiple fusions in a single run further highlights the benefits of the SureSeq Myeloid Fusion Panel over FISH or single-gene assays like qPCR, allowing users to combine multiple tests into a single assay.

Discussion

Traditionally, detection of fusion events in research sample cohorts has been challenged by the need for individual testing of samples as well as the need to focus on a select group of fusions for efficiency.

Our research shows that the SureSeq Myeloid Fusion Panel is highly accurate, detecting all known fusion events in positive samples and no detection in negative samples. Our results highlight the confidence researchers can attain in our panel when replacing their existing FISH, qPCR and/or G-band with the SureSeq Myeloid Fusion Panel, and thereby negating the need for individual testing of samples by enabling the simultaneous detection of multiple fusion genes within a single assay. Additionally, the ability to simultaneously detect *MECOM* overexpression serves to further streamline user's workflows by easily folding this detection into the existing workflow.



Fusion detected	Gene	Exon/Intron	Location
KMT2A::MLLT3	*KMT2A*	Intron 8	118482496
KMT2A::MLLT3	*KMT2A*	Exon 9	118484312
KMT2A::AFF1	*KMT2A*	Exon 9	118482496
KMT2A::AFF1	*KMT2A*	Exon 9	118484184
KMT2A::MLLT10	*KMT2A*	Exon 9	118482496
KMT2A::MLLT3	*MLLT3*	Exon 6	20,365,744
KMT2A::MLLT3	*MLLT3*	Exon 6	20365744
KMT2A::MLLT10	*MLLT10*	Exon 10	21670449
KMT2A::AFF1	*AFF1*	Exon 5	87084120
KMT2A::AFF1	*AFF1*	Exon 4	87047594

Figure 4. Partner- agnostic fusion detection for multiple breakpoints

Ready to experience the benefits of the SureSeq Myeloid Fusion Panel?

Learn more about the SureSeq Myeloid Fusion Panel

References

[1] Taniue K and Akimitsu N. Noncoding RNA 2021;7:10. doi: 10.3390/ncrna7010010 [2] Birdwell C *et al.* Blood Cancer J 2021;11:64. doi: 10.1038/s41408-021-00457-9 [3] Thol F. Haematologica 2022;107:44–45. doi: 10.3324/haematol.2021.278983 [4] Kerbs P *et al.* Haematologica 2022;107:100–111. doi: 10.3324/haematol.2021.278436 [5] Matsukawa T and Aplan PD. Stem Cells 2020;38:1366–1374. doi: 10.1002/stem.3263 [6] van Dongen J *et al.* Leukemia 1999;13:1901–1928. doi: 10.1038/sj.leu.2401592 [7] Sanz M A *et al.* Blood 2009;113:1875–1891. doi: 10.1182/blood-2008-04-150250 [8] Hoogstrate Y *et al.* GigaScience 2021;10;giab080. doi: 10.1093/gigascience/giab080

Oxford Gene Technology
A Sysmex Group Company

What binds us, makes us.

Ordering information
UK +44 (0) 1865 856800
US +1 914 467 5285
contact@ogt.com
ogt.com

Oxford Gene Technology Ltd., Unit 5, Oxford Technology Park, 4A Technology Drive, Kidlington, Oxfordshire, OX5 1GN, UK

© Oxford Gene Technology IP Limited – 2024. All rights reserved.
SureSeq™: For research use only; not for use in diagnostic procedures.