Enhanced Sensitivity for Fusion Gene Detection in AML with QIAseq™ RNAscan and Next Generation Sequencing

Song Tian[1], Michelle Baird[1], Frank Reinecke[1], John DiCarlo[1], Raed Samara[1], Hong Yuen Wong[2], Christopher S. Hourigan[2] and Eric Lader[1]

[1] QIAGEN Sciences Inc., Frederick, MD, 21703

[2] Myeloid Malignancies Section, Hematology Branch, National Heart, Lung and Blood Institute, NIH, Bethesda, MD 20892

Abstract

Acute myeloid leukemia (AML) is genetically heterogeneous and is often characterized by chromosomal rearrangements that produce fusion proteins with aberrant transcriptional regulatory activities. Reliable detection of fusion genes as a biomarker in AML is an ongoing effort, both for determining prognostics and tracking response to treatment. Metaphase cytogenetics, fluorescence in situ hybridization (FISH), immunohistochemistry (IHC) and RT-PCR have been widely used for fusion detection with varying levels of success. Unfortunately, significant limitations exist for current methods, such as subjective results and no breakpoint/fusion partner characterization limit utility.

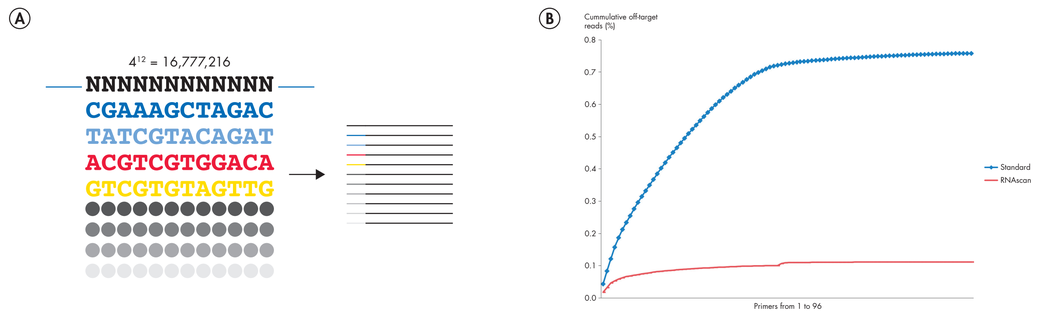
In recent years, next generation sequencing (NGS) has become a powerful tool for fusion gene study. Though several targeted library generation methods have been developed and are commercially available, significant improvements for sample input limitation and detection rate are required to satisfy the need for clinical diagnostics.

We have developed a comprehensive targeted RNA sequencing workflow, making use of unique molecular indices (UMI), single primer extensions (SPE) and a highly optimized bioinformatics pipeline for sensitive and accurate targeted fusion transcript detection in RNA.

AML samples previously characterized by cytogenetic testing were evaluated with the QIAseq Targeted RNAscan Human Oncology Panel to confirm high concordance with clinical metaphase cytogenetic testing. With RNAscan NGS, the existence of chromosome translocations were confirmed, and additional characterization of the AML fusions, such as CBFB/MYH1, KMT2A/MLLT3 RNA fusion junction information was obtained. By reducing the input of RNA sample to as little as 1 ng, we demonstrate the high sensitivity of QIAseq Targeted RNAscan for fusion detection and its value as a tool for RNA fusion transcript studies.

Dealing with NGS Challenges with UMI and SPE

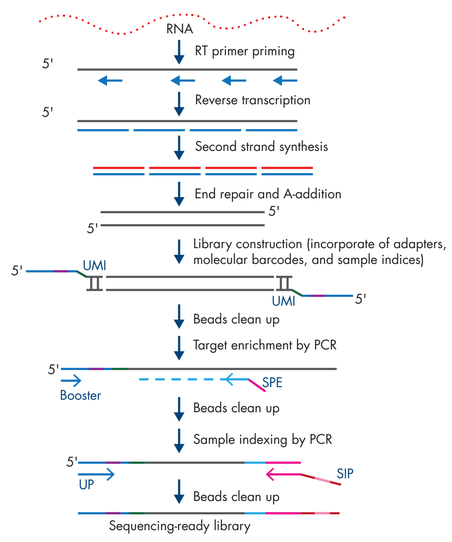
Targeted NGS has been widely used for the detection and confirmation of genetic changes. However, there are technical challenges when relying on precise quantification and characterization, related to clinical decision making. Library construction bias, errors in amplification and sequencing errors can confound results. Similarly, multiplex PCR and nested PCR have primer design and specificity challenges for amplicon-based targeted enrichment. Dealing with all these challenges, QIAseq Targeted RNAscan incorporates a UMI error-correction strategy for PCR and sequencing errors (1) and SPE technology for flexibility and high-specificity performance for target enrichment. The concept of UMI is that prior to any amplification, each original target molecule is 'tagged by' a unique 'barcode' sequence. This DNA sequence must be long enough to provide sufficient permutations to assign each founder molecule a unique barcode. In the current implementation, a 12-base random sequence provides 4^{12} = 16,777,216 unique molecular tags for each target molecule in the sample.



Key technologies used in QIAseq Targeted RNAscan. Ⓐ UMI structure and assignment. UMI analysis recognizes reads with the same barcode and counts them as replicates. Ⓑ Comparison of standard primer design and chemistry (blue line) and SPE primer design and chemistry (red line) on accumulative nonspecific signal level (Y axis). SPE primer design and chemistry provides the extraordinary specificity for target of interest enrichment.

QIAseq Targeted RNAscan Workflow

QIAseq Targeted RNAscan panels use two advanced technologies – single primer extension (SPE) and unique molecular indices (UMI) in NGS. Together, they help to identify and characterize fusion gene events at the RNA level, with high sensitivity and flexibility. The whole workflow and related chemistry are optimized to provide highly efficient RNA conversion, gene-specific target enrichment and library amplification for sensitive fusion gene detection.



QIAseq Targeted RNAscan Panel workflow. The figure shows the smooth workflow of RNAscan from sample to library sequencing. A free cloud-based RNAscan data analysis portal is available at **qiagen.com** for fusion calling after sequencing data is generated. **UMI**: Unique molecular indices. **SPE**: Single primer extension for target enrichment with genespecific primers. **SIP**: Sample index primer. **UP**: Universal PCR primer.

Detection of Fusion in AML Samples

AML is a heterogeneous collection of myeloid malignancies with diverse genetic etiology (2). Precision medicine needs personalized therapy with the detection of informative genetic changes – including gene fusion. Eight AML samples were used for delicate fusion detection with QIAseq Targeted RNAscan. The input was 40 ng total RNA (RIN 3.5–8.1). The standard RNAscan protocol was used with a Human Oncology Panel. An Illumina® MiSeq® was used for sequencing and Illumina NextSeq® 500 was used to confirm the fusion call with higher read depth.

Ⓐ

	1	2	3	4	5	6	7	8
RIN	5.8	3.5	4.9	5.2	6.6	8.3	7.6	8.1

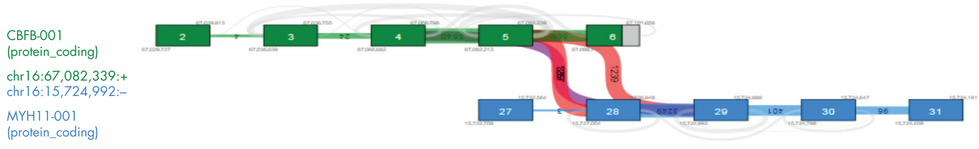
Ⓑ

	1	2	3	4	5	6	7	8
Fusion	VTI1A-TCF7L2	CBFB-MYH11	DEK-NUP214	DEK-NUP214	RUNX1--RUNX1T1	KMT2A--MLLT3	CBFA2T3--PABPN1L	CBFB--MYH1

Ⓒ

						Crossing			Spanning		
Role	**Gene(s)**	**5' feature**	**5' genomic location**	**3' feature**	**3' genomic location**	**Fragments**	**Tags**	**Sizes**	**Fragments**	**Tags**	**Sizes**
Fusion	CBFB-MYH11	Exon 5	Chr 16 67,082,339 +	Exon 29	Chr 16 15,724,992 –	10,172	2835	173	0	0	0
Normal 5'	CBFB	Exon 5	Chr 16 67,082,339 +	Exon 6	Chr 16 67,098,710 +	3162	875	278	4227	1212	305
Normal 3'	MYH11	Exon 28	Chr 16 15,726,848 –	Exon 29	Chr 16 15,724,992 –	11,967	3255	395	10,244	2980	311

Ⓓ



Fusion detection of AML samples. Ⓐ Sample RIN number. Ⓑ AML samples fusion call. Four matched with the samples with clear cytogenetics data, one sample without clear cytogenetic changes was detected with fusion, two samples with negative or complicated cytogenetic changes showed read-through fusions, only one sample showed a different translocation result (RUNX1-RUNX1T1), but the NGS evidence is strong and clear. Ⓒ Representative fusion report after cloud-based data analysis. Fusion breakpoint information and reference transcript information are provided. Ⓓ Representative image shows relevant exon–exon links supported by the largest number of UMI (molecular tags).

Successful Fusion Detection with Low-Input Samples

Fusion detection with low input is one of the challenges for many studies. The limitation of sample availability and limitation on assay efficiency in many detection products blocked fusion discovery and validation in related diseases.

Though RNAscan has a general input limit of 10–250 ng, a low-input test was performed using four AML samples with clear cytogenetic evidence. 10 ng and 1 ng per reaction was used with the RNAscan low-input protocol and the Human Oncology Panel. MiSeq was used for sequencing with paired-end reads. As fusions are called from RNA in multiple steps, handling quality and instrument error can affect the results. Integrating UMI and a barcode-aware algorithm into fusion calling allows correction for PCR amplification bias, thereby identifying quality issues with internal control assays, while providing high sensitivity for fusion call.

Ⓐ

Cytogenetic details	Fusion (10 ng)	Reads/UMI/Size (10 ng)	Fusion (1 ng)	Reads/UMI/Size (1 ng)
Inv16	CBFB-MYH11, t(16;16)	2461/563/171	CBFB-MYH11, t(16;16)	3755/121/59
t(6;9)	DEK-NUP214, t(6;9)	2040/747/77	DEK-NUP214, t(6;9)	1502/101/28
t(6;9)	DEK-NUP214, t(6;9)	2174/634/86	DEK-NUP214, t(6;9)	1440/193/39
46,XY,inv(16)(p13.1q22)	CBFB--MYH1, t(16;6)	2492/1042/251	CBFB--MYH1, t(16;6)	1880/146/75

Ⓑ

	S1	S2	S3
Reads	2586958	2717265	9725916
Fusion	Neg	Neg	KMT2A-MMLT3
Average reference control UMI	14	27	18947

Ⓒ

	S1	S2	S3
Reads	1864946	1871238	1347626
Fusion	DEK-NUP214	DEK-NUP214	KMT2A-MMLT3
Average reference control UMI	805	946	700

Successful fusion calling with low input challenge. Data showed that with 10 ng, or even 1 ng input, RNAscan still provided consistent detection of fusion with strong evidence that matched the cytogenetic data. Ⓐ Data showed successful fusion calling with RNAscan for as little as 1 ng input RNA with good efficiency and results matching cytogenetic data. Ⓑ Identification of false-negative results with molecular barcode count. Different samples were processed with the Human Oncology Panel and results showed two negative and one positive gene fusions. By checking the UMI count for the reference target control, two negative samples were recognized as potential false-negatives, as the extreme low MTs indicating an input issue or handling error for library generation for samples S1 and S2. Ⓒ Confirmation of false-negative results in Ⓑ. The same samples were processed again with adjusted input and standard library generation procedure with fixed instrument. All the samples were identified with gene fusions, and all results matched the samples' cytogenetic data.

Conclusions

- The QIAseq Targeted RNAscan Human Oncology Panel provides a powerful tool for fusion gene detection with NGS. It detects gene fusions with good sensitivity and provides detailed fusion gene information compared with traditional methods, such as cytogenetics by metaphase karyotyping.
- With UMI and SPE technologies, the optimized RNAscan workflow and data analysis tool provide:
  - Confidence in fusion calling – after removing PCR and sequencing errors
  - Flexibility for fusion detection primer design – for highly efficient target enrichment.
- For AML fusion detection, RNAscan provides good tolerance and sensitivity for low-input fusion detection.
- A customized AML fusion detection panel will improve AML-focused fusion detection and biomarker discovery, with a higher sensitivity in fusion detection, and has potential utility in dealing with challenges such as determination of remission status.

Acknowledgements

- This research was supported in part by the Intramural Research Program of the NIH, National Heart Lung and Blood Institute.
- Quan Peng and Yexun Wang for technologies development.

Abstract ID: 914

References

1. Peng, Q., Vijaya Satya, R., Lewis, M., Randad, P., Wang, Y. (2015) Reducing amplification artifacts in high multiplex amplicon sequencing by using molecular barcodes. BMC Genomics **16**, 589.
2. Lai, C., Karp, J.E., Hourigan, C.S. (2016) Precision medicine for acute myeloid leukemia. Expert Rev. Hematol. **9**, 1.

1106178 02/2017

Sample to Insight