APPLICATION NOTE

Evaluation of a targeted RNA sequencing workflow for variant calling from FFPE samples in non-small cell lung cancer (NSCLC)

BACKGROUND AND OVERVIEW

In the oncology field, many research laboratories are moving beyond just interrogating the genome for mutational analysis and seeking more dynamic biomarkers. RNA-based workflows paired with analytical solutions for molecular analysis of tumor-derived samples provide a deeper understanding of tumor biology in the context of non-small cell lung cancer (NSCLC).

The typical biopsy specimen taken from a patient is usually a very limited amount of material for molecular testing. To improve outcomes, it is critical to provide a thorough characterization with limited material available.

At the medical center AZ Sint-Lucas, located in Ghent, Belgium, molecular biologists Koen Jacobs, Ph.D. and Katrien De Mulder, Ph.D. are working to improve the molecular analysis of NSCLC samples by developing RNA-based workflows. In a feasibility study, they retrospectively evaluated Twist's RNA analysis solutions in their laboratory. Most samples are excised pieces of tumor tissue obtained as formalin-fixed paraffin-embedded (FFPE) blocks, a sample type that is ideal for histological analysis, but often presents a challenge to molecular analysis as the genetic material can be degraded or highly crosslinked. Additionally, the amount of tissue available can frequently be limited which can make it refractory to use in an NGS workflow[1,2]. As a whole, NSCLC presents a challenging use case due to the difficulty of working with these fixed specimens for NGS analysis[3]. Therefore, an RNA sequencing workflow should be capable of working with low sample input, while also enabling the sensitive discovery of both expected and unexpected variants.

The AZ Sint-Lucas team had some key requirements for any solution that they would implement in their lab. First and most importantly, they required a library preparation method that would minimize the introduction of artifacts into sequencing data that could potentially lead to a spurious conclusion. Second, the solution had to robustly detect low-expressing genes. In some cases, they would be testing samples in which there is a low percentage of neoplastic cells relative to normal cells, and therefore sensitive detection of low-expressing genes is needed to boost their overall ability to detect the tumor's expression profile. Finally, the solution they were seeking needed to be able to detect not just the genetic variation they were expecting but also identify novel changes such as alternative splicing events and novel gene fusion variants.

TESTIMONIAL

"The transcriptome provides closer access to the real biological status of the tumor cell, in contrast to the more 'static' nature of DNA. The parallel analysis of gene expression, fusion detection, variant calling, tumor signatures, and much more, possible with one library prep and starting from three to five FFPE tissue slices makes the RNA exome kit a promising assay for future-proof implementation into our clinical routine."

Koen Jacobs, Molecular Biologists, AZ Sint-Lucas Ghent

RNA-SEQ ANALYSIS WORKFLOW AND METHODOLOGY

Target Enrichment With a Custom RNA Panel Design

Twist's Targeted RNA workflow was evaluated with a Twist custom RNA target enrichment panel based on targets defined by the AZ Sint-Lucas lab. The custom RNA panel, designated "RNA-O", was targeted against 272 cancer-related genes relevant to NSCLC (**Table 1**). The custom RNA panel design was targeted to coding regions and is therefore not expected to be capable of detecting regions that are not expressed into RNA, such as gene promoters or some PGx variants located in non-transcribed regions. It should also be noted that the RNA panel incorporates Twist's novel "Exon-Aware design strategy" that avoids positioning probes at exon junctions. The advantages of this are a more complete targeting of the coding regions within the transcriptome and the ability to capture novel fusion events which typically present with fusions joined at exon-exon junctions in the fusion transcript.

The performance of the Twist Targeted RNA workflow was compared with two independent workflows that AZ Sint-Lucas routinely uses on NSCLC tumor specimens. The first is a DNA sequencing workflow for tumor variant calling and the second is an RNA workflow specific to fusion detection. The evaluation used the same set of FFPE samples derived from tumor and normal tissue to evaluate the feasibility of a single RNA-based workflow to replace their routine test for DNA (for variant calling) and RNA (for RNA-fusion detection) analysis in NSCLC.

Sample Cohorts

The initial evaluation cohort included 32 FFPE samples: 29 NSCLC samples and 3 derived from normal lung tissue. The tumor samples were selected based on the presence of known consequential variants including missense, indels, CNV, and gene fusions across a range of sample quality. Sequencing results were then compared to that of routine DNA and RNA sequencing to assess variant and fusion transcript detection, respectively. Such a comparison is intended to evaluate the feasibility of using a single RNA sequencing workflow in place of a more complex and resource-intensive approach that involves both DNA and RNA sequencing.

PANEL DESIGN FEATURE	RNA - FUSION PANEL	TWIST CUSTOM "RNA-O" RNA PANEL
Genes Covered	84	272
MSI-Screening	N/A	39
Chromosomal Aberration	N/A	1,921 SNPs
Probe Count	7,279	10,123
Design Target Region Size	0.36 Mb	0.84 Mb

Table 1. Characteristics of RNA panels utilized in the study. The lab utilizes a fusion-specific panel in routine testing along with a DNA-based sequencing panel. The "RNA-O" custom panel has a broader scope of coverage that targets regions of the transcriptome beyond just fusions, including variant, MSI, and chromosomal aberrations.

Library Preparation

Briefly, RNA was extracted from 3-5 FFPE curls of 5-micron thickness and processed for total RNA extraction. Subsequently, all RNA samples were treated with the optional FFPE repair module included in the Twist RNA library preparation kit before proceeding with the standard RNA library preparation protocol which can be completed within 5 hours. A summary of the RNA integrity and extraction yields is shown in Figure 1. Overall, the samples demonstrated a wide range of degradation profiles and yields based on DV200 score and RNA quantification by TapeStation system (Agilent) and Quantus (Promega) fluorometric measurement[4,5].

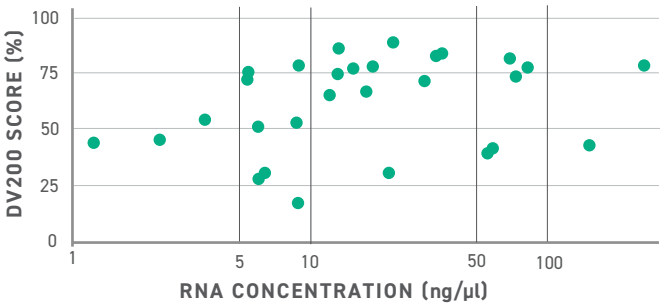

Figure 1. Cohort of FFPE RNA extractions demonstrating a range of RNA fragmentation profiles based on DV200 measurements. A sample with a DV200 calculation of >70% is considered high quality, 50-70% is considered acceptable quality, and 30-50% is considered poor quality (<30% is advised as too poor quality for NGS analysis)[4,5].

Target Enrichment and Sequencing

The prepared libraries were then enriched with the "RNA-O" custom RNA panel overnight with Twist's Standard Hyb v2 protocol and then sequenced across multiple sequencing runs to compare performance against the routine workflow involving DNA and RNA sequencing. All sequencing data was processed with a custom pipeline for data processing and variant calling, the performance of which was compared against prior sequencing analysis results from the same tumor specimens.

RNA WORKFLOW EVALUATION RESULTS

While evaluating Twist's RNA sequencing solutions, the team utilized a baseline of variant and fusion calls from their routine DNA and RNA sequencing workflow. Importantly, neither routine workflow was capable of detecting both variants and fusion transcripts on its own, therefore reliance on the incumbent technology necessitates the use of precious sample material for both RNA and DNA sequencing. The Twist Custom RNA-O panel proved capable of detecting both coding variants and fusion transcripts. The only variants that RNA-O was unable to detect were those that lay outside of coding regions and were therefore not converted into RNA transcripts. Variants were identified in several clinically relevant genes, including *SMARCA4, KRAS, TP53, PIK3CA, MAP2K1, BRAF, RB1, ATM, POLE, ARID1A, SMAD4,* and *MET*. Additionally, transcript variants, including two *EGFR* exon 19 deletions and one *EGFR* exon 20 insertion, were included. Notably, detection of a splice site variant in *KEAP1* (c.1709-1G>T) when RNA-O was used indicates the capture of pre-mRNA.

Alterations were classified as single nucleotide variant (SNV) (n=37), splice site variant (SS) (n=1), insertion or deletion (INDEL) (n=4), germline pharmacogenetic variant (PG) (n=3), gene promoter mutation (GP) (n=2), or fusion gene (F) (n=9). A summary of the results is presented in Table 2. These variants were present in several genes that were adequately detected with "RNA-O". Detection of fusion genes: *ALK, RET, MET* exon 14 skipping, *NTRK3*, and *FGFR3* were also concordant between the custom panel design and the routine DNA workflow.

Concerning pharmacogenetic variants in a coding sequence (one *DPYD*, two *SLCO1B1*), only one variant (*DPYD*) out of three could be detected with the RNA-O panel. However, it is notable that *SLCO1B1* is not known to be expressed in the two sample tissues where no *SLCO1B1* RNA coverage was present as shown in Human Protein Atlas[6], therefore this result is potentially spurious because this gene is not known to be normally expressed within the lung. Two gene promoter mutations are investigated with the solid tumor DNA panel: *TERT* and *UGT1A1*. Since both variants are not part of a coding sequence, none of these variants could be detected with the RNA-O panel. While this may appear to be a limitation of the panel, the mutational status of the promoter is not a critical clinical finding in the context of tumor profiling in NSCLC. Additionally, the mutational status of a promoter is typically utilized to infer the aberrant expression of a gene. In the context of an RNA-based analysis, the expression of the gene itself can be measured and utilized as an alternative marker to promoter sequence analysis. Overall, the custom RNA panel designed by Twist shows strong concordance with the routine DNA and RNA panels regarding tumor-associated variant calling across a range of variant classes including detection of fusion events.

The use of the custom RNA target enrichment panel that includes targets for both mutation detection and fusion calling makes for a fast and reliable sequencing workflow. Given that all coding genes are included and tumor-specific expression patterns could be identified in the context of signatures, this approach is highly relevant and future-proof to the investigation of additional genetic biomarkers.

VARIANT CLASS	ROUTINE TESTING		TWIST RNA EVALUATION	CONCORDANCE OF VARIANTS CALLS
	DNA	RNA	CUSTOM RNA PANEL (RNA-O)	
Single Nucleotide Variant (SNV)	37	N/A	37	37/37
Splice Site Variant (SS)	1	N/A	1	1/1
Insertion / Deletion (INDEL)	4	N/A	4	4/4
Germline Pharmacogenetic Variant (PG)	3	N/A	1	1/3
Gene Promoter Mutation (GP)	2	N/A	N/A	0/2
Fusion Variant (F)	N/A	9	9	9/9

"N/A" indicates that the workflow is not designed to capture or target these variants. See text for additional details.

Table 2. Summary metrics on the concordance of the Twist Custom RNA enrichment panels vs. the standard stand-alone routine testing protocols for DNA sequencing or RNA sequencing. The Twist enrichment workflow with the RNA-O panel demonstrates concordance with variant classes that are targeted with the routine panel. For detection metrics amongst the 9 known fusions detected by routine RNA testing, the RNA-O panel design was able to also detect the fusions in addition to other variant classes.

REFERENCES

1. De Maglio, G. et al. The storm of NGS in NSCLC diagnostic-therapeutic pathway: How to sun the real clinical practice. Crit. Rev. Oncol. Hematol. 169, 103561 (2022).

2. Penault-Llorca, F. et al. Expert opinion on NSCLC small specimen biomarker testing — Part 1: Tissue collection and management. Virchows Arch. 481, 335-350 (2022).

3. Arreaza, G. et al. Pre-Analytical Considerations for Successful Next-Generation Sequencing (NGS): Challenges and Opportunities for Formalin-Fixed and Paraffin-Embedded Tumor Tissue (FFPE) Samples. Int. J. Mol. Sci. 17, 1579 (2016).

4. Matsubara, T. et al. DV200 Index for Assessing RNA Integrity in Next-Generation Sequencing. Biomed Res. Int. 2020, 9349132 (2020).

5. Evaluating RNA Quality from FFPE Samples. Illumina https://www.illumina.com/content/dam/illumina-marketing/documents/products/technotes/evaluating-rna-quality-from-ffpe-samples-technical-note-470-2014-001.pdf (2016).

6. Thul, P. J. et al. A subcellular map of the human proteome. Science. 356, 6340 (2017).