A UNIFIED AND STREAMLINED TARGETED SEQUENCING SYSTEM FOR THE QUANTIFICATION OF DNA MUTATIONS AND RNA EXPRESSION MARKERS IN LUNG CANCER

Gary J Latham[1], Julie Krosting[1], Michael Dodge[1], Robert Zeigler[1], Liangjing Chen[1], Jason Plyler[1], Shobha Gokul[1], Junya Fujimoto[2], Vassiliki A Papadimitrakopoulou[3], Ignacio I Wistuba[2], Richard Blidner[1], and Brian C Haynes[1]
[1]Asuragen, Inc., Austin, TX, [2]Department of Translational Molecular Pathology, [3]Department of Thoracic/Head and Neck Medical Oncology, Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX

SUMMARY

- Both DNA and RNA variants can drive cancer initiation and progression, and provide targets for precision medicine.
- A unified NGS technology is needed to identify "driver" variants from low-quality and low-quantity tumor biopsies.
- We describe two targeted NGS workflows that can report DNA mutations, indels, RNA expression markers and fusions in NSCLC, and assess the methodology with 112 residual clinical FFPE specimens.
- Our data demonstrate that a novel workflow that unifies DNA and RNA sequencing can unveil multi-omic data from a single TNA specimen, resulting in faster and simpler sample processing and more comprehensive analysis options, particularly from mutation-negative samples.

INTRODUCTION

The promise of precision medicine relies on the identification of DNA and RNA markers that can individualize patient management. Methods such as next-generation sequencing (NGS) can deduce DNA or RNA sequences, but both types of nucleic acid have not been efficiently and effectively combined into a single NGS workflow. We describe a comprehensive methodology for targeted clinical NGS that reports DNA and RNA variants, provides a streamlined workflow, and accommodates low-input total nucleic acid (TNA) from challenging clinical specimens.

MATERIALS AND METHODS

FFPE specimens used in this study were provided by MDACC, or procured through tissue banks (Asuragen and Asterand). Sample QC was performed using a novel multiplex qPCR assay that quantifies discrete populations of amplifiable DNA (QuantideX® qPCR QC Assay) or RNA (prototype RT-qPCR assay), or both, from TNA material. Cancer-relevant sequences were enriched using multiplex PCR followed by tag PCR to append sample-specific barcode sequences and NGS adaptor sequences. Sample libraries were then purified using magnetic bead-based size selection and pooled and normalized based on library yield (QuantideX® NGS reagents, Asuragen). After 2x220 sequencing on the Illumina MiSeq, each demultiplexed library was analyzed using a novel bioinformatics pipeline (QuantideX® NGS Reporter) informed by the pre-analytical copy number of each sample to improve the calling of DNA point mutations, indels, structural variants, RNA expression and gene fusions.

Presented at AACR 2016

RESULTS

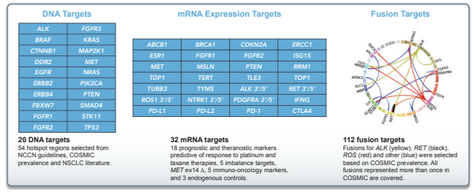


Figure 1. DNA and RNA content for the NSCLC lung cancer panels evaluated in this study.

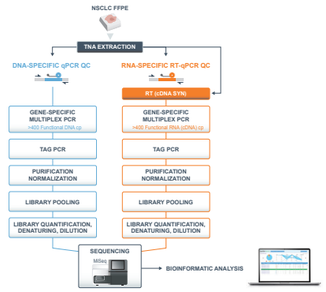


Figure 2. Parallel workflow for targeted DNA and RNA NGS of lung cancer FFPE specimens. A single total nucleic acid (TNA) sample was used to query both DNA and RNA targets. A qPCR assay provided quantification of the functional (amplifiable) copy number of DNA, whereas a separate RT-qPCR assay was used to quantify functional RNA using a marker demonstrated to have stable expression across lung cancer and normal lung. At least 200-400 amplifiable copies of each nucleic acid from TNA was input into the RT or PCR steps that were processed in parallel workflow streams through NGS.

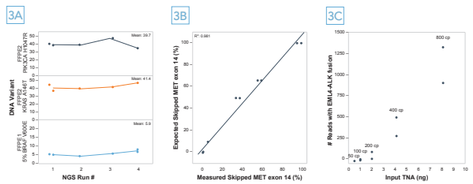
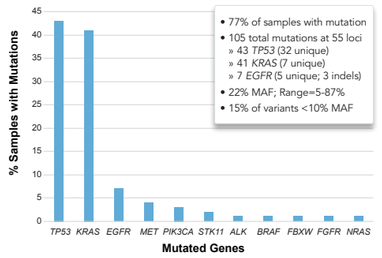


Figure 3. Analytical assessment of targeted DNA and RNA NGS. A) Evaluation of residual clinical FFPE samples across 4 successive NGS runs demonstrated consistent quantification of known DNA variants down to at least 5% MAF. Variants were quantified with 10% (40% MAF), 11% (41% MAF), and 23% CV (5.8% MAF; 5% by ddPCR). Inputs were 400 amplifiable copies in each case, which translated to ~10 ng of bulk DNA. **B)** MET exon 14 skipping was quantified above background down to a 1% mass fraction. **C)** A known EML4-ALK fusion was detected down to <1 ng TNA (50 copies of amplifiable RNA) in an FFPE sample.

4A



4B

Sample	Fusion	% Reads/cntrl Run1	% Reads/cntrl Run2	3'/5' Imbalance	CE Confirmation
AD54	EML4:ALK	8%	10%	Yes	No
AD57	EZR:ROS1	44%	17%	No	Yes
AD58	CD74:ROS1	81%	51%	No	Yes

Figure 4. Analysis of 112 NSCLC FFPE specimens reveals a spectrum of driver mutations in DNA and RNA. A total of 73 adenocarcinoma and 39 squamous cell carcinomas were sequenced for DNA and RNA variants. QC analysis of DNA isolated from two 5 um FFPE sections yielded a median of 369 cp/ul (50 ul elution) with 82% providing >100 cp/ul. A similar analysis of the RNA fraction showed a median of 237 bp/ul with 80% providing >50 cp/ul. Analysis of both DNA and RNA identified 108 COSMIC variants present at >5% MAF, including three fusions. All variants were confirmed by replicate testing or testing using an orthogonal assay.

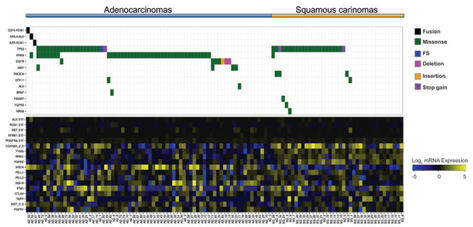
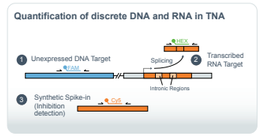


Figure 5. Multi-omic characterization of NSCLC FFPE specimens: 73 adenocarcinomas (AD) and 39 squamous (SQ). Variants associated with each specimen were cataloged by type, and mRNA expression from imbalance genes, housekeeping controls or other targeted transcripts were coded into the heatmap shown. The profile of variants was distinctive to each tumor type (AD or SQ), consistent with results from large-scale studies such as TCGA. Note that specimen AD-54, which was positive for an EML4-ALK fusion, showed a strong 3'/5' imbalance of *ALK*. Note that AD42, AD72, AD74, and AD75 were lung metastases from colonic primary tumors. SQ09 was characterized as "poorly differentiated NSCLC with squamous cell carcinoma features" and SQ28 was indicated as "poorly differentiated NSCLC with adenocarcinoma features".

6A



6B

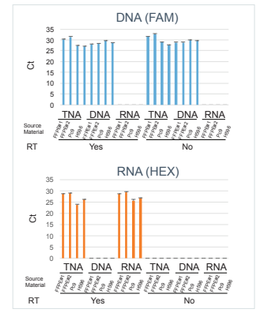


Figure 6. Independent quantification of functional DNA and RNA from TNA in a single-tube multiplexed RT-qPCR assay. A) Specific DNA and RNA populations were quantified using a novel DNA target that was completely absent RNA expression, and a separate RNA target that was stably expressed across lung cancers and normal lung tissue. Multiplexed RT-qPCR primer and probe sets were engineered to avoid DNA detection in RNA (via an exon-spanning design) with independent amplification of the unexpressed DNA target. An anti-genomic sequence was included to monitor any functional inhibition from the nucleic acid isolate. **B)** Real-time RT-PCR data demonstrated the specificity of DNA or RNA detection.

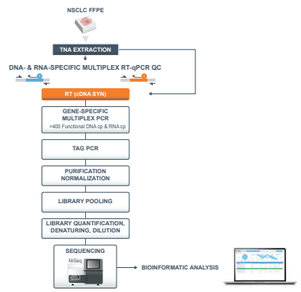


Figure 7. A unified workflow for concurrent targeted DNA and RNA NGS of lung cancer FFPE specimens. A single total nucleic acid (TNA) sample was used to query both DNA and RNA targets in a 186-plex single-tube PCR. Functional DNA and RNA quantification was achieved using the assay described in Figure 6. A total of 289 primers were included in the PCR step to achieve simultaneous amplification of DNA and RNA targets without the need for separate process streams.

8A

Sample ID	Type	Nucleic Acid: TNA	DNA	RNA	Mutation or Fusion	Detected by NGS	NGS Panel(s) with Detection
PC9	Cells	Yes	-	-	EGFR.746_750delEREA	Yes	DNA+RNA*
H2228	Cells	Yes	-	-	EML4-ALK	Yes	RNA & DNA+RNA
H2228	Cells	-	Yes	-	EML4-ALK	No (expected)	
H2228	Cells	-	-	Yes	EML4-ALK	Yes	RNA & DNA+RNA
RT112	Cells	Yes	-	-	FGFR3:TACC3	Yes	RNA & DNA+RNA
RT112	Cells	-	Yes	-	FGFR3:TACC3	Yes	DNA & DNA+RNA
RT112	Cells	-	-	Yes	FGFR3:TACC3	Yes	RNA & DNA+RNA
LC2	Cells	Yes	-	-	CCDC6-RET	Yes	DNA+RNA*
2044	FFPE	Yes	-	-	PIK3CA.E542K, BRAF.V600E	Yes	DNA & DNA+RNA
2046	FFPE	-	Yes	-	PIK3CA.E542K, BRAF.V600E	Yes	DNA & DNA+RNA
2051	FFPE	-	-	Yes	PIK3CA.E542K, BRAF.V600E	No (expected)	
2057	FFPE	Yes	-	-	PIK3CA.H1047	Yes	DNA & DNA+RNA
2059	FFPE	-	Yes	-	PIK3CA.H1047	Yes	DNA & DNA+RNA
2066	FFPE	-	-	Yes	PIK3CA.H1047	No (expected)	
2067	FFPE	Yes	-	-	KRAS.G12C, PIK3CA.E545K	Yes	DNA+RNA*
2182	FFPE	Yes	-	-	EML4-ALK	Yes	DNA+RNA*
2189	FFPE	Yes	-	-	KRAS.Q61L	Yes	DNA+RNA*
H2228/2044 (3:7)	Blend	Yes	-	-	EML4-ALK; PIK3CA.E542K, BRAF.V600E	Yes (EML4-ALK)	RNA & DNA+RNA
H2228/2044 (3:7)	Blend	Yes	-	-	EML4-ALK; PIK3CA.E542K, BRAF.V600E	Yes (PIK3CA, BRAF)	DNA & DNA+RNA
H2228/2044 (7:3)	Blend	Yes	-	-	EML4-ALK; PIK3CA.E542K, BRAF.V600E	Yes (EML4-ALK)	RNA & DNA+RNA
H2228/2044 (7:3)	Blend	Yes	-	-	EML4-ALK; PIK3CA.E542K, BRAF.V600E	Yes (PIK3CA, BRAF)	DNA & DNA+RNA

8B

Figure 8. A unified NGS technology for co-amplification of DNA and RNA produces sensitive and specific results. RNA, DNA, and/or TNA with known RNA fusions or DNA mutations were extracted from 5 FFPE and 4 cell-line samples. Libraries were created using the DNA/RNA workflow shown in Figure 7 and compared in several cases with the matching DNA or RNA panels alone. **A)** All known DNA and RNA variants were detected (2.2-86% MAF). RT112, which harbors an FGFR3-TACC3 fusion, was detected in both the RNA and DNA (ie, no RT step), consistent with the primer design. *Only panel that was run. **B)** Quantification of each mutation demonstrated high specificity across the sample set.

CONCLUSIONS

- High performance targeted DNA and RNA panels were developed that can detect low-abundance DNA mutations and indels, transcripts with skipped exons or aberrant expression, and gene rearrangements of clinical relevance to NSCLC from low-input FFPE DNA, RNA, or TNA.
- Parallel DNA and RNA sequencing produced profiles of somatic cancer drivers from 112 NSCLC FFPE samples consistent with previous studies of adenocarcinoma and squamous cell carcinoma.
- A single-tube 3-plex RT-qPCR assay quantified independent populations of DNA and RNA from a single TNA input, thereby enabling comprehensive guidance of functional DNA/RNA into NGS.
- Feasibility for a unified TNA NGS workflow using 289 primers in single-tube PCR was demonstrated. This technology reported both DNA mutations and RNA fusions from the same TNA isolate with sensitivity and specificity.

THE UNIVERSITY OF TEXAS MD Anderson Cancer Center

Asuragen®
www.asuragen.com