A modular next-generation sequencing technology that couples the detection of RNA structural variants and DNA mutations in lung cancer

Richard Blidner, Shobha Gokul, Brian C Haynes, Huiping Zhu, and Gary J Latham
Asuragen, Inc., Austin, TX

Summary

- The accurate detection of cancer-related genetic variants from both DNA and RNA by next-generation sequencing (NGS) requires reliable and integrated methods to enable more comprehensive analyses for personalized medicine approaches.
- Pooling NGS libraries across different targeted panels offers workflow flexibility and efficiency to help manage routine analysis.
- Using QuantideX® NGS chemistries, we demonstrate high-performance results in a single sequencing run by combining libraries that target DNA or RNA cancer genes in solid tumors and liquid biopsies.

Introduction

Non-small cell lung cancer (NSCLC) demonstrates remarkable molecular diversity, highlighting the need to detect both RNA and DNA markers for personalized medicine. Mutational categories associated with NSCLC initiation and progression include single-nucleotide variants (SNV), DNA insertions and deletions (indels) and copy number variations (CNV), as well as RNA fusions and post-transcriptional splicing variants. We describe a novel targeted next-generation sequencing (NGS) approach with systems-level integration that achieves independent or combined detection of RNA and DNA variants using optimized workflows and analytics.

Materials and Methods

Total nucleic acid was isolated from 21 residual FFPE NSCLC biopsies and three cancer cell lines that were previously annotated for gene fusion and mutation status. Three targeted NGS panels were evaluated: 1) the prototype QuantideX® NGS DNA Hotspot 21 Kit* (DNA Hotspot 21), 2) the QuantideX® NGS RNA Lung Cancer Kit (RUO)† (RNA Lung), and 3) the prototype QuantideX® NGS Liquid Biopsy‡ (Liquid Biopsy) that can accurately detect variants down to 0.5% MAF (see poster ST115). Seraseq™ ctDNA v2 reference materials (SeraCare) that mimic cfDNA with mutations at 2%, 1%, 0.5%, and 0% MAF (wildtype) were used as inputs for the Liquid Biopsy prototype.

3' Fusion Genes	# of Fusions	3'-5' Imbalance
ALK	53	*ALK*
ROS1	22	*ROS1*
RET	12	*RET*
FGFR3	7	*NTRK1*
NTRK3	3	*PDGFRA*
NTRK1	4	
NRG1	2	
FGFR1	1	**Exon Skipping Event**
FGFR2	1	*MET* e13:e14
MBIP	1	*MET* e14:e15
PDGFRA	1	*MET* e13:e15

Gene	Codon Range	Gene	Codon Range	Gene	Codon Range	Gene	Codon Range
ABL	249-258	*ERBB2*	755-769	*IDH1*	122-134	*NRAS*	9-20
	303-319		774-788	*IDH2*	122-134		55-67
AKT1	16-27		839-847		163-174		110-119
AKT2	16-26		877-883	*JAK2*	607-620		144-150
ALK1	1174-1196	*FGFR1*	123-136	*KIT*	557	*PDGFRA*	560-572
	1274-1278		250-262		815		840-852
BRAF	465-474	*FGFR3*	247-260	*KRAS*	9-20	*PIK3CA*	540-551
	591-612		363-374		55-65		1038-1049
EGFR	486-493		638-653		104-118	*RET*	916-926
	709-722	*FLT3*	829-840		137-148		
	737-761	*HRAS*	9-20	*MET*	1245-12-56		
	767-798		59-76				
	849-861		113-121				

Table 1. QuantideX® NGS Panel Content. Content for the QuantideX® NGS RNA Lung Cancer Kit (ORANGE) includes specific gene fusion breakpoints, 3'-5' imbalances, and *MET* exon 14 skipping. The QuantideX® NGS DNA Hotspot 21 Kit (BLUE) interrogates 46 gene regions within 21 genes. Content for the QuantideX® NGS Liquid Biopsy is not shown (see poster ST115 for more information).

*This product is under development. Future availability and performance cannot be ensured.
†Research Use Only. Not for use in diagnostic procedures.
‡Proof-of-concept data only. The performance characteristics of this assay have not yet been established.

Presented at AMP 2017 - ST109

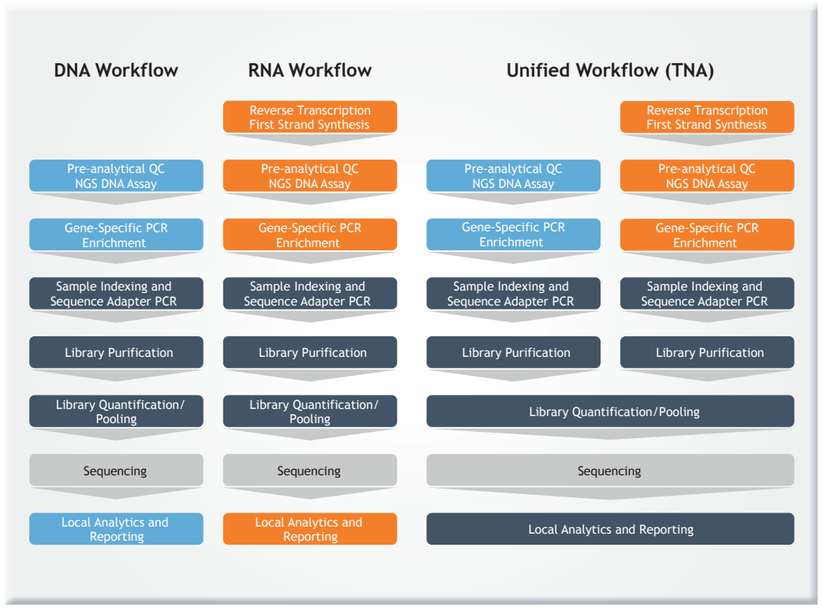

Figure 1. NGS Assay Workflow. DNA and RNA NGS workflows are aligned to enable parallel processing of samples. RNA libraries require one additional step to generate cDNA from RNA or total nucleic acid (TNA: DNA+RNA) isolates. Each assay includes a pre-analytical QC assay to assess the quantity and quality of the input and inform the downstream analysis pipelines. Purified libraries were combined onto a single MiSeq V3 flow cell (Illumina) and sequenced simultaneously. Analysis of the results are performed by separate bioinformatics pipelines bundled within a unified software architecture. Pooling of libraries prior to sequencing was adjusted to achieve a target coverage for each library as well as a target for the total number of reads generated from the flow cell. This strategy was used to yield an equivalent number of reads for the RNA and DNA panels, as well as a tiered approach to yield differences in coverage for each library type (DNA Hotspot 21, RNA Lung, Liquid Biopsy) based on coverage requirements for each assay.

Results

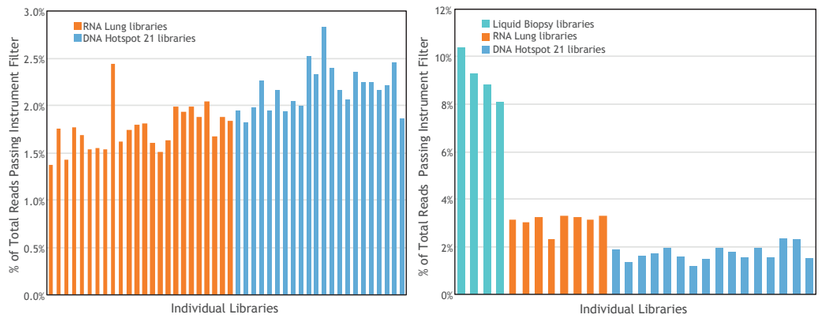

Figure 2. A Library Pooling Strategy Achieves Consistent Coverage Appropriate for Each Library Type. The percentage of the total reads attributed to each library correlated with the relative molar concentration of each library after pooling. For equimolar pooling of DNA and RNA libraries **(LEFT)** coverage was nearly uniform across all libraries (RNA only: 6% CV; DNA only: 11% CV; all libraries: 16% CV). For uneven pooling **(RIGHT)**, coverage of each library was proportional to the ratio of the concentration of each library type to the total pooled library concentration. This approach allows pooling of libraries with greater depth of coverage requirements (Liquid Biopsy targeted 2M reads/library) with libraries that require less coverage (DNA Hotspot 21 targeted 250K reads/library and RNA Lung targeted 500K reads/library) to achieve more efficiency utilization of NGS capacity. For each library type, the recommended minimum coverages were achieved for all libraries.

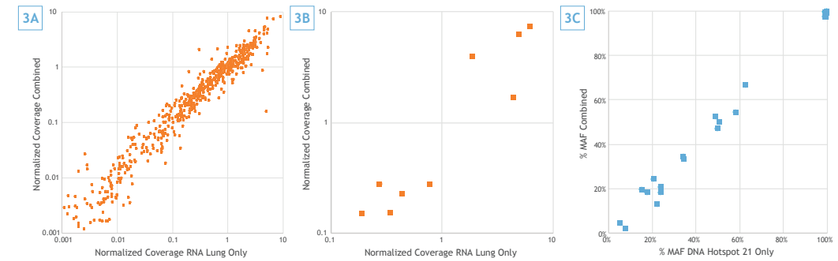

Figure 3. Relative Target-by-Target Coverage and Variant Fraction is Maintained in Multi-Panel NGS Runs. Normalized (within sample) coverage of mRNA expression markers **(A)** and detected fusions and splice variants **(B)** within the RNA Lung libraries was consistent between single-panel and multiple-panel pools. The mutant allele frequency of detected DNA variants **(C)** within the DNA Hotspot 21 libraries was also consistent between single-panel multiple-panel pools.

Sample	Expected RNA	RNA call	Imbalance	Expected DNA	DNA call
# 1	-	-	-	*PIK3CA*	*PIK3CA*
# 2	-	-	-	*BRAF/PIK3CA*	*BRAF/PIK3CA*
# 3	-	-	-	*KRAS/PIK3CA*	*KRAS/PIK3CA*
# 4	-	-	-	*BRAF*	*BRAF*
# 5	-	-	-	*NRAS*	*NRAS*
# 6	-	-	-	*KRAS*	*KRAS*
# 7	*ALK*	*ALK*	*ALK*	-	-
# 8	-	-	-	*KRAS*	*KRAS*
# 9	-	-	-	*PIK3CA*	*PIK3CA*
# 10	*MET*ex14	*MET*ex14	-	-	-
# 11	-	-	-	*EGFR*	*EGFR*
# 12	*MET*ex14	*MET*ex14	-	-	-
# 13	-	-	-	*KRAS*	*KRAS*
# 14	*RET*	*RET*	*RET*	-	-
# 15	-	-	-	-	-
# 16	*NRG1*	*NRG1*	-	*EGFR*	*EGFR*
# 17	-	-	-	*EGFR*	*EGFR*
# 18	-	-	-	*EGFR/PIK3CA*	*EGFR/PIK3CA*
# 19	*RET*	*RET*	*RET*	-	-
# 20	-	-	-	-	-
# 21	-	-	-	*FGFR3*	*FGFR3*
# 22	*FGFR3*	*FGFR3*	-	UNK	*EGFR*
# 23	*MET*ex14	*MET*ex14	-	Not Tested	Not Tested
# 24	*ROS1*	*ROS1*	*ROS1*	Not Tested	Not Tested
NTC	No Call	No Call	No Call	No Call	No Call

Table 2. NGS Calls for DNA and RNA Variants are Preserved in Multi-Panel Runs. Call results from RNA and DNA panels in combined panel NGS runs were 100% consistent with single panel pools and expected results.

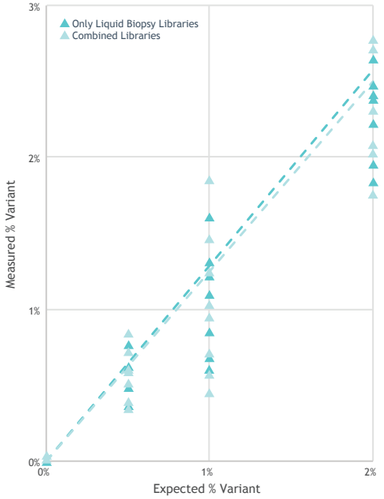

Figure 4. Variant Quantification in Liquid Biopsy Samples. Mutant allele frequencies were consistent between multiple and single panel pools, as well as for variants in control samples.

Conclusions

- We find the the number of reads/library is proportional to the normalized loading capacity, thereby allowing different library types with different depth requirements to be sequenced simultaneously.
- Relative intra-library coverage and absolute variant calling were preserved between single-type and multi-type library sequencing runs.
- Post-sequence analysis using QuantideX® Reporter was unaffected by mixed library NGS runs and generated the expected results.
- The data show that QuantideX® NGS libraries can by combined for the reliable, versatile, and concurrent detection of genetic variants in solid tumor DNA, solid tumor RNA, and liquid biopsy samples.

Asuragen®
www.asuragen.com