

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Practical Laboratory Medicine

journal homepage: [www.elsevier.com/locate/plabm](http://www.elsevier.com/locate/plabm)

## Key considerations for comprehensive validation of an RNA fusion NGS panel

Subit Barua<sup>1</sup>, Gary Wang<sup>1</sup>, Mahesh Mansukhani, Susan Hsiao, Helen Fernandes<sup>\*</sup>

Department of Pathology and Cell Biology, Columbia University Medical Center, New York, USA

### ARTICLE INFO

#### Keywords:

Next generation sequencing  
Precision medicine  
Molecular oncology  
Assay  
Validation  
RNA Fusions

### ABSTRACT

**Objectives:** Validation of RNA-based NGS assays for the detection of therapeutically targetable gene fusions is challenging. Here, we report systematic validation and quality control monitoring of our targeted fusion panel for the detection of 17 clinically relevant fusion transcripts across several tumor types. We implemented this RNA Fusion Panel as a reflex test for tumors lacking DNA driver mutations.

**Design:** Forty-four formalin-fixed, paraffin-embedded (FFPE) or fresh-frozen lung, brain, soft tissue and skin tumors were used to determine the accuracy of the assay. Additional fusion-positive specimens and a calibrated reference standard were used to establish the precision, reproducibility and sensitivity of the assay. All aspects of the validation, including quality control metrics, were performed according to New York State guidelines.

**Results:** For the RNA fusion panel, accuracy, reproducibility and precision studies were above 99%. Reproducibility and sensitivity studies with the reference standard were helpful in identifying inconsistencies. The limit of detection for most RNA fusion transcripts was 50 copies. Application of the RNA fusion assay as a reflex test to 450 tumor samples lacking DNA driver mutations resulted in a 10% increase in diagnostic yield with minimal additional processing time.

**Conclusions:** The validated RNA fusion panel provides clinical utility in therapy selection for patients with solid tumors. By using a sequential testing approach, the RNA fusion assay complements the DNA hotspot assay in identifying clinically relevant variants across many tumor types with minimal additional increase in processing time.

## 1. Introduction

Gene fusions are hybrid genes formed by the breakage and re-joining of two different chromosomes, or from intra chromosomal rearrangements due to deletions, insertions or inversions. Gene fusions are often oncogenic, and act as drivers for tumorigenesis through the constitutive activation of protein kinases, deregulation of proto-oncogenes, or disruption of tumor suppressor genes. Detection of defined gene fusions in specific tumor types can guide targeted therapy and provide risk stratification [1,2]. In addition, identification of novel gene fusions can provide insights into the mechanism of tumorigenesis, allow for sub-classification of histologically similar tumors and may serve as useful biomarkers for disease progression.

The technologies for the detection of gene fusions have evolved from contemporary sequencing techniques, fluorescence *in situ*

\* Corresponding author.

E-mail address: [hf2340@cumc.columbia.edu](mailto:hf2340@cumc.columbia.edu) (H. Fernandes).

<sup>1</sup> Authors contributed equally.

<https://doi.org/10.1016/j.plabm.2020.e00173>

Received 14 October 2019; Received in revised form 19 April 2020; Accepted 28 May 2020

2352-5517/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

hybridization (FISH), immunohistochemistry (IHC), and reverse-transcriptase PCR (RT-PCR) analysis to multiplexed Next-Generation Sequencing (NGS) [3,4]. There are inherent drawbacks associated with each method of detection. For example, paucity of tumor cells can increase the likelihood of false-negatives by FISH. RT-PCR is limited to the detection of previously characterized fusion partners. IHC interpretation of different gene fusion partners can be a subject of inter-observer variability. NGS testing is prohibitively expensive for routine clinical use, and its interpretation often requires a sophisticated bioinformatics pipeline [5–7]. Irrespective of the technology, identification of gene fusions together with DNA variants in solid tumors is extremely informative for the treating physician.

At our institution, we routinely utilize a targeted, cancer specific DNA hotspot panel for the detection of single nucleotide variant (SNV), insertions and deletions (indels) from Formalin-fixed paraffin embedded (FFPE) or frozen tissue. Assays with larger panels, such as cancer Whole Exome Sequencing (cWES) or transcriptome analysis, are reserved for pediatric patients or patients with advanced disease. Recently, a custom RNA Fusion Panel using Anchored multiplex PCR (AMP) technology (ArcherDX, Boulder, CO, USA) was introduced into our testing menu. This technology targets chromosomal breakpoints, which can fuse with known or novel partners, thus facilitating the identification of new cancer driver fusions [8]. Several studies have documented the application of this technology for detection of fusions in solid tumors and hematological malignancies [9–11].

In this report, we successfully validated an RNA fusion panel based on AMP technology targeting 17 cancer driver genes with clinically relevant fusions in lung, thyroid, brain, breast and gastrointestinal tumors (Table 1). As part of the validation, we undertook an in-depth examination of the quality indicators of the fusion panel as per the guidelines from the New York State Department of Health (NYSDOH) [12]. Our validation also incorporated the prerequisites detailed in the consensus recommendation of the Association for Molecular Pathology and College of American Pathologists [13]. This report provides details for addressing the four main components of validation, namely accuracy, precision, reproducibility and limit of detection or sensitivity. In addition, we provide our experience with a sequential testing algorithm, whereby samples negative for DNA variants (based on a hotspot panel) were then reflexed to RNA fusion testing. By adopting this process, we were able to increase the overall diagnostic yield with minimal increase in turn-around time and reduction in overall cost. Finally, we discuss the clinical utility of RNA-based fusion detection in an academic oncology practice.

## 2. Materials and methods

### 2.1. Patient samples and nucleic acid extraction

A total of forty-four formalin-fixed, paraffin-embedded (FFPE) or fresh-frozen tissue, 34 of which had confirmed fusions and 10 fusion-negative samples were used to establish the accuracy and the specificity of the assay. Of the 44 patient specimens, 37 were slides with FFPE tumor tissue from surgical and/or cytology specimens, and the remaining 7 patient samples were from fresh and/or frozen tissue. The samples represented non-small cell lung cancer (NSCLC), thyroid cancer, skin, glioblastomas and gastrointestinal tumors containing at least 20% tumor. The validation samples had fusions that were previously confirmed by orthogonal assays performed in accredited clinical laboratories including FISH (N = 5), WES (N = 6), Real-time PCR (N = 9), NGS based targeted assays (N = 5), RNA-fusion assays (N = 7), and transcriptome analysis (N = 12).

5-10 unstained slides with 5- $\mu$ m sections were obtained from FFPE tumor tissue. The tumor-enriched area was encircled on the hematoxylin and eosin-stained slides and the neoplastic cellularity of the selected area was estimated by a molecular genetic pathologist. The scored slide was used to guide the macrodissection of tumor cells from the unstained slides. Microdissection was performed on cases where the tissue was fragmented and/or pauci-cellular. The tumor percentage in the samples selected ranged from 20 to 80%. Tumor DNA and RNA was extracted from FFPE tissue using the Qiagen AllPrep DNA/RNA FFPE kit and Qiacube protocol (QIAGEN, Hilden, Germany). RNA from frozen tissue was extracted using Qiagen RNeasy RNA extraction kit. After extraction, RNA was quantified using the Qubit fluorometric quantification system (Thermo Fisher Scientific, Waltham, MA).

**Table 1**  
Gene details for fusion drivers in the panel.

Target Gene	NCBI_reference_sequence	Target_exons
<i>ALK</i>	NM_004304	2,4,6,10,16,17,18,19,20,21,22,23,26
<i>AXL</i>	NM_021913	18,19,20
<i>BRAF</i>	NM_004333	1,2,3,7,8,9,10,11,12,13,15,16
<i>EGFR</i>	NM_005228	1,7,8,9,16,19,20,24,25
<i>FGFR1</i>	NM_015850	2,3,4,5,6,7,8,9,10,11,12,17
<i>FGFR2</i>	NM_000141	2,5,7,8,9,10,16,17
<i>FGFR3</i>	NM_000142	3,5,8,9,10,16,17,18
<i>MET</i>	NM_000245	2,4,5,6,13,14,15,16,17,21
<i>MYB</i>	NM_001130173	7,8,9,11,12,13,14,15,16
<i>NRG1</i>	NM_004495	1,2,3,4,6,8
<i>NTRK1</i>	NM_002529	2,4,6,8,10,11,12,13
<i>NTRK2</i>	NM_006180	5,7,9,11,12,13,14,15,16,17
<i>NTRK3</i>	NM_002530	4,7,10,12,13,14,15,16
<i>PDGFRA</i>	NM_006206	7,10,11,12,13,14,15
<i>PPARG</i>	NM_015869	1,2,3
<i>RET</i>	NM_020975	2,4,6,8,9,10,11,12,13,14
<i>ROS1</i>	NM_002944	2,4,7,31,32,33,34,35,36,37
<i>THADA</i>	NM_022065	24,25,26,27,28,29,30,36,37

## 2.2. Analysis of gene fusions

Total input RNA (range 20–50 ng) from each sample was used for cDNA library preparation using custom primers with Archer FusionPlex reagents (ArcherDX, Boulder, CO, USA) per manufacturer's recommendation. Briefly, after preparation of cDNA with random priming, the library was subjected to a cDNA quality check (PreSeq quality check) using a quantitative PCR (qPCR) assay that uses KAPA SYBR® FAST qPCR Master Mix (KAPA Biosystems, Wilmington, MA, USA). AMP chemistry then uses molecular barcodes along with universal primer binding sites to ligate with the double stranded cDNA library for selection of fragments with known and novel fusion partners. Subsequent annealing with gene-specific primers that target the 17 genes in the fusion panel results in PCR amplification of the fusion product. The resulting NGS library is quantified using KAPA SYBR® FAST qPCR Master Mix, normalized and then sequenced on the MiSeq (Illumina) using the MiSeq v2 or MiSeq v3 Reagent Kit according to manufacturer's instructions (Illumina, San Diego, CA). Briefly, pooled libraries were denatured, diluted to 10–15 pmol/L and loaded on a flow cell. Data Analysis for the identification of relevant fusions was performed using Archer Analysis software version 5.2 (ArcherDX, Boulder, CO, USA). Fastq files were also parsed through an updated version of the software (v 6.0.2).

## 2.3. Controls for monitoring quality and performance

**Internal controls** for the purpose of monitoring the overall quality of the sample are built into the assay. These include house-keeping genes *CHMP2A*, *RABA7A*, *GPI* and *VCP*. Amplification of these genes is used by the Archer software to assess RNA quality. The inability of the software to detect these genes is considered a run failure.

**Reference standards** for monitoring the analytical performance of the assay were incorporated into every run. The Sereaseq Fusion RNA Mix v3 (SeraCare Life Sciences, Milford, MA, USA) is a reference standard containing a total of 16 clinically relevant fusion transcripts (14 gene fusions and 2 oncogenic isoforms) that are engineered into a well-characterized cell line backbone and quantified with digital PCR (Supplementary Table\_ S1). The percentage of reads that spanned the fusion junction served as a quality control parameter for assay performance. This reference standard was also used for assessing the analytical sensitivity of the assay. Please note that our RNA fusion panel did not interrogate fusions in the *ERG* gene and was unable to detect the Tmprss2-ERG fusion present in the Sereaseq reference control. We therefore used the control to monitor 15 transcripts, including 13 gene fusions and 2 oncogenic isoforms.

## 2.4. Quality metrics for RNA fusions

To ensure optimal sequencing quality for variant analysis, all samples had to meet the following criteria (Fig. 1A): PreSeq quality check (PreSeq score): Ct (crossing threshold) value by qPCR  $\leq 31$  for cDNA; At least 500,000 total reads; and  $\geq 10$  unique RNA Start Sites (SS) for per Gene Specific Primer 2 (GSP 2).

Samples with detectable fusions had to meet the following additional criteria (Fig. 1B): Number of supporting reads spanning the fusion junction  $\geq 5$

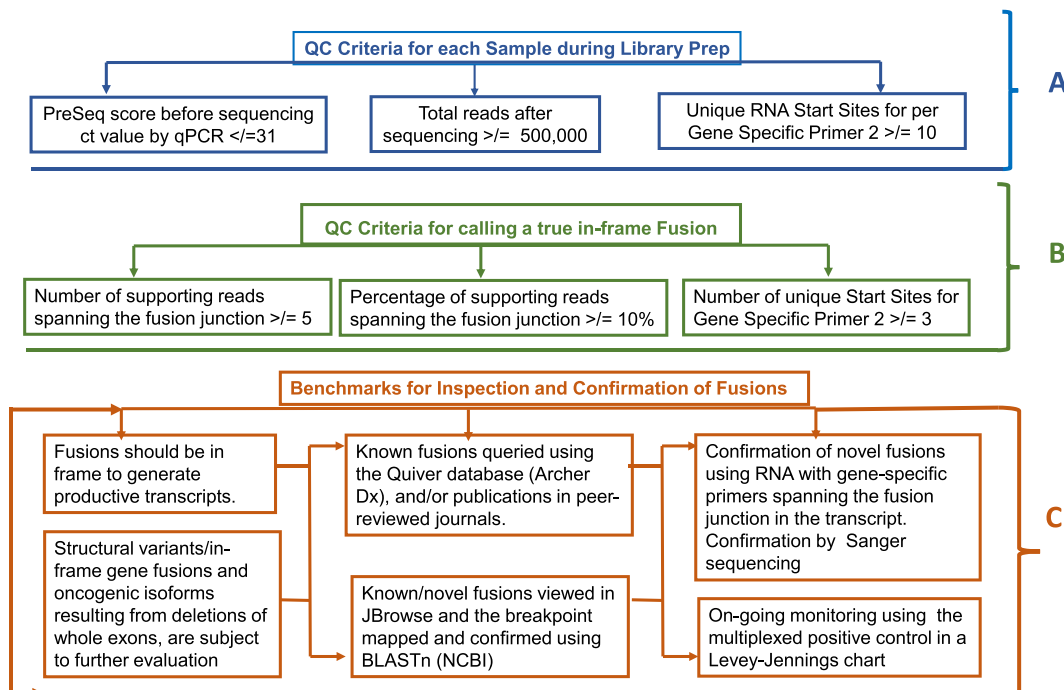


Fig. 1. QC algorithm applied to prioritize and assess the fusions from the Targeted RNA NGS Fusion Assay.

fusion junction  $\geq 5$ ; Percentage of supporting reads spanning the fusion junction  $\geq 10\%$ ; Number of unique start sites (SS) for the fusion sequenced specific primer  $\geq 3$ .

### 2.5. Inspection of RNA fusions identified: Fusions; gene translocations; rearrangements; oncogenic isoforms

RNA fusions need to be in frame in order to generate productive transcripts. All structural variants, including in-frame gene fusions, rearrangements and translocations as well as oncogenic isoforms resulting from deletions of whole exons, were subject to further evaluation (Fig. 1C). Databases with details for documented fusions were used to authenticate the fusion sequence identified. Known fusions were queried using Quiver (a database owned and maintained by ArcherDX), fusion database and/or publications from peer-reviewed journals. The fusion transcript sequenced from both known and novel fusions was viewed in JBrowse and the exact position of sequences were accurately mapped and confirmed using BLASTn using the human genomic plus transcript database [14,15]. Identification of both gene partners in the sequence were used as primary evidence for the identification of the fusion. For confirmation of novel fusions, gene-specific primers spanning the fusion junction in the transcript were synthesized and the sequences of amplified product were confirmed using Sanger sequencing. When needed, fusions identified by Quiver, especially for *MET* exon 14 skipping events that did not meet the set quality metrics were then reflexed to a targeted NGS hot spot panel to detect splice site variants resulting in *MET* exon14 skipping. Two other variants, one with a *NTRK3* fusion and the other with a *ROS1* fusion were confirmed by RT-PCR.

### 2.6. On-going monitoring for quality assurance

To ensure optimal ongoing performance of the assay, the percentage of reads spanning the fusion junction of the 13 fusions and 2 oncogenic isoforms in the reference standard were monitored in every run in a Levey-Jennings chart [16]. The initial parameters of mean  $\pm$  3SD (standard deviations) were set using data obtained from fifteen individual runs of the Seraseq Fusion RNA reference standard (Fig. 3). The percentage reads for all 15 variants in the reference standard are expected to fall within the set range. Outliers that are out of range for more than 2 consecutive runs were further evaluated.

### 2.7. Algorithm for identification of DNA and RNA variants in a single sample

In contrast to assays that interrogate DNA and RNA variants concurrently, we developed a sequential approach for identification of both types of variants. The algorithm used is depicted in Fig. 2. Briefly, after enrichment for tumor cells, both DNA and RNA were extracted simultaneously. This was followed by DNA variant analysis (Day 2–4) and reporting of results on Day 5. At this point, cases with no identifiable drivers were processed for the fusion assay (Day 6–8) and resulted on Day 8 or 9.

## 3. Results

### 3.1. Accuracy of the fusion assay

Ten of the seventeen fusion genes and the two oncogenic isoforms were fully validated as per the NYSDOH guidelines [12]. Fusions in *ALK*, *BRAF*, *EGFR*, *FGFR1*, *FGFR3*, *MET*, *NTRK1*, *NTRK3*, *RET*, *ROS1* genes, as well as oncogenic isoforms resulting from deletions in exons 2–7 in *EGFR* (EGFRvIII) and exon 14 for *MET* (exon 14 skipping) were tested on three patient samples with previously confirmed fusions by orthogonal assays in accredited clinical laboratories.

All forty-four sequenced samples passed pre-sequencing and sequencing quality metrics and were considered informative. The median Ct values for PreSeq for all cases was 27.9 (range 24.3–30.92). The average supporting reads across the 34 validation samples was 560 (range 8–3540) and the average % of reads was 52 (range 14–100). The average number of unique start sites across the samples was 91 (range 7–378). Additional quality metrics including total number of fragments sequenced, were also met for all specimens (see

### Algorithm for fusion testing

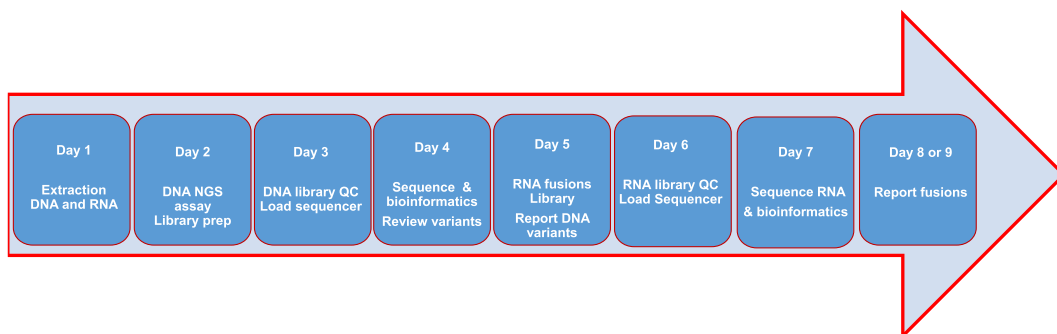


Fig. 2. Overview of RNA fusion transcript reflex testing scheme.

Monitoring of Reference Control using Percentage(%) Reads

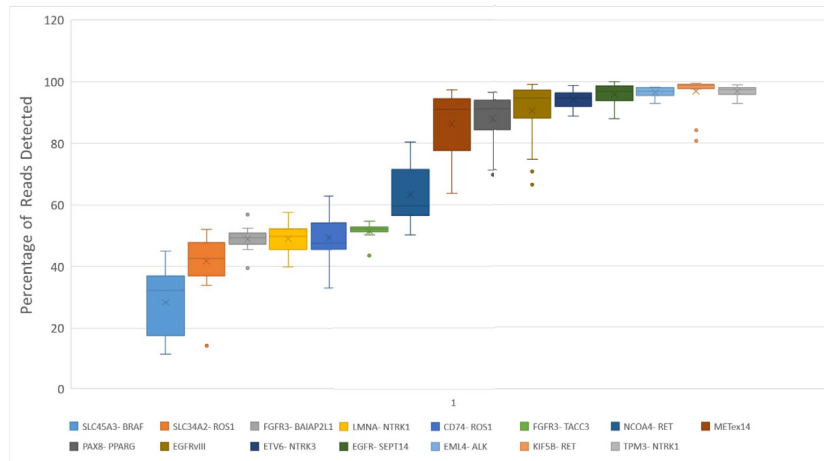


Fig. 3. Percentage of reads corresponding to individual fusions in the reference standard. Data shown are mean percentage of reads ± SDs from fifteen individual runs.

Supplementary Table S2 for details). The specific exons involved in the fusion transcript identified are provided for all samples. A sample with a EZR - ROS1 fusion showed a lower % reads than recommended. This sample also had a less than optimal Preseq value of 31.45. We included this sample because it was not possible to obtain additional samples with known ROS1 fusions. The specificity of the assay for the identification of variants is 100%. One specimen with a positive ALK fusion as determined by FISH was shown to have fused with a new gene partner, SLMAP. The SLMAP-ALK gene fusion was confirmed by Sanger sequencing, and has since been reported [17].

Table 2  
Precision and reproducibility of the RNA fusion Panel.

Replicate	Gene Fusion	Chromosomal Breakpoint	Read%	Unique SS	Preseq-CT	Average Read%	Std dev
1	FGFR3 (ex 15)-TACC3 (ex 10)	chr4:1808661,chr4:1739325	9.8	45	27.67	<b>8.73</b>	<b>2.67</b>
Mix 1	KIAA1549 (ex15)-BRAF (ex 9)	chr7:138552721,chr7:140487384	25.6	34		<b>24.60</b>	<b>1.65</b>
2	FGFR3 (ex 15) -TACC3 (ex 10)	chr4:1808661,chr4:1739325	10.7	37	27.31		
Mix 1	KIAA1549 (ex15) - BRAF (ex 9)	chr7:138552721,chr7:140487384	22.7	21			
3	FGFR3 (ex 15) - TACC3 (ex 10)	chr4:1808661,chr4:1739325	5.7	22	27.28		
Mix 1	KIAA1549 (ex15) - BRAF (ex 9)	chr7:138552721,chr7:140487384	25.5	27			
1	TPM4 (ex 7) - ALK (ex 20)	ch19:16204563,chr2:29446394	81.8	15	28.79	<b>86.70</b>	<b>4.67</b>
Mix 2	ETV6(ex 5) - NTRK3 (ex 14)	chr12:12022903,chr15:88483984	68.5	13		<b>65.73</b>	<b>6.27</b>
2	TPM4 (ex 7) - ALK (ex 20)	ch19:16204563,chr2:29446394	87.2	107	25.92		
Mix 2	ETV6(ex 5) - NTRK3 (ex 14)	chr12:12022903,chr15:88483984	69.2	66			
3	TPM4 (ex 7) - ALK (ex 20)	ch19:16204563,chr2:29446394	91.1	88	26.01		
Mix 2	ETV6(ex 5) - NTRK3 (ex 14)	chr12:12022903,chr15:88483984	69.5	58			
1	EGFRvIII (ex1-ex8)	chr7:138552721,chr7:140487384	29.2	22	29.1	<b>32.87</b>	<b>3.36</b>
2	EGFRvIII (ex1-ex8)	chr7:138552721,chr7:140487384	35.8	23	29		
3	EGFRvIII (ex1-ex8)	chr7:138552721,chr7:140487384	33.6	25	28.61		
1	MET ex14 (ex13 - ex15)	chr7:116411708,chr7:116414935	46.4	19	27.81	<b>39.50</b>	<b>7.89</b>
2	MET ex14 (ex13 - ex15)	chr7:116411708,chr7:116414935	30.9	14	27.97		
3	MET ex14 (ex13 - ex15)	chr7:116411708,chr7:116414935	41.2	22	27.31		
1	EZR(ex9) - ROS1(ex34)	chr6:159191796,chr6:117645578	43.5	19	28.69	<b>46.75</b>	<b>4.60</b>
2	EZR(ex9) - ROS1(ex34)	chr6:159191796,chr6:117645578	50	15	28.9		
3	EZR(ex9) - ROS1(ex34)	chr6:159191796,chr6:117645578	ND	ND	ND		
1	CCDC6(ex1) - RET(ex12)	chr10:61665880,chr10:43612032	91.3	133	25.81	<b>89.10</b>	<b>2.03</b>
2	CCDC6(ex1) - RET(ex12)	chr10:61665880,chr10:43612032	87.3	109	25.8		
3	CCDC6(ex1) - RET(ex12)	chr10:61665880,chr10:43612032	88.7	150	25.84		

ND: Not Done.

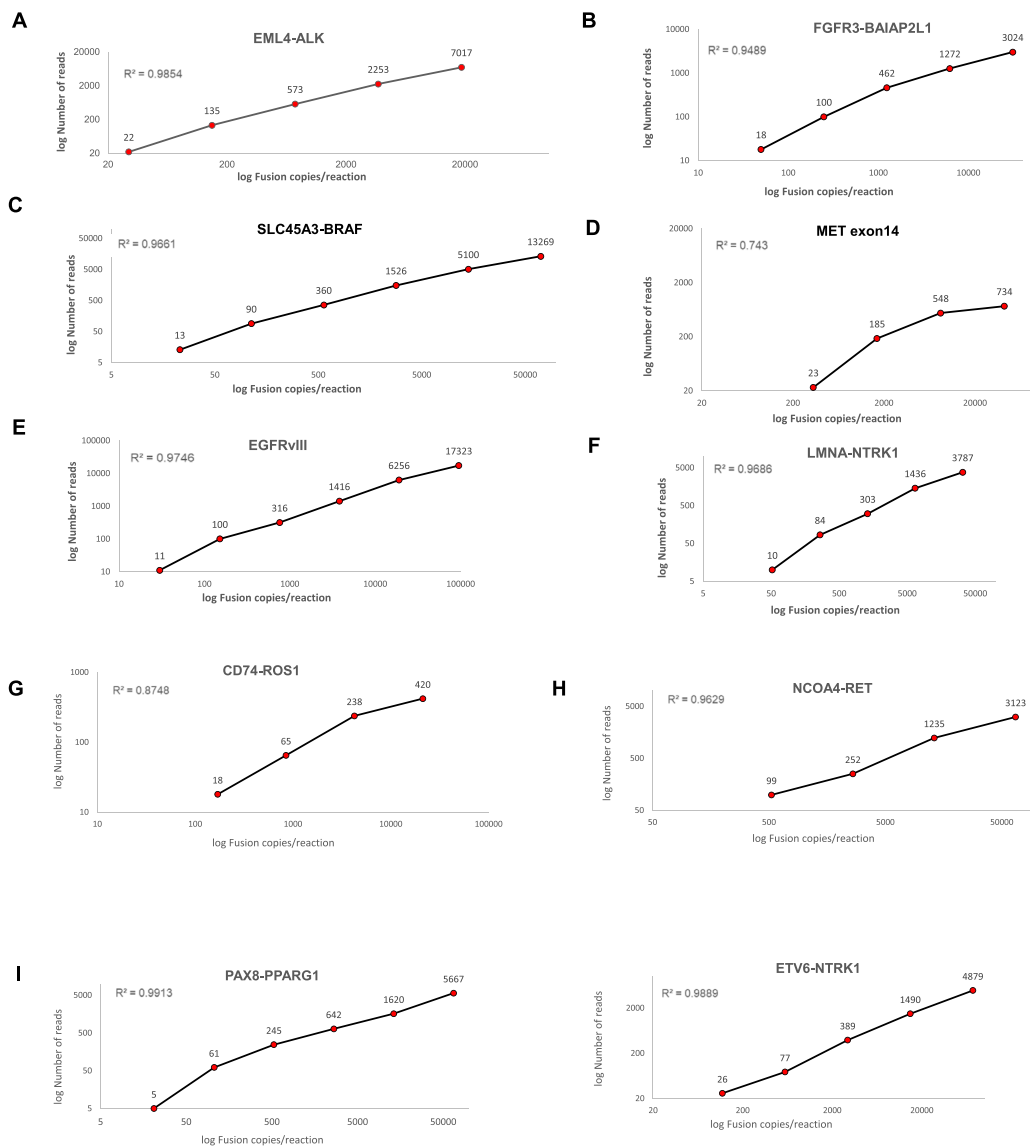
Mix 1: Multiplexed fusion transcripts.

Mix 2: Multiplexed fusion transcripts.

Average Read%: The average of the 3 replicates for each fusion transcript.

### 3.2. Reproducibility and precision

All 10 genes that were fully validated for accuracy were also validated for precision and reproducibility as per NYSDOH guidelines [12]. This included running three replicates for each fusion, on three different days/runs, using two different technologists (Table 2). This was accomplished by testing gene fusions individually or in combination as mixtures. Patient specimens with confirmed fusions in *ALK*, and *NTRK* genes were included in Fusion Mix 1. Fusion Mix 2 harbored *BRAF* and *FGFR3* gene fusions. These mixtures were created by mixing equal amounts of RNA from previously run specimens, and the selected cases showed comparable quality metrics with similar Preseq Ct, unique reads, unique start sites and percent (%) reads. Performing precision and reproducibility studies on mixtures rather than individual gene-fusions is practical from a cost and time perspective as it reduces the number of required runs. The precision and reproducibility of fusion identification in the reference standard was assessed with data obtained from fifteen different runs that had incorporated controls. The performance of the fusions in the reference standard was similar to that expected as per manufacturer data (Fig. 3). The standard deviation bars point towards the excellent reproducibility of the control for most fusions. The greatest standard deviation was observed for *MET* exon14 oncogenic isoform.



**Fig. 4.** Correlation between number of fusion copies obtained from reference standards using digital droplet PCR with the number of reads obtained in the RNA fusion NGS assay. Fusion copies/reaction, number or reads are represented for each standard. Values are represented in logarithmic scale.

### 3.3. Sensitivity of the fusion assay

Unlike the DNA hotspot NGS assays, quantification of the RNA fusions and assessment of assay sensitivity is challenging. To circumvent this, we utilized SeraSeq reference standards and quantified fusion copy numbers using digital droplet PCR. As shown in Fig. 4- number of copies detected ranged from 0 to 100,000 per reaction. The known number of copies were correlated with the number of reads obtained in the fusion NGS assay. The linearity depicts the overall robustness of the assay for identification of fusions in specimens that harbor a wide linear range of fusion transcripts. Notably, the total number of reads spanning the junction of the *ROS1* fusion and the *MET* oncogenic isoform were significantly lower, and these fusions showed lower degree of linear correlation with copy number. Interestingly, results obtained during validation and after implementation of the assay often showed a transcript in the *ROS1* gene that the software reads as “unaligned”. Though these samples meet the assay QC metrics, these results are likely false-positive as the transcript was out-of-frame. Similar spurious events were also seen in samples with lower levels of input RNA.

### 3.4. Performance of the NGS RNA fusion assay for patient management

The 17 gene fusion panel has been successfully incorporated into the routine testing menu of our NYS certified clinical laboratory. To make the testing informative for patient care and cost-effective for the laboratory, all lung and thyroid cancer samples lacking somatic driver mutations based on DNA hotspot panel were then routinely interrogated for RNA fusion drivers.

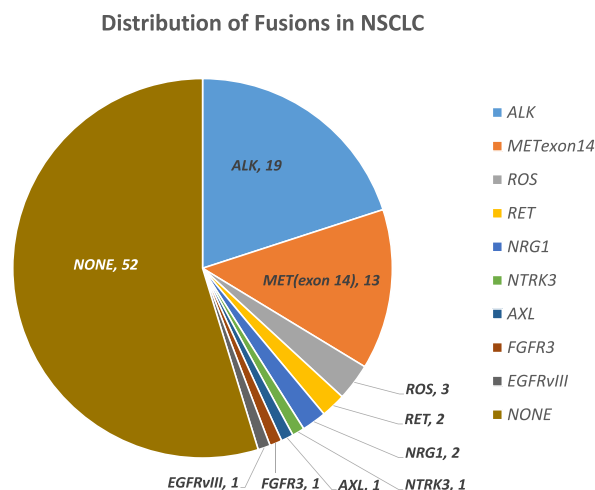
A total of 407 lung adenocarcinomas underwent DNA hotspot panel testing during a two-year period (September 1, 2017 to September 1, 2019). Of these, 95 (23%) were negative for somatic driver mutations and underwent successful sequencing for the detection of fusion transcripts. Fusions were seen in 43 cases. These included 19 *ALK* (all positive by IHC and FISH); 13 *MET* exon14 skipping; 3 *ROS1*; 2 *RET*, 2 *NRG1*, one of each *AXL*, *NTRK3*, *FGFR3* and *EGFRvIII* (Fig. 5). One novel *ALK* fusion transcript (SLMAP-*ALK*) was detected and subsequently confirmed with Sanger sequencing [17]. Interestingly, the fusion assay also detected 2 *MET* exon 14 skipping events that were not otherwise identified with the our targeted-NGS DNA assay. Further examination of primer details revealed a potential for the DNA assay to miss *MET*-exon14 mutants due to lack of coverage in intron 13 and the 3' splice site, compared to the 5' splice site of intron 14.

A total of 48 thyroid atypia of undetermined significance (AUS) underwent DNA hotspot panel testing during the two-year period. Of these, 38 (79%) were negative for somatic driver mutations, and 22 underwent successful sequencing for the detection of fusion transcripts. Three clinically relevant fusions, one *PAX8-PPARG*, one *ESYT2-BRAF* (novel to thyroid tumors) and one uncommon *STRN-ALK* fusion were identified.

## 4. Discussion

Targeted assays using AMP technology are increasingly used by clinical laboratories to identify fusions in FFPE tumor tissue for patients with various cancers [18,19]. Since AMP technology enables the detection of novel fusions at the RNA transcript level, positive findings help to expand the landscape of tumor drivers and identify therapeutic targets. Our panel covered clinically relevant RNA fusions and oncogenic isoforms recommended for testing according to the NSCLC professional guidelines [20]. The custom panel we utilized was aimed at incorporating the detection of gene fusions into our routine diagnostic algorithm in order maximize patient benefit in a cost-effective manner.

The accuracy of the assay for detection of fusions and rearrangements involving deletions and splice variants is noteworthy.



**Fig. 5.** Pie diagram depicting number of DNA driver-negative NSCLC cases with positive fusion transcripts detected. Data shown is obtained from a total of 95 cases.

Reproducibility, robustness and sensitivity studies indicate that this assay can reliably detect 50 digital copies of the fusion transcript.

The QC parameters that dictate the metrics required for qualifying a clinical assay are of utmost importance for measuring the analytical performance of the assay. DNA targeted assays appear to have risen to good standard of performance as measured by proficiency testing parameters instituted by CAP and other regulatory bodies [13,21]. In contrast, testing for RNA fusions utilizes different principle of detection and as such currently lacks standardization. For example, the FDA approved OncoPrint Focus Assay (ThermoFisher scientific) employs primers that span the chromosomal breakpoints of known fusion transcripts. This targeted NGS assay, by virtue of its chemistry generates significantly greater number of supporting reads that span the interrogated gene fusion. On the other hand, whole transcriptome analysis has the ability to detect all potential fusions in the genome, though the sensitivity of the assay can be limited by pre and post-analytical factors such as RNA input and sophisticated bioinformatics [22,23]. Therefore, unlike DNA based assays that detect single nucleotide variants (SNV) and insertions or deletions (Indels) in cancer, comparisons using common denominators such as mutant allelic fraction (MAF) cannot be used in the currently utilized RNA fusion assays.

Hence, choosing the right QC metrics and using reference standards for understanding and evaluating the performance of the assay is extremely valuable. Using the quantified reference standard as a determinant of sensitivity has helped us to identify potential weak spots in our fusion assay. While weak suboptimal performance in a DNA assay can be easily measured by uniformity and depth of coverage, the AMP assay technology does not lend itself to these options. Instead, mispriming, leading to transcripts that are unaligned or out-of-frame is sometimes evident. These artifacts are more apparent in samples with very low levels of fusion transcripts and fusions with multiple breakpoints within the same exon of the fusion partner. In our hands, the latter issue is evident for *ROS1* gene fusions. Monitoring the performance of the assay with a well characterized reference standard is central to troubleshooting and interpretation of results. As seen in this study, the fusions with the greatest deviation from the mean, likely depict transcripts that are more sensitive to assay variability, including kinetics of the assay. Our data suggests that using a single fusion gene target is not ideal to assess the performance of an NGS assay, hence we advocate the use of multiple gene targets.

Examination of the fusion transcript for alignments to the reference genome is an essential measure of quality control. This is particularly true for new and novel fusion partners that are detected by the AMP assay. Over the past year, we have detected two novel fusions, SLMAP-ALK in lung cancer [17] and an ESYT2-BRAF fusion in a thyroid tumor. Both fusions were valid and confirmed by Sanger sequencing. Future *in vitro* studies are needed to better understand the significance of these novel fusion genes.

Prior to implementation of the fusion assay, our targeted 47 gene NGS panel was used for the detection of DNA variants in patients with NSCLC. We assessed the value of adding the fusion panel to our battery of tests for patients with NSCLC. When analyzed independently, our targeted DNA NGS panel was able to detect clinically informative variants (*BRAF*, *EGFR*, *KRAS*, *ERBB4*, *MET*, *STK11*) in 77% of NSCLC patients (312 out of 407 cases). After implementation of the fusion panel, fusion transcripts were identified in 43 out of 95 cases, increasing the percentage of clinically informative variants to 87%. Likewise, in patients with thyroid cancer, implementation of the fusion assay increased the percentage of clinically relevant variants from 21% to 27%. On another note, when compared to platforms that interrogate DNA and RNA concurrently, reflex RNA testing on samples negative for driver mutations on the DNA panel reduces the number of tests performed by 30–35%. This indeed results in cost savings for the performing laboratory.

Excluded from this analysis is the fact that identification of *ALK* fusions are also performed using IHC and FISH. While these methodologies are rapid and robust, they may not be able to distinguish between classic and novel fusion partners. In our study, the patient with non-*EML4* partner to *ALK* showed diffuse granular cytoplasmic staining by IHC and was positive for rearrangement by FISH with *ALK* break-apart probe. The main advantage of the specificity of gene fusion partners at the exon and chromosomal breakpoint level, is to monitor the patient for minimal residual disease and/or the evolution of new or drug resistant clones. In this regard, monitoring of fusions in circulating tumor DNA (ctDNA) in plasma is proving to be a useful tool.

## 5. Conclusion

In summary, our validation systematically delineates NYSDOH requirements and AMP/CAP guidelines for a sequential DNA/RNA approach to identify gene fusions in a cost-effective manner. Quality control criteria utilized at library prep and sequencing stages support the detection and identification of true fusions. Benchmarks for querying the fusion add to the confidence level of a positive result. Integration of a multiplexed reference standard in every run helps to monitor the quality of the run and the integrity of the assay reagents. Finally, using a sequential approach for detection of SNVs and indels in DNA, followed by RNA fusions appears to provide a comprehensive solution for the identification of known and novel targetable variants in FFPE samples from cancer patients.

## Funding

This research did not receive any specific grant support.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRedit authorship contribution statement

**Subit Barua:** Methodology, Validation, Data curation, Writing - original draft. **Gary Wang:** Writing - original draft, Investigation,



Data curation, Writing - review & editing. **Mahesh Mansukhani**: Supervision, Writing - original draft. **Susan Hsiao**: Software, Writing - review & editing. **Helen Fernandes**: Conceptualization, Investigation, Data curation, Writing - review & editing.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.plabm.2020.e00173>.

## References

- [1] Q. Gao, W.W. Liang, S.M. Foltz, et al., Driver fusions and their implications in the development and treatment of human cancers, *Cell Rep.* 23 (2018) 227–238, <https://doi.org/10.1016/j.celrep.2018.03.050>, e223.
- [2] X. Dai, R. Theobald, H. Cheng, et al., Fusion genes: a promising tool combating against cancer, *Biochim. Biophys. Acta Rev. Canc* 1869 (2018) 149–160, <https://doi.org/10.1016/j.bbcan.2017.12.003>.
- [3] H. Ni, X.T. Wang, R. Fang, et al., Gene fusions in human cancers: a review focused on diagnostic biomarkers, method selections, and treatments, *Crit. Rev. Oncog.* 22 (2017) 403–410, <https://doi.org/10.1615/CritRevOncog.2017020550>.
- [4] F. Mertens, B. Johansson, T. Fioretos, et al., The emerging complexity of gene fusions in cancer, *Nat. Rev. Canc.* 15 (2015) 371–381, <https://doi.org/10.1038/nrc3947>.
- [5] A. Scattoni, A. Catino, L. Schirosi, et al., Discordance between FISH, IHC, and NGS analysis of ALK status in advanced non-small cell lung cancer (NSCLC): a brief report of 7 cases, *Transl Oncol* 12 (2019) 389–395, <https://doi.org/10.1016/j.tranon.2018.11.006>.
- [6] C. Lin, X. Shi, S. Yang, et al., Comparison of ALK detection by FISH, IHC and NGS to predict benefit from crizotinib in advanced non-small-cell lung cancer, *Lung Canc.* 131 (2019) 62–68, <https://doi.org/10.1016/j.lungcan.2019.03.018>.
- [7] F. Cabillic, P. Hofman, M. Ilie, et al., ALK IHC and FISH discordant results in patients with NSCLC and treatment response: for discussion of the question-to treat or not to treat? *ESMO Open* 3 (2018), e000419 <https://doi.org/10.1136/esmoopen-2018-000419>.
- [8] Z. Zheng, M. Liebers, B. Zhelyazkova, et al., Anchored multiplex PCR for targeted next-generation sequencing, *Nat. Med.* 20 (2014) 1479–1484, <https://doi.org/10.1038/nm.3729>.
- [9] S.W. Lam, A.M. Cleton-Jansen, A.H.G. Cleven, et al., Molecular analysis of gene fusions in bone and soft tissue tumors by anchored multiplex PCR-based targeted next-generation sequencing, *J. Mol. Diagn.* 20 (2018) 653–663, <https://doi.org/10.1016/j.jmoldx.2018.05.007>.
- [10] S. Afrin, C.R.C. Zhang, C. Meyer, et al., Targeted next-generation sequencing for detecting MLL gene fusions in leukemia, *Mol. Canc. Res.* 16 (2018) 279–285, <https://doi.org/10.1158/1541-7786.MCR-17-0569>.
- [11] X. Li, M. Anand, J.D. Haimes, et al., The application of next-generation sequencing-based molecular diagnostics in endometrial stromal sarcoma, *Histopathology* 69 (2016) 551–559, <https://doi.org/10.1111/his.12966>.
- [12] New York State Department of Health, NYSDOH. <https://www.wadsworth.org/sites/default/files/WebDoc/3NextGenSeqONCOGuidelines%2012318.pdf>. Accessed September 2019.
- [13] L.J. Jennings, M.E. Arcila, C. Corless, et al., Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the association for molecular pathology and College of American Pathologists, *J. Mol. Diagn.* 19 (2017) 341–365, <https://doi.org/10.1016/j.jmoldx.2017.01.011>.
- [14] S.F. Altschul, W. Gish, W. Miller, et al., Basic local alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410, [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- [15] M.E. Skinner, A.V. Uzilov, L.D. Stein, et al., JBrowse: a next-generation genome browser, *Genome Res.* 19 (2009) 1630–1638, <https://doi.org/10.1101/gr.094607.109>.
- [16] S. Levey, E.R. Jennings, The use of control charts in the clinical laboratory, *Am. J. Clin. Pathol.* 20 (1950) 1059–1066, [https://doi.org/10.1093/ajcp/20.11\\_ts.1059](https://doi.org/10.1093/ajcp/20.11_ts.1059).
- [17] C. Pagan, S. Barua, S.J. Hsiao, et al., Targeting SLMAP-ALK-a novel gene fusion in lung adenocarcinoma, *Cold Spring Harb Mol Case Stud* 5 (2019), <https://doi.org/10.1101/mcs.a003939>.
- [18] J.G. Mantilla, R.W. Ricciotti, E. Chen, et al., Detecting disease-defining gene fusions in unclassified round cell sarcomas using anchored multiplex PCR/targeted RNA next-generation sequencing-Molecular and clinicopathological characterization of 16 cases, *Genes Chromosomes Cancer* 58 (2019) 713–722, <https://doi.org/10.1002/gcc.22763>.
- [19] F. Chang, F. Lin, K. Cao, et al., Development and clinical validation of a large fusion gene panel for pediatric cancers, *J. Mol. Diagn.* 21 (2019) 873–883, <https://doi.org/10.1016/j.jmoldx.2019.05.006>.
- [20] K.M. Kerr, L. Bubendorf, M.J. Edelman, et al., Second ESMO consensus conference on lung cancer: pathology and molecular biomarkers for non-small-cell lung cancer, *Ann. Oncol.* 25 (2014) 1681–1690, <https://doi.org/10.1093/annonc/mdu145>.
- [21] N. Aziz, Q. Zhao, L. Bry, et al., College of American Pathologists' laboratory standards for next-generation sequencing clinical tests, *Arch. Pathol. Lab Med.* 139 (2015) 481–493, <https://doi.org/10.5858/arpa.2014-0250-CP>.
- [22] S. Liu, W.H. Tsai, Y. Ding, et al., Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data, *Nucleic Acids Res.* 44 (2016) e47, <https://doi.org/10.1093/nar/gkv1234>.
- [23] E.E. Heyer, I.W. Deveson, D. Wooi, et al., Diagnosis of fusion genes using targeted RNA sequencing, *Nat. Commun.* 10 (2019) 1388, <https://doi.org/10.1038/s41467-019-09374-9>.