Comprehensive validation of RNA sequencing for clinical NGS fusion genes and RNA expression reporting

Jun Hu[1], Jerod Parsons[1], Brittany Mineo[1], Joshua SK Bell[1], Jenna Malinauskas[1], Joshua Drews[1], Jackson Michuda[1], Calvin McCarter[1], Rasika Dhond[1], Jack C Tyndall[1], Nithya Sreenivasan[1], Sheeri Hanjra[1], Shannon Gallagher[1], Ankit Jambusaria[1], Naihui Zhou[1], Catherine Igartua[1], Nike Beaubier[1], Richard A Blidner[1], Robert Tell[1]

[1]Tempus Labs, Chicago, IL

INTRODUCTION

Tumor genome sequencing has emerged as a powerful tool for identifying biomarkers for targeted cancer therapies. While DNA sequencing is a well-established method and considered a gold standard for mutation events and genomic biomarkers, RNA sequencing (RNA-seq) can identify anomalies in gene transcription, regulation of gene expression, and gene fusions and have critical diagnostic and therapeutic impacts.

PANEL DESIGN: The Tempus 2nd generation RNA whole-transcriptome panel uses IDT xGen Exome Research Panel v2 backbone, which consists of >415K individually synthesized probes and spans a 34 Mb target region (19,433 genes) of the human genome. Additional Tempus-specific custom spike-ins probes are included to enhance target region detection (e.g., fusion and viral probes).

VALIDATION METHODS

RNA Fusion:

- **Clinical Accuracy:** 82 formalin-fixed paraffin embedded (FFPE) tumor samples were sequenced. Variants were confirmed with an orthogonal reference method, including 62 samples with targeted fusions and 20 with un-targeted fusions. Concordance acceptability was measured by the presence or absence of a fusion call. Sensitivity is computed as TP/(TP+FN).
- **Clinical Intra-assay Precision:** 7 fusion positive samples (2 un-targeted fusions, and 5 with targeted fusions) were analyzed in triplicate within the same run using the same reagent lot with different barcodes.
- **Clinical Inter-assay Precision:** 12 fusion positive samples (2 un-targeted and 10 with targeted fusions) were analyzed in 3 separate runs using different barcodes on different days by at least 2 different technologists. Samples were sequenced on different sequencers, and two different reagent lot numbers were used.
- **Limit of Detection (LOD) - Total Input:** 3 targeted fusion positive and 3 un-targeted fusion positive samples were tested in duplicate at 100 ng, 50 ng, and 25 ng starting input of total RNA.
- **LOD - Tumor purity:** 2 targeted fusion positive and 3 un-targeted fusion positive samples were titrated with fusion negative samples at 50%, 10%, 5%, 2.5%, and 1% tumor purity. Undiluted fusion positive and negative (diluent) samples were also run. All samples were tested at 100 ng starting input.
- **Analytical Specificity:** 12 replicates of GM12878, a sample with no expected fusions, were sequenced on 3 separate runs using different barcodes on separate days by at least 2 technologists. These samples were evaluated for reportability by clinical scientists.

RNA Expression:

- **Linearity Study:** 88 samples were utilized in this study. Samples represented multiple cancer types with expression values near the low, midpoint, and high values of the analytical measurement range. qPCR with TaqMan probes was run for 18 genes, 2 housekeeping genes and 2 negative control genes for all samples. Delta CT values were compared to RNA-Seq gene expression normalized values.
- **Concordance with DNA copy number variant (CNV):** Among the top 10 most frequently amplified and deleted genes, we compared expression among patients with deep deletions (0 copies) and patients with amplifications (≥8 copies).
- **UHR correlation to published qPCR data:** 21 replicates of the universal human reference (UHR) were run using different lot numbers, sequencers, and runs over different days. Replicates were compared to an external qPCR reference for 17,321 available genes on our panel (Wang et al. 2006).

RESULTS

Table 1: Fusion Validation Results

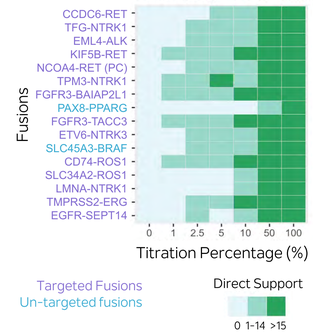
Metric	Target Fusions	Un-targeted Fusions	RNA Expression
Clinical Sensitivity	100%	97%	NA
Clinical Specificity	99.996%	N/A	NA
Intra-assay Precision	100%	100%	0.99 (0.9971 - 0.9979)
Inter-assay Precision	100%	100%	0.97 (0.9778 - 0.9975)
LOD Total Input	< 25ng	< 25ng	< 25ng 0.98 (0.986 - 0.992)
Analytical Specificity	100%	100%	NA

Table 2: Tumor Purity LOD - Clinical Samples

Fusion	Detection Limit with Evidence	Detection Limit with Strong Evidence
NCOA4-RET	2.5%	25%
NAB2-STAT6	0.90%	0.90%
VCL-CTNNA3	1.75%	35%
NUP214-ABL1	2.25%	90%
DOCK1-MLLT10	2.25%	45%

With Evidence: fusion has >=1 direct support reads
Strong Evidence: fusion has >=15 high quality spanning reads

Figure 1: Tumor Purity LOD - Positive Control



Read support for fusions in 0%, 1%, 2.5%, 5%, 10%, 50%, and 100% titration levels in SeraSeq Fusion RNA Mix v4 positive control sample. All targeted fusions were detected in as low as the 10% dilution, with all targeted and un-targeted fusions detected as the 50% dilution.

Figure 2: Expression Linearity Analysis

Correlation between mean ΔCT values and normalized gene expression in clinical samples for 18 oncogenes **(A)** or in UHR across 17,321 genes **(B)**. For **A**, mean ΔCT values for each sample/gene were normalized to the mean of the average CT values of AAMP and CANX for that sample.

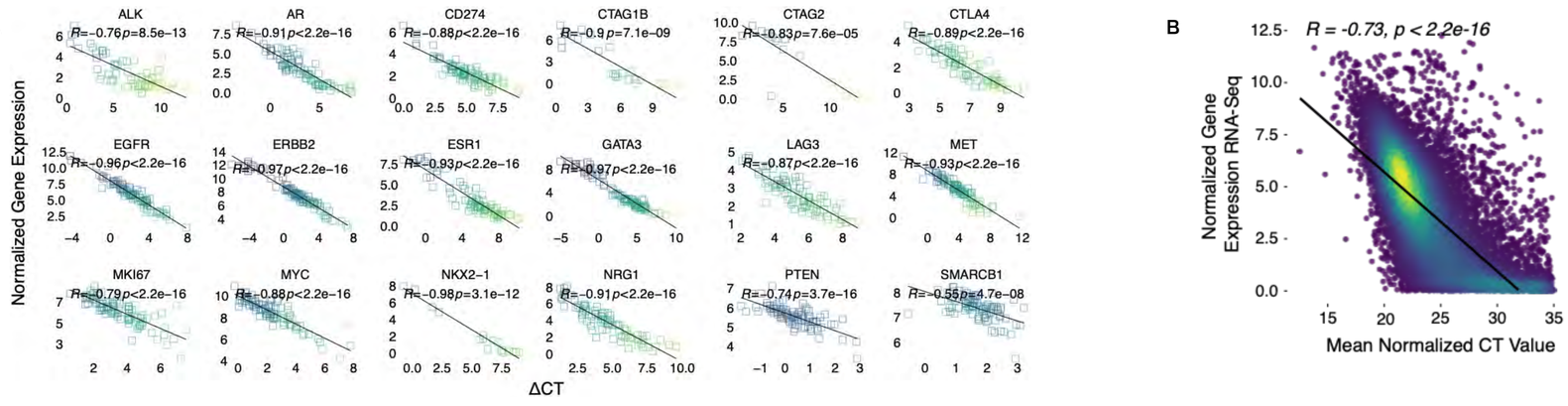
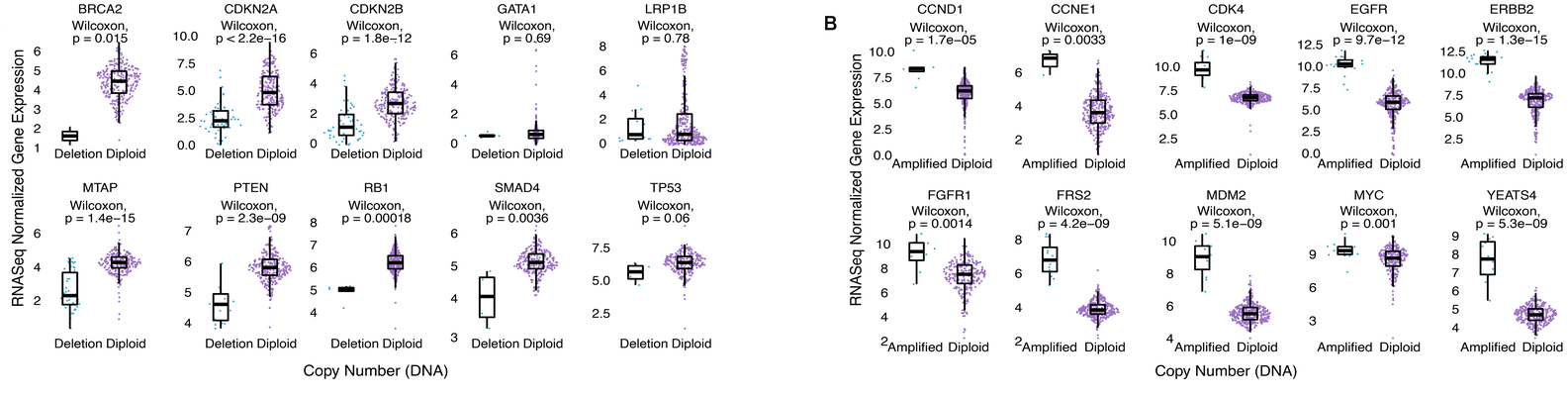


Figure 3: CNV Concordance Analysis

Top 10 genes with most reported **(A)** deletions (9/12 deleted genes resulted in Wilcoxon p < 0.05) or **(B)** amplifications (10/10 amplified genes results in Wilcoxon p < 0.05).



CONCLUSIONS

The Tempus whole-transcriptome hybrid-capture NGS assay has been enhanced to improve fusion detection and RNA expression calls. We clinically and analytically validated fusion calls for gene rearrangement detection. Our results show strong sensitivity and specificity for both targeted and untargeted fusions and significantly improved LOD for target fusions. We also analytically validated RNA expression counts for research purposes only. Our validation results show strong linear correlation with orthogonal qPCR results and high concordance with CNV results from our DNA panel. In conclusion, the whole-transcriptome assay offers clinically and analytically validated unbiased detection of common and novel gene rearrangements, as well as analytically validated gene expression data for comprehensive research analyses.

ACKNOWLEDGEMENTS: We thank Namratha Sastry, Ph.D. and the Tempus Scientific Communications team for data visualization guidelines and review of poster.

TEMPUS