

Exploratory Whole Transcriptome Sequencing Reveals Ancillary Fusion Transcripts in Driver-Positive Tumors

Nikkan Chakraborty^{1,2}, Raja Kumar^{1,2}, Rajan Kumar Singh³, Spandan Chaudhary², Naman Mangukia²

¹Faculty of Applied Science, Parul University, Waghodia, Vadodara, Gujrat, 391760, India. ²Unipath Specialty laboratory Ltd. Ahmedabad, Gujarat, 380015, India. ³Department of Biotechnology, Parul Institute of Technology, Faculty of Engineering and Technology, Parul University, Waghodia, Vadodara, Gujrat, 391760, India.

Abstract

Background: Gene fusions are important oncogenic drivers and therapeutic biomarkers across multiple malignancies. While targeted RNA sequencing panels reliably detect established rearrangements, the broader landscape of ancillary fusion transcripts remains incompletely characterized. This exploratory study aimed to characterize ancillary fusion transcripts identified by whole transcriptome sequencing (WTS) in tumors harbouring known driver fusions. **Materials and Methods:** WTS was performed on nine FFPE tumor samples with known *ALK*, *RET*, *ROS1*, or *FGFR* rearrangements. Fusion detection employed a consensus pipeline integrating STAR-Fusion, Arriba, and FusionCatcher, and high-confidence candidate fusions were prioritized using concordant junction and spanning read support. **Results:** Thirty-one candidate ancillary fusion transcripts were identified, demonstrating substantial inter-sample heterogeneity. Several involved genes associated with mTOR, *MAPK*, and calcium signalling pathways, although their functional significance remains uncertain. Within this limited cohort, FGFR-rearranged tumors exhibited a relatively higher ancillary fusion burden; however, no statistical comparisons were performed. **Conclusion:** WTS enabled exploratory characterization of transcriptomic fusion complexity beyond canonical driver events. Given the limited sample size, absence of orthogonal validation, and methodological differences between sequencing workflows, these findings should be considered hypothesis-generating and require confirmation in larger independent cohorts.

Keywords: Whole transcriptome sequencing- Gene fusions- RNA sequencing- Tumor heterogeneity

Asian Pac J Cancer Biol, **11** (4), 1101-1113

Submission Date: 07/11/2026

Acceptance Date: 09/10/2026

Introduction

Gene fusions are important oncogenic drivers and clinically relevant biomarkers in multiple solid tumors, with significant implications for diagnosis, prognosis, and targeted therapy selection. These chimeric transcripts arise from chromosomal rearrangements such as translocations, inversions, or deletions and may result in constitutive kinase activation or altered transcriptional regulation. Accurate detection of fusion transcripts is therefore an essential component of modern molecular oncology diagnostics procedures [1, 2]. Compared with DNA-based approaches, RNA sequencing directly evaluates expressed fusion transcripts, enabling identification of transcriptionally active rearrangements while reducing false-positive

detection of non-expressed structural variants [2, 3]. Advancements in next-generation sequencing (NGS) technologies have facilitated the use of multiplexed RNA-based fusion detection, replacing older single-target methods like fluorescence in situ hybridization and reverse transcription PCR. In the realm of clinical and translational oncology, RNA sequencing approaches are mainly divided into two types: targeted RNA panel sequencing and whole transcriptome sequencing (WTS). Targeted RNA panels focus on enriching specific genes or fusion hotspots to provide deep read coverage, which improves the analytical sensitivity for detecting low-abundance transcripts and degraded RNA, often found in formalin-fixed paraffin-

Corresponding Author:

Dr. Rajan Kumar Singh

Head and Assistant Professor, Department of Biotechnology, Parul Institute of Technology, Faculty of Engineering and Technology, Parul University, Waghodia, Vadodara, Gujrat, 391760, India.

Email: rajankumar.singh22658@paruluniversity.ac.in

embedded (FFPE) samples [2, 4, 5]. These assays typically employ amplicon or hybrid-capture techniques and are fine-tuned to identify clinically significant rearrangements, making them ideal for regular diagnostic use. Conversely, whole transcriptome sequencing offers an impartial, comprehensive examination of RNA expression and transcript structure across the genome. This method allows for the concurrent evaluation of gene expression patterns, alternative splicing, the discovery of novel fusions, and the diversity of transcript isoforms, thus providing an all-encompassing perspective of the transcriptomics framework of tumors [3, 6]. Despite these advantages, WTS distributes sequencing reads across thousands of transcripts, resulting in lower effective depth for individual genes and fusion junctions compared with targeted assays [7]. Consequently, detection sensitivity may be affected in samples with limited RNA quality, low tumor cellularity, or low transcript expression [8]. Different sequencing platforms and library preparation strategies possess distinct analytical characteristics that may influence read quality, fusion breakpoint detection, transcript coverage, and overall analytical performance [9, 10]. Fluorescence-based and semiconductor-based sequencing technologies each offer specific technical advantages and limitations related to sequencing accuracy, turnaround time, computational analysis, and workflow integration in clinical laboratories [11-13]. In addition, transcriptome-based RNA sequencing has become increasingly important in molecular oncology because it enables detection of transcriptionally active fusion transcripts and may identify clinically relevant rearrangements not captured by conventional genomic approaches [14]. Even with their distinct technical profiles, both sequencing methodologies have proven their worth in the clinical arena. In practice, the decision to use one over the other rarely hinges on a single feature; instead, it is a strategic balancing act involving sample volume (throughput), urgency (turnaround time), budgetary constraints, and the unique technical demands of the diagnostic test being performed [15, 16]. Emerging research highlights the increasing importance of transcriptome analysis in clinical oncology, particularly for the evaluation of expressed fusion transcripts and transcriptomic alterations with potential clinical relevance [17-19]. A variety of technical parameters including the intensity of sequencing (depth), the specific chemistry used to prepare libraries, the length of individual reads, and the choice of computational alignment tools can profoundly affect how accurately fusions are identified and how precisely gene expression is measured. Extensive benchmarking research consistently identifies sequencing depth as the most critical factor for sensitivity. This is especially true when attempting to catch fusion events in genes with low expression levels or in complex tumor samples where cancerous cells are diluted by healthy tissue [20]. In addition, technical variables such as sequencing depth, library preparation strategy, read length, and bioinformatics pipelines can substantially influence fusion detection performance and transcript quantification. The complexity of the transcriptome, including alternative splicing and transcriptional noise, further necessitates

the use of robust computational approaches for accurate fusion identification [21]. While RNA sequencing has substantially expanded fusion discovery in cancer, the broader landscape of ancillary fusion transcripts beyond canonical oncogenic drivers remains incompletely characterized. In this exploratory study, we employed whole transcriptome sequencing to investigate ancillary fusion events in nine FFPE tumor samples harboring known driver rearrangements involving *ALK*, *RET*, *ROS1*, and *FGFR* genes. Using a consensus-based fusion detection pipeline integrating STAR-Fusion, Arriba, and FusionCatcher, we aimed to characterize ancillary fusion transcripts predicted by whole transcriptome sequencing in tumors harboring known oncogenic driver fusions. In addition, we sought to examine the transcriptomic heterogeneity of fusion events detected beyond canonical rearrangements using a consensus-based fusion detection approach integrating multiple algorithms.

Materials and Methods

Study Design and Ethical Approval

This retrospective study involved nine archival formalin-fixed paraffin-embedded (FFPE) tumor specimens, all of which were previously identified as fusion-positive via standard clinical genomic screening. These samples were derived from patients with confirmed lung cancer who had undergone molecular profiling. This study received approval from the institutional ethics review board (Approval No. IECHR-SAINATH HOSPITAL/AHMC/178) and adhered to all established ethical guidelines. By selecting samples with known positives, we established a gold standard to measure the comparative performance of the two sequencing platforms.

Sample Selection and Histopathological Assessment

Archived FFPE tissue blocks were stored under routine pathology conditions at controlled room temperature (18–25 °C) prior to nucleic acid extraction. Tumor cellularity ranged from 25% to 100%. A minimum tumor cellularity threshold of 20% was established as the inclusion criterion [22]. To ensure accurate sample selection, a board-certified pathologist evaluated Hematoxylin and Eosin (H&E)-stained slides to determine the percentage of tumor cells. High-density tumor regions were manually outlined, and only samples meeting the >20% cellularity threshold within these annotated zones were moved forward. For the molecular workflow, three to five tissue curls each roughly 10 mm in length were sectioned from each paraffin block [23].

RNA Extraction and Quality Assessment

Total RNA was extracted from two FFPE tissue curls using the MagMAX™ FFPE DNA/RNA Ultra Kit (Thermo Fisher Scientific), which is optimized for recovery of fragmented nucleic acids from archival FFPE specimens [24]. RNA concentration was quantified using the Qubit RNA High Sensitivity (HS) Assay (Thermo Fisher Scientific), and RNA fragment size distribution was evaluated using the Agilent 4200 TapeStation System.

RNA quality assessment was performed as part of the routine laboratory workflow prior to library preparation [25]. RNA Integrity Number (RIN) values were recorded when technically feasible; however, standardized DV200 measurements were not available for all retrospectively archived FFPE samples included in this exploratory study because this study utilized retrospectively archived FFPE material processed according to routine institutional workflows. Variability in RNA integrity related to formalin fixation and long-term archival storage is a recognized limitation of FFPE-derived nucleic acid analysis. Previously characterized driver-positive clinical samples identified during routine molecular diagnostic testing served as reference-positive cases for fusion assessment. Dedicated spike-in controls or reference cell line standards were not included.

Targeted RNA Sequencing and Library Preparation

Previously generated targeted RNA sequencing data were obtained using the Oncomine Focus Assay, Oncomine Comprehensive Assay Plus, and Oncomine Precision Assay performed on the Ion Torrent S5 Plus platform according to manufacturer-recommended protocols with alignment to the human reference genome hg19 (GRCh37) followed by library preparation in which target regions are amplified using Ion AmpliSeq technology or fragmented cDNA is ligated with platform-specific adapters and barcodes to generate sequencing-ready libraries. These libraries are purified, quantified, and then subjected to template preparation using the Ion Chef System, where emulsion PCR clonally amplifies DNA fragments onto ion sphere particles (ISPs), enriches template-positive beads, and loads them onto semiconductor sequencing chips. Primary analysis software Torrent Suite Software include base calling, alignment using TMAP and Variant calling while annotation, reporting and fusion analysis were done in secondary analysis software namely Ion Reporter.

Whole Transcriptome Library Preparation and Sequencing

Whole transcriptome sequencing libraries were prepared using the KAPA RNA HyperPrep Kit with RiboErase (Roche) according to the manufacturer's protocol. RNA input ranged from 100–200 ng per sample. Ribosomal and globin RNA depletion was performed using RNase H-based depletion chemistry to improve transcriptome coverage in degraded FFPE-derived RNA samples [26]. Following depletion, RNA underwent fragmentation, first- and second-strand cDNA synthesis, end repair, adapter ligation, PCR amplification, and library quality assessment. Sequencing was performed on the Illumina NovaSeq X Plus platform using paired-end sequencing chemistry. Available sequencing quality metrics, including total read counts and mapping statistics, were evaluated as part of routine sequencing quality control analysis [27]. This platform was chosen for its ability to produce massive volumes of high-accuracy, short-read data with the benefit of consistent read lengths and paired-end support.

Bioinformatics Pipeline for Fusion Detection

Raw Data Processing

To refine the raw sequencing data, we subjected the initial FASTQ files to quality control and preprocessing using Fastp (v0.20.0). This step effectively stripped away adapter sequences and filtered out low-quality base calls to ensure data integrity. Following this cleanup, the processed reads were mapped to the human reference genome (GRCh38) using the STAR aligner (v2.7.10a). We specifically configured the alignment parameters to enhance the detection of chimeric reads, which is critical for identifying potential gene fusions [28].

Fusion Detection Algorithms

Fusion identification was performed using three complementary tools to enhance robustness and minimize platform-specific bias: STAR-Fusion (v1.13.1) detects fusion transcripts by leveraging chimeric junction reads generated during STAR alignment [29]. Arriba (v2.4.0) utilizes STAR output and applies advanced filtering strategies to balance sensitivity and specificity, reducing false-positive predictions in clinical datasets [30]. FusionCatcher (v1.33) employs a database-driven strategy integrating multiple reference fusion datasets and filtering algorithms to minimize artefactual fusion calls [31]. All tools were installed and executed on a local computational server using default parameters. Consensus fusion calls and cross-validation across algorithms were performed to strengthen confidence in reported rearrangements.

Exploratory Fusion Profiling Framework

Whole transcriptome sequencing findings were descriptively evaluated alongside previously available targeted RNA sequencing results generated using the Oncomine Focus, Comprehensive Plus, and Precision assays on the Ion Torrent S5 Plus platform. The analysis focused on characterization of canonical driver fusions, identification of ancillary fusion transcripts, and assessment of inter-sample fusion heterogeneity. Concordance of clinically established driver rearrangements and supporting read evidence across fusion detection algorithms were also examined. Because the targeted RNA sequencing and WTS workflows utilized different sequencing platforms, enrichment strategies, and analytical pipelines, the analysis was interpreted as exploratory rather than a formal comparative performance evaluation. The overall study design and analytical workflow are summarized in Figure 1.

Results

In the cohort of nine previously validated fusion-positive FFPE specimens, whole transcriptome sequencing (WTS) on the Illumina NovaSeq X Plus platform achieved a 100% detection rate for the known driver rearrangements. Each sample yielded an average of 50 million paired-end reads, providing sufficient depth for robust bioinformatics analysis across all cases. Fusion calling was performed using three established algorithms STAR-Fusion, Arriba, FusionCatcher. Across all nine

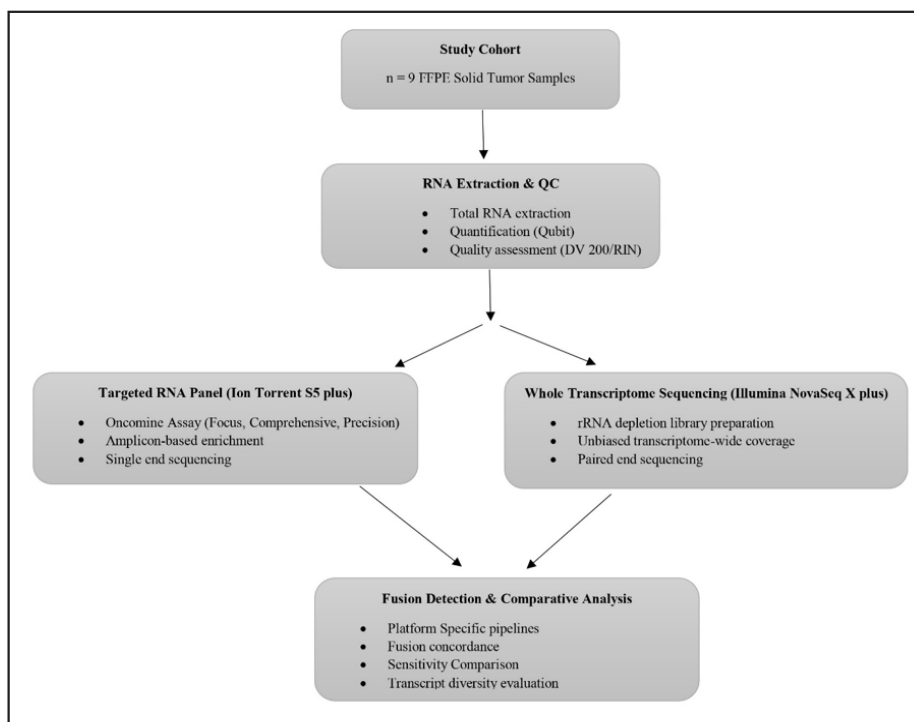


Figure 1. Study Design and Analytical Workflow. Nine FFPE solid tumor samples were subjected to RNA extraction and quality assessment prior to parallel sequencing using targeted RNA panels on the Ion Torrent S5 Plus platform and whole transcriptome sequencing on the Illumina NovaSeq X Plus platform. Fusion transcripts were identified using platform-specific bioinformatics pipelines, and results were compared for concordance, analytical sensitivity, and transcript diversity.

samples, WTS successfully identified the previously reported recurrent fusion events detected by targeted RNA panel sequencing. The use of three distinct callers (STAR-Fusion, Arriba, and FusionCatcher) provided high-confidence results, as each algorithm uses different logic to filter out transcriptional noise. WTS successfully resolved the specific exon-to-exon junctions, matching the sensitivity levels previously only attributed to focused targeted panels. The recurrent fusion distribution were as follows *RET* Fusions (n=3), all three cases were identified as the recurrent KIF5B (15)::*RET*(12) variant, a common driver in lung adenocarcinoma. *ALK* Fusions (n=3), this group consisted entirely of *EML4*::*ALK* rearrangements, including specific variants such as *EML4* (6)::*ALK* and *EML4* (13)::*ALK* (20). *FGFR* Fusions (n=2), two different *FGFR* family members were involved one *FGFR1* (17)::*TACC1* (7) and another *FGFR3* (17)::*TACC3* (6). *ROS1* Fusions (n=1), a single case of CD74 (6)::*ROS1* (34) were identified as mentioned in Figure 2.

The clinicopathological characteristics of the study cohort and the fusion events identified by targeted RNA panel sequencing are summarized in Table 1 which outlines the total number of fusion-positive cases along with the corresponding clinical information, demographic details (age and sex), and tumor cellularity for each sample. All nine formalin-fixed paraffin-embedded (FFPE) tumor samples harboured previously reported oncogenic fusion events. All previously identified clinically established driver fusions were detected by both workflows, demonstrating concordance for canonical driver rearrangements. However, ancillary fusion transcript profiles identified through WTS showed

substantial variability beyond the primary driver events.

Based on the data provided of the demographic summary, the patient cohort has the following characteristics: Gender Distribution- the study included 6 females (66.7%) and 3 males (33.3%). Age Profile- the mean age was approximately 58.2 years (range: 40–80 years), with a median age of 56 years as mentioned in (Supplementary Figure 1). Based on the above data we have also generated a boxplot (with individual data points overlaid) comparing the age ranges of male and female participants. It shows the median age and the spread of the data, which is useful for demonstrating that your cohort covers a wide age range as showed in (Supplementary Figure 2). All nine archival specimens passed initial pathological quality control, with H&E-stained section reviews confirming their suitability for downstream molecular analysis. The cohort demonstrated substantial tumor enrichment, yielding a median cellularity of 60% (IQR: 50%–70%). Crucially, every sample exceeded the study's pre-defined 20% inclusion threshold, providing the high neoplastic content necessary for sensitive fusion detection. To evaluate the relationship between sample composition and molecular findings, we analysed the distribution of tumor cellularity across the different fusion categories. The scatter plot reveals that diagnostic fusion events were successfully captured across the entire observed cellularity range (40%–70%). Notably, common fusions such as *EML4*::*ALK* were detected in samples with varying neoplastic content, confirming that our whole transcriptome sequencing approach provided sufficient sensitivity regardless of minor fluctuations in tumor purity as mentioned in Figure 3.

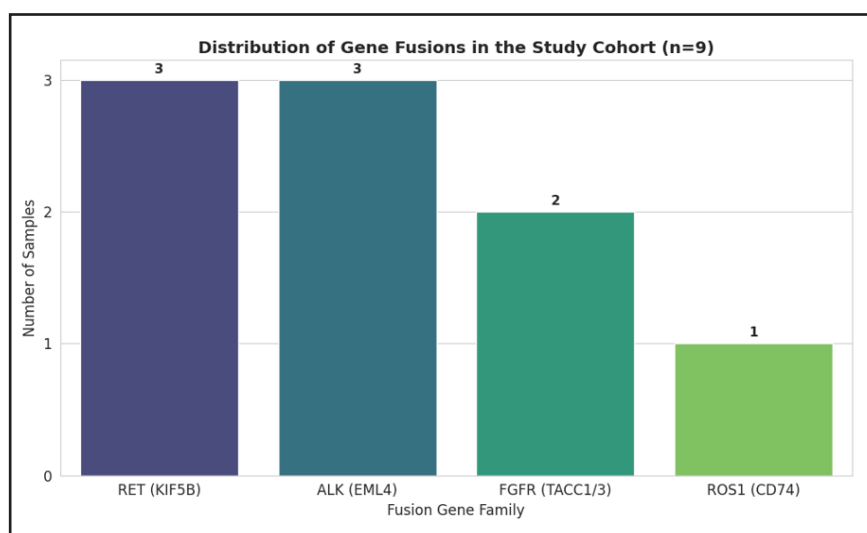


Figure 2. Distribution of Recurrent Fusion Events in the Study Cohort of Nine Fusion-Positive FFPE Solid Tumor Samples (n = 9).

Table 1. Clinicopathological Characteristics and Fusion Events Detected by Targeted RNA Sequencing in the Study Cohort

S.no	Sample ID	Age/Sex	Cellularity (%)	Clinical history	Fusion Detected	Primary Driver Fusion Concordance*
1	S1	71/F	50	Metastatic lung adenocarcinoma	KIF5B (15)::RET (12)	Yes
2	S2	80/F	60	Moderately differentiated lung adenocarcinoma	EML4::ALK	Yes
3	S3	63/M	50	Metastatic lung carcinoma	KIF5B (15)::RET (12)	Yes
4	S4	40/F	60	Metastatic high-grade spindle cell sarcoma (lung)	FGFR1 (17)::TACC1 (7)	Yes
5	S5	54/F	40	Oesophageal squamous cell carcinoma	EML4 (6)::ALK	Yes
6	S6	56/M	70	Papillary adenocarcinoma, gall bladder	FGFR3 (17)::TACC3 (6)	Yes
7	S7	52/F	65	Lung cancer	CD74 (6)::ROS1 (34)	Yes
8	S8	50/M	70	Primary lung cancer	KIF5B (15)::RET (12)	Yes
9	S9	58/F	70	Lung carcinoma with vertebral metastasis	EML4 (13)::ALK (20)	Yes

* Concordance refers specifically to detection of clinically established primary driver fusions and does not imply concordance across ancillary fusion transcript profiles identified by WTS.

A total of nine specimens were analyzed using both a targeted panel and whole transcriptome sequencing (WTS) to evaluate the detection of oncogenic fusion transcripts. Whole transcriptome sequencing analysis of nine tumor samples identified multiple gene fusion events across the cohort. Fusion detection was performed using a combination of algorithms, including STAR-Fusion, Arriba, and FusionCatcher, ensuring robust identification of candidate fusion transcripts. Beyond the primary oncogenic drivers, WTS identified a substantial number of additional fusion transcripts not captured by the targeted panel. Across the nine-sample cohort, a total of 31 unique ancillary fusions were detected using a multi-tool bioinformatics pipeline consisting of STAR-Fusion, Arriba, and FusionCatcher as mentioned in Table 2. The high-confidence detection of these fusion events were supported by the consistent results across all three

bioinformatics tools employed for every sample in the cohort. This robust multi-algorithm verification ensured the reliable identification of both established clinical drivers and other chimeric transcripts. Chromosomal breakpoint informations of all the samples along with fusion types were also provided with reference to fusion points based on (hg38 genome coordinates) in Table 3.

Whole transcriptome sequencing (WTS) revealed a diverse landscape of ancillary fusion events co-existing with primary oncogenic drivers across all nine specimens. While targeted panel sequencing reliably identified canonical driver fusions, including *KIF5B::RET*, *EML4::ALK*, *FGFR1::TACC1*, *FGFR3::TACC3*, and *CD74::ROS1*, the WTS pipeline identified a total of 31 additional chimeric transcripts that were not detected by targeted approaches. The distribution of these ancillary fusion events varied markedly across the cohort, indicating

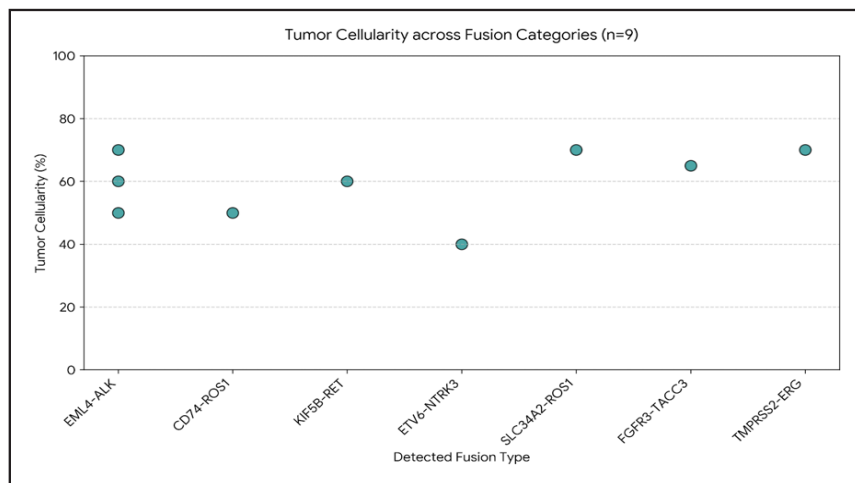


Figure 3. Distribution of Tumor Cellularity across Fusion Categories by Scatter Plot where x-axis Represents the Various Fusion Types Identified within the Cohort. Y-axis shows the tumor cellularity percentage for each specific case and each point represents a single specimen, allowing for an easy visual comparison of cellularity across different molecular subgroups.

a high degree of inter-patient molecular heterogeneity. Notably, genomic complexity was more pronounced in samples harboring fibroblast growth factor receptor (*FGFR*) alterations. Sample S6 (*FGFR3::TACC3*) exhibited the highest fusion burden, with 11 ancillary fusion events, including *SLC25A23::SHLD1* and *STX5::CHRM1*. Similarly, sample S4 (*FGFR1::TACC1*) harbored five additional fusion transcripts. In contrast, samples S3 and S8 (both *KIF5B::RET*-positive) demonstrated relatively lower fusion burden, each with a single ancillary event (*CRYBG3::NCOR1* and *SPIDR::LINC00824*, respectively) as illustrated in

Figure 4. Collectively, these findings suggest that fusion burden derived from WTS may provide a more comprehensive representation of tumor transcriptomic complexity beyond the identification of primary driver alterations.

In addition to primary oncogenic drivers, whole transcriptome sequencing (WTS) revealed a diverse spectrum of ancillary fusion transcripts across the cohort. As detailed in Table 4, a total of 31 events were annotated according to reading frame and bioinformatics confidence.

While all primary drivers were confirmed to be in-frame, the ancillary landscape exhibited significant

Table 2. Summary of Driver and Ancillary Fusions Detected by Whole Transcriptome Sequencing

S. no.	Sample ID	Fusion detected by targeted panel sequencing	Fusion detected by RNA seq	Tools used for fusion detection
1	S1	KIF5B (15)::RET (12)	BMS1::KIF5B,PRKCZ::HSPG2,TRIO::TASP1	Star fusion, Arriba, Fusion catcher
2	S2	EML4::ALK	AKIRIN2::CDKAL1, ZFAND3::CASC15	Star fusion, Arriba, Fusion catcher
3	S3	KIF5B (15)::RET (12)	CRYBG3::NCOR1	Star fusion, Arriba, Fusion catcher
4	S4	FGFR1 (17)::TACC1 (7)	ESYT1::ANO6,GDF11::EP400,RAB31::PTPRB,VAPA::PPTC7 ZNF140::TMTC2	Star fusion, Arriba, Fusion catcher
5	S5	EML4 (6)::ALK	FNIP2::UGT2B27P,PPP3CA::SH3D19,UBE2D3::SPOCK3	Star fusion, Arriba, Fusion catcher
6	S6	FGFR3 (17)::TACC3 (6)	KCTD2::F12,MIR23AHG::PODNL1,NUDCD3::NACAD,OCLN::SAMHD1 SLC25A23::SHLD1,SLC25A23::THOP1,SPINT2::SIPA1L3,STX5::CHRM1 TAOX1::PIPOX,TSC2::MAPK8IP3,WWP2::SNTB2	Star fusion, Arriba, Fusion catcher
7	S7	CD74 (6)::ROS1 (34)	MAP3K3::FAM20A, PAFAH1B1::YWHAE	Star fusion, Arriba, Fusion catcher
8	S8	KIF5B (15)::RET (12)	SPIDR::LINC00824	Star fusion, Arriba, Fusion catcher
9	S9	EML4 (13)::ALK (20)	MTSS1::SLC25A21,PHRF1::DEPTOR,ZBED3::PRKCE	Star fusion, Arriba, Fusion catcher

Table 3. Chromosomal Breakpoints of Fusion Transcripts (hg38) with Corresponding Fusion Types

Sr. No	Sample Id	Ancillary fusions	Left Breakpoint	Right Breakpoint	Fusion type
1	S1	BMS1::KIF5B	chr10:42802218:+	chr10:32023036:-	Inversion
		PRKCZ::HSPG2	chr1:2150978:+	chr1:21896310:-	Inversion
		TRIO::TASPI	chr5:14143882:+	chr20:13483337:-	Translocation
2	S2	AKIRIN2::CDKAL1	chr6:87681620:-	chr6:20535351:+	Inversion
		ZFAND3::CASC15	chr6:38116739:+	chr6:22110747:+	Duplication
3	S3	CRYBG3::NCOR1	chr3:97936908:+	chr17:16047093:-	Translocation
4	S4	ESYT1::ANO6	chr12:56134428:+	chr12:45331295:+	Duplication
		GDF11::EP400	chr12:55743761:+	chr12:132076516:+	Deletion
		RAB31::PTPRB	chr18:9815222:+	chr12:70544663:-	Translocation
		VAPA::PPTC7	chr18:9914335:+	chr12:110551968:-	Translocation
5	S5	ZNF140::TMTC2	chr12:133083561:+	chr12:83132210:+	Duplication
		FNIP2::UGT2B27P	chr4:158833628:+	chr4:69018387:-	Inversion
		PPP3CA::SH3D19	chr4:101195916:-	chr4:151149561:-	Duplication
		UBE2D3::SPOCK3	chr4:102826485:-	chr4:167062537:-	Duplication
6	S6	KCTD2::F12	chr17:75062245:+	chr5:177406061:-	Translocation
		MIR23AHG::PODNL1	chr19:13842776:-	chr19:13938006:-	Duplication
		NUDCD3::NACAD	chr7:44484968:-	chr7:45086112:-	Duplication
		OCLN::SAMHD1	chr5:69504294:+	chr20:36912552:-	Translocation
		SLC25A23::SHLD1	chr19:6456420:-	chr20:5863024:+	Translocation
		SLC25A23::THOP1	chr19:6453981:-	chr19:2806917:+	Inversion
		SPINT2::SIPA1L3	chr19:38290575:+	chr19:38162253:+	Duplication
		STX5::CHRM1	chr11:62824166:-	chr11:62911178:-	Duplication
		TAOK1::PIPOX	chr17:29482288:+	chr17:29044859:+	Duplication
		TSC2::MAPK8IP3	chr16:2070578:+	chr16:1764115:+	Duplication
		WWP2::SNTB2	chr16:69799295:+	chr16:69299590:+	Duplication
7	S7	MAP3K3::FAM20A	chr17:63652656:+	chr17:68555743:-	Inversion
		PAFAH1B1::YWHAE	chr17:2638320:+	chr17:1365058:-	Inversion
8	S8	SPIDR::LINC00824	chr8:47440542:+	chr8:128428384:-	Inversion
9	S9	MTSS1::SLC25A21	chr8:124585087:-	chr14:36875004:-	Translocation
		PHRF1::DEPTOR	chr11:582081:+	chr8:120049576:+	Translocation
		ZBED3::PRKCE	chr5:77087111:-	chr2:46086208:+	Translocation

variation. Out of the 31 ancillary fusions, 12 (38.7%) were identified as In-Frame, suggesting a potential for translated protein products that could influence tumor biology. The remaining 19 fusions (61.3%) were Out-of-Frame, representing non-productive transcriptional events that nonetheless serve as high-confidence markers of genomic instability. The vast majority of these calls (28/31; 90.3%) were categorized as High Confidence, with only three events in samples S4 and S6 being of medium confidence, further validating the robustness of the multi-tool detection pipeline. While the primary oncogenic drivers (*ALK*, *RET*, *ROS1*) were confirmed to be in-frame, a majority of the ancillary fusions identified via WTS were classified as out-of-frame. Despite their lack of coding potential, these events were detected with high confidence serving as robust indicators of localized genomic instability within the cohort. Sample S6, which possessed an *FGFR3::TACC3* primary driver, exhibited the highest degree of complexity, containing 5 in-frame and 6 out-of-frame ancillary fusions. This

specimen-specific density of rearrangements highlights the superior resolution of the WTS pipeline compared to standard targeted panels. In Figure 5, Stacked bar chart illustrating the distribution of 31 novel ancillary fusions across the study cohort (S1–S9). Bars are categorized by Sample ID and color-coded by the predicted reading frame: green represents in-frame fusions with potential functional significance, while orange represents out-of-frame transcripts indicative of genomic instability.

The analysis of ancillary fusion distribution across driver types revealed distinct patterns of genomic complexity. Specimens with *FGFR* family drivers exhibited the highest burden of ancillary events. The *FGFR3::TACC3* driver (Sample S6) was associated with a median of 5 in-frame and 6 out-of-frame fusions, while the *FGFR1::TACCI* driver (Sample S4) showed a median of 3 in-frame events alongside 2 out-of-frame events. In contrast, *EML4::ALK* driven samples showed a median of 1 in-frame fusion and 2 out-of-frame fusions, while *KIF5B::RET* driven samples displayed a

Table 4. Detailed Profile of Ancillary Fusion Transcripts Detected by WTS

Sr. No	Sample Id	Ancillary fusions	Confidence	Reading frame
1	S1	BMS1::KIF5B	High	Out of Frame
		PRKCZ::HSPG2	High	In Frame
		TRIO::TASP1	High	In Frame
2	S2	AKIRIN2::CDKAL1	High	Out of Frame
		ZFAND3::CASC15	High	Out of Frame
3	S3	CRYBG3::NCOR1	High	Out of Frame
4	S4	ESYT1::ANO6	High	In Frame
		GDF11::EP400	Medium	Out of Frame
		RAB31::PTPRB	High	In Frame
		VAPA::PPTC7	High	In Frame
		ZNF140::TMTC2	Medium	Out of Frame
5	S5	FNIP2::UGT2B27P	High	Out of Frame
		PPP3CA::SH3D19	High	Out of Frame
		UBE2D3::SPOCK3	High	In Frame
6	S6	KCTD2::F12	High	Out of Frame
		MIR23AHG::PODNL1	High	In Frame
		NUDCD3::NACAD	Medium	Out of Frame
		OCNL::SAMHD1	High	Out of Frame
		SLC25A23::SHLD1	High	Out of Frame
		SLC25A23::THOP1	High	In Frame
		SPINT2::SIPA1L3	High	In Frame
		STX5::CHRM1	High	Out of Frame
		TAOK1::PIPOX	High	Out of Frame
		TSC2::MAPK8IP3	High	In Frame
		WWP2::SNTB2	High	In Frame
		MAP3K3::FAM20A	High	Out of Frame
7	S7	PAFAH1B1::YWHAE	High	Out of Frame
8	S8	SPIDR::LINC00824	High	Out of Frame
9	S9	MTSS1::SLC25A21	High	In Frame
		PHRF1::DEPTOR	High	Out of Frame
		ZBED3::PRKCE	High	In Frame

median of 0 in-frame ancillary events and 1 out-of-frame events, indicating relatively higher genomic stability in these subgroups as mentioned in Table 5. Across all driver types, the median number of out-of-frame fusions remained consistently present (ranging from a median of 1 to 6 per driver), further supporting the use of these events as a universal marker for transcriptional noise in fusion-positive tumors.

Technical Validation of High-Complexity Sample (S6) ancillary fusion landscape

The Whole Transcriptome Sequencing (WTS) pipeline revealed a high-density landscape of 11 ancillary fusions in Sample S6, all of which were independently validated by STAR-Fusion, Arriba, and FusionCatcher. As shown in Table 6, these events were supported by robust junction (split) read counts across all algorithms, confirming the

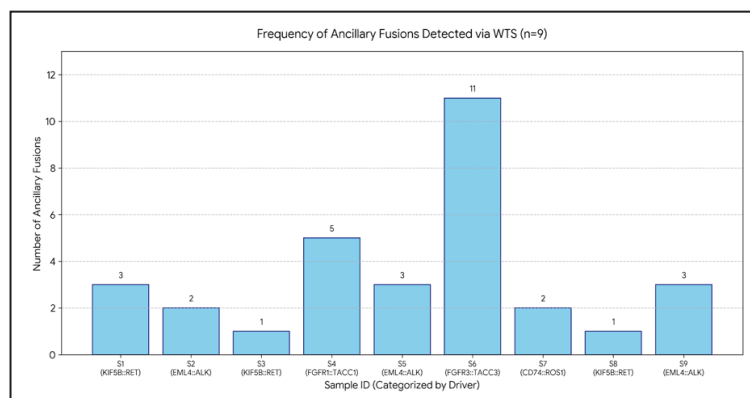


Figure 4. Distribution of Ancillary Fusion Events across the Study Cohort

Table 5. Statistical Summary of Ancillary Fusion Burden by Driver Subgroup

Primary Driver	n (Samples)	Median In-Frame Fusions (Range)	Median Out-of-Frame Fusions (Range)
<i>FGFR3::TACC3</i>	1	5 (5–5)	6 (6–6)
<i>FGFR1::TACC1</i>	1	3 (3–3)	2 (2–2)
<i>EML4::ALK</i>	3	1 (0–2)	2 (1–2)
<i>CD74::ROS1</i>	1	0 (0–0)	2 (2–2)
<i>KIF5B::RET</i>	3	0 (0–2)	1 (1–1)

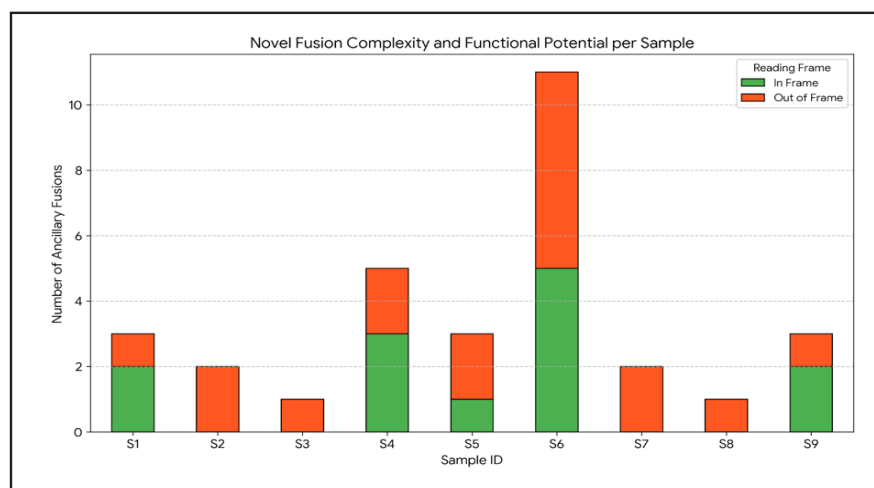


Figure 5. Molecular Characterization of New Ancillary Fusion Transcripts

high sensitivity and reproducibility of the consensus approach. In particular, certain fusions identified by Arriba demonstrated absent or limited spanning fragment support despite positive evidence reported by other callers. These differences likely reflect variations in algorithm-specific filtering strategies, read prioritization methods, and confidence thresholds used during fusion prediction. Accordingly, low-support ancillary fusion events were interpreted cautiously in this exploratory analysis. However, the consistent detection of high-count spanning pairs in FusionCatcher and junction reads in STAR-Fusion confirms that these represent genuine structural rearrangements rather than sequencing artifacts. Variability in supporting read evidence was observed across fusion detection algorithms for several ancillary

fusion events, and low-support calls were interpreted cautiously within this exploratory analysis.

Of the 11 ancillary fusions, 5 (45.4%) were predicted to be in-frame, including the high-support SPINT2::SIPA1L3 and MIR23AHG::PODNL1 events. The SPINT2::SIPA1L3 fusion demonstrated the strongest transcriptional support, with a cumulative total of over 100 supporting fragments across the three tools. Every identified fusion in this high-complexity sample was cross-verified by all three tools (3/3 concordance), providing a rigorous benchmark for the genomic instability observed in this specimen. Differences in junction and spanning read support across fusion callers may reflect algorithm-specific filtering criteria and confidence thresholds. Low-support ancillary fusion events should be interpreted cautiously. Because

Table 6. Comprehensive Technical Support for All 11 Ancillary Fusions in Sample S6

Ancillary Fusion	STAR-Fusion (J/S)	Arriba (J/S)	FusionCatcher (J/S)	Reading Frame
SPINT2::SIPA1L3	34 / 4	28 / 3	24 / 86	In Frame
OCN::SAMHD1	31 / 2	27 / 2	18 / 60	Out of Frame
MIR23AHG::PODNL1	18 / 2	16 / 0	20 / 55	In Frame
SLC25A23::THOP1	13 / 0	11 / 0	14 / 39	In Frame
KCTD2::F12	10 / 0	9 / 0	12 / 27	Out of Frame
WWP2::SNTB2	6 / 0	5 / 0	6 / 14	In Frame
SLC25A23::SHLD1	5 / 0	5 / 0	7 / 7	Out of Frame
TSC2::MAPK8IP3	5 / 1	5 / 0	6 / 6	In Frame
NUDCD3::NACAD	4 / 0	4 / 0	6 / 13	Out of Frame
STX5::CHRM1	4 / 0	4 / 0	3 / 12	Out of Frame
TAOK1::PIPOX	4 / 0	4 / 0	5 / 17	Out of Frame

*J/S = Junction Reads / Spanning Fragments

orthogonal experimental validation was not performed, ancillary fusion transcripts identified in this study should be interpreted as computationally predicted candidate events within an exploratory analytical framework.

Discussion

In this study, whole transcriptome sequencing (WTS) revealed a complex and heterogeneous landscape of fusion transcripts across nine tumor samples, extending beyond canonical oncogenic drivers identified by targeted panel sequencing. By integrating multiple fusion detection tools, including StumoTAR-Fusion, Arriba, and FusionCatcher, we identified a total of 31 ancillary fusion events, highlighting the expanded detection capability of RNA sequencing approaches. These findings are consistent with previous reports demonstrating that RNA-seq enables unbiased, transcriptome-wide discovery of gene fusions, including low-abundance and previously uncharacterized events [32, 33]. A notable observation in this study is the limited concordance between targeted panel sequencing and WTS-derived fusion calls. While targeted approaches reliably identified clinically established drivers such as *KIF5B::RET*, *EML4::ALK*, and *FGFR3::TACC3*, WTS predominantly uncovered additional, non-overlapping fusion transcripts. This discrepancy likely reflects fundamental methodological differences: targeted panels are optimized for sensitivity toward predefined fusion events, whereas RNA-seq provides a broader but expression-dependent detection framework. Several ancillary fusion transcripts were retained for analysis despite lacking clear protein-coding potential, owing to their high-confidence detection supported by split reads and spanning pairs. The presence of these non-productive fusion transcripts likely reflects underlying chromosomal instability and transcriptional noise, a phenomenon that has been widely reported in both normal and cancerous tissues [34, 35]. Such events, although not directly contributing to oncogenic transformation, may provide insight into the structural and transcriptional dynamics of tumor genomes. The fusion landscape observed across the cohort was highly heterogeneous, with most events being non-recurrent and sample-specific. While many of these fusions may represent passenger alterations, a subset involved genes with established roles in oncogenic signaling and regulatory pathways. Notably, the detection of *PAFAH1B1::YWHAE*, a previously characterized oncogenic fusion, supports the robustness of the RNA-seq pipeline. In addition, some ancillary fusion transcripts identified in this study involved genes associated with pathways such as mTOR, *MAPK*, and calcium signaling. However, the biological and functional relevance of these candidate fusion events remains uncertain because no orthogonal or functional validation was performed. The identification of fusion events involving genes associated with mTOR and *MAPK* signaling pathways warrants further investigation to determine potential biological relevance [36–38]. The identification of *CRYBG3::NCOR1* further suggests disruption of transcriptional regulatory mechanisms, which may contribute to aberrant gene

expression patterns. The convergence of fusion partners on critical signaling pathways supports the biological plausibility of a subset of detected events. In particular, the potential interaction between mTOR and *MAPK* signaling via *TSC2::MAPK8IP3* highlights a possible mechanism for pathway cross-talk in tumor progression. Similarly, alterations involving *PPP3CA* may influence calcium-dependent signaling processes, which are increasingly recognized as important regulators of tumor growth and cellular homeostasis. Interestingly, *FGFR*-altered samples within this exploratory cohort demonstrated a relatively higher number of ancillary fusion transcripts compared with other cases. However, because only a limited number of *FGFR*-positive samples were included, this observation should be interpreted cautiously and requires validation in larger independent cohorts. This observation is consistent with previous reports linking *FGFR* rearrangements to increased genomic instability and structural variation, suggesting that different oncogenic drivers may be associated with distinct mechanisms of chromosomal rearrangement and transcriptomic complexity [39]. Despite the advantages of WTS, RNA-seq-based fusion detection remains susceptible to false positives arising from alignment artifacts, read-through transcription, and sequencing-related biases. To address this limitation, we employed a consensus-based approach incorporating multiple fusion detection algorithms and prioritized high-confidence events supported across tools. Such strategies have been recommended in previous methodological studies to improve the reliability of fusion detection [40, 41]. Overall, our findings highlight the complementary role of RNA sequencing in expanding the detectable fusion landscape beyond canonical drivers. While targeted panels remain essential for identifying clinically actionable fusions, WTS provides a broader perspective on tumor transcriptomic complexity and may facilitate the discovery of novel fusion events with potential biological or therapeutic significance. Furthermore, there is a growing consensus that the cumulative fusion burden encompassing both functionally productive and non-productive transcripts acts as a proxy for systemic genomic instability and clonal evolution. Contemporary pan-cancer studies underscore that an elevated frequency of structural variants correlates with a heightened diversity of fusion events, providing a high-resolution signature of continuous chromosomal flux and disrupted transcriptional architecture [42, 43]. Recent progress in transcriptome-wide fusion discovery underscores the necessity of converging computational algorithms with biological frameworks. Such an integrated approach is vital for delineating functionally active chimeric transcripts from the background of passenger events that arise during genomic destabilization [44, 45]. By synthesizing RNA-seq findings with genomic and epigenomic datasets, we can achieve a deeper understanding of the structural and regulatory complexity of the tumor. This synergy between diverse data types facilitates a more precise identification of molecular drivers with direct clinical relevance [46]. Although clinically actionable driver fusions have been

extensively studied, the broader landscape of ancillary fusion transcripts detected through transcriptome-wide sequencing remains incompletely characterized. Understanding the diversity and heterogeneity of these secondary fusion events may provide additional insight into transcriptomic complexity in fusion-positive tumors.

Limitations

This study has several limitations, including the small cohort size and retrospective FFPE-based design, which limited statistical analysis and broader generalizability of subgroup-specific observations. Orthogonal experimental validation of ancillary fusion transcripts was not performed, and RNA-seq-based fusion detection remains susceptible to false-positive calls arising from alignment artefacts, transcriptional noise, and degraded FFPE-derived RNA despite the use of a consensus multi-caller approach. In addition, targeted RNA sequencing and whole transcriptome sequencing were performed using different sequencing platforms and analytical workflows, limiting direct comparative interpretation between methods. Therefore, the findings should be considered exploratory and hypothesis-generating pending validation in larger independent cohorts.

In conclusion, this study demonstrates that whole transcriptome sequencing providing a comprehensive and high-resolution view of the fusion landscape in driver-positive tumors, extending beyond the identification of canonical oncogenic rearrangements. By uncovering a diverse repertoire of ancillary fusion transcripts, our findings highlight the substantial inter-sample heterogeneity and underscore the complexity of tumor transcriptome that is not captured by targeted panel sequencing alone. The identification of 31 additional fusion events, many of which were not detected by conventional approaches, emphasizes the added value of RNA sequencing in revealing previously unrecognized transcriptomic alterations. Although a significant proportion of these ancillary fusions may represent passenger events or transcriptional noise, the involvement of genes associated with key oncogenic pathways including mTOR, *MAPK*, and calcium signalling suggests that a subset may have functional or biological relevance. However, the biological and functional relevance of these candidate fusion events remains uncertain because no orthogonal or functional validation was performed. Notably, the increased fusion burden observed in *FGFR*-altered samples points toward a potential link between specific oncogenic drivers and underlying genomic instability, indicating that fusion complexity may serve as an additional layer of tumor heterogeneity. These observations highlight the importance of considering not only primary driver events but also the broader fusion landscape when characterizing tumor biology. *FGFR*-altered samples within this limited cohort demonstrated a relatively higher ancillary fusion burden; however, this observation requires validation in larger cohorts. The high-confidence events identified in this study provide a valuable resource for future mechanistic investigations. Integrating transcriptomic

findings with genomic and clinical data in larger cohorts will be essential to determine their potential translational relevance. Overall, this study reinforces the complementary role of whole transcriptome sequencing in expanding the detectable spectrum of gene fusions and provides important insights into the complexity and diversity of fusion events in cancer. These findings lay the groundwork for future studies aimed at understanding the biological significance of secondary fusion burden and its potential implications in cancer diagnostics and therapeutic targeting. Collectively, these findings expand the current understanding of the fusion landscape and provide a resource for future studies ultimately contribute to the identification of novel biomarkers and therapeutic targets in fusion-driven malignancies. We explored ancillary fusion transcript diversity in a small real-world FFPE cohort using WTS and observed substantial transcriptomic heterogeneity beyond canonical drivers. The potential biological relevance of these candidate fusion transcripts remains uncertain and requires further functional investigation.

Acknowledgements

The author gratefully acknowledges the guidance of Dr. Rajan Kumar Singh and the support of Unipath Specialty Laboratory Ltd. for providing access to laboratory facilities and advanced NGS platforms, which were critical for the successful completion of this work. My sincere thanks also goes to Dr. Spandan Chaudhary & Dr. Naman Mangukia for their technical support, collaborative spirit, and enriching scientific discussions.

Conflict of Interest

The author(s) declare that there is no conflict of interest regarding the publication of this paper.

Funding

This study did not receive any specific funding support from funding agencies in the public, commercial or not-for-profit sectors.

Ethics Approval

This study was approved by the Hospital Ethics Committee with the approval number IECHR-SAINATH HOSPITAL/AHMC/178.

References

1. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, Temple-Smolkin RL, Voelkerding KV, Nikiforova MN. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels. *The Journal of Molecular Diagnostics*. 2017 05;19(3):341-365. <https://doi.org/10.1016/j.jmoldx.2017.01.011>
2. R. Ptashkin, H. Won, J. Chang, D. Halpenny, and M. Alison, "Downloaded from clincancerres.aacrjournals.org on April 26, 2019. © 2019 American Association for Cancer Research.," 2019, doi: 10.1158/1078-0432.CCR-19-0225.
3. Mertens F, Johansson B, Fioretos T, Mitelman F. The emerging complexity of gene fusions in cancer. *Nature Reviews*

- Cancer. 2015 06;15(6):371-381. <https://doi.org/10.1038/nrc3947>
4. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, Berlin AM, Sivachenko A, et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nature Methods*. 2013 07;10(7):623-629. <https://doi.org/10.1038/nmeth.2483>
 5. Hedegaard J, Thorsen K, Lund MK, Hein AK, Hamilton-Dutoit SJ, Vang S, Nordentoft I, et al. Next-Generation Sequencing of RNA and DNA Isolated from Paired Fresh-Frozen and Formalin-Fixed Paraffin-Embedded Samples of Human Cancer and Normal Tissue. *PLoS ONE*. 2014 05 30;9(5):e98187. <https://doi.org/10.1371/journal.pone.0098187>
 6. Costa V, Aprile M, Esposito R, Ciccociola A. RNA-Seq and human complex diseases: recent accomplishments and future perspectives. *European Journal of Human Genetics*. 2013 02;21(2):134-142. <https://doi.org/10.1038/ejhg.2012.129>
 7. Tarazona S, Furió-Tari P, Turrà D, Pietro AD, Nueda MJ, Ferrer A, Conesa A. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Research*. 2015 07 16;:gkv711. <https://doi.org/10.1093/nar/gkv711>
 8. Srinivasan M, Sedmak D, Jewell S. Effect of Fixatives and Tissue Processing on the Content and Integrity of Nucleic Acids. *The American Journal of Pathology*. 2002 Dec;161(6):1961-1971. [https://doi.org/10.1016/S0002-9440\(10\)64472-0](https://doi.org/10.1016/S0002-9440(10)64472-0)
 9. Lahens NF, Ricciotti E, Smirnova O, Toorens E, Kim EJ, Baruzzo G, Hayer KE, et al. A comparison of Illumina and Ion Torrent sequencing platforms in the context of differential gene expression. *BMC Genomics*. 2017 Dec;18(1):602. <https://doi.org/10.1186/s12864-017-4011-0>
 10. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011 07;475(7356):348-352. <https://doi.org/10.1038/nature10242>
 11. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology*. 2012 05;30(5):434-439. <https://doi.org/10.1038/nbt.2198>
 12. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: Comparison of ion torrent, pacific biosciences and illumina miseq sequencers. *BMC Genomics*. 2012;13(1):341. <https://doi.org/10.1186/1471-2164-13-341>
 13. Salipante S. Performance Comparison of Illumina and Ion Torrent Next- Generation Sequencing Platforms for 16S rRNA-Based Bacterial Community Profiling. 2014;. <https://doi.org/10.1128/AEM.02206-14>
 14. Zhao S. Comparison of stranded and non-stranded RNA-seq transcriptome profiling and investigation of gene overlap. *BMC Genomics*. 2015;16(1):1-14. <https://doi.org/10.1186/s12864-015-1876-7>
 15. Goodwin S, McPherson J, McCombie W. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet*. 2016;17(6):333-351. <https://doi.org/10.1038/nrg.2016.49>
 16. Metzker M. Sequencing technologies the next generation. *Nat. Rev. Genet*. 2010;11(1):31-46. <https://doi.org/10.1038/nrg2626>
 17. Maher C. Transcriptome sequencing to detect gene fusions in cancer. *Nature*. 2009;458(7234):97-101. <https://doi.org/10.1038/nature07638>
 18. Wang Q, Xia J, Jia P, Pao W, Zhao Z. Application of next generation sequencing to human gene fusion detection: Computational tools, features and perspectives. *Brief. Bioinform*. 2013;14(4):506-519. <https://doi.org/10.1093/bib/bbs044>
 19. Drilon A. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *N. Engl. J. Med*. 2018;378(8):731-739. <https://doi.org/10.1056/nejmoa1714448>
 20. Liu S. 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*. 2015;44(5):47. <https://doi.org/10.1093/nar/gkv1234>
 21. Conesa A. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016;17(1):1-19. <https://doi.org/10.1186/s13059-016-0881-8>
 22. Luthra R. A Targeted High-Throughput Next-Generation Sequencing Panel for Clinical Screening of Mutations, Gene Amplifications, and Fusions in Solid Tumors. *J. Mol. Diagnostics*. 2017;19(2):255-264. <https://doi.org/10.1016/j.jmoldx.2016.09.011>
 23. Peng H. Development and Validation of an RNA Sequencing Assay for Gene Fusion Detection in Formalin-Fixed, Paraffin-Embedded Tumors. *J. Mol. Diagnostics*. 2021;23(2):223-233. <https://doi.org/10.1016/j.jmoldx.2020.11.005>
 24. Esteve-Codina A. A comparison of RNA-Seq results from paired formalin-fixed paraffin-embedded and fresh-frozen glioblastoma tissue samples. *PLoS One*. 2017;12(1):1-18. <https://doi.org/10.1371/journal.pone.0170632>
 25. Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA quality from FFPE samples. *PLoS One*. 2007;2(12):1-7. <https://doi.org/10.1371/journal.pone.0001261>
 26. Zhao W, He X, Hoadley K, Parker J, Hayes D, Perou C. Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics*. 2014;15(1):1-11. <https://doi.org/10.1186/1471-2164-15-419>
 27. Bentley D. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008;456(7218):53-59. <https://doi.org/10.1038/nature07517>
 28. Dobin A. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. <https://doi.org/10.1093/bioinformatics/bts635>
 29. Haas B. STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq. *bioRxiv*. 2017;:120295.
 30. Uhrig S. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res*. 2021;31(3):448-460. <https://doi.org/10.1101/GR.257246.119>
 31. Nicorici D. FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014;:011650.
 32. Hu X. TumorFusions: An integrative resource for cancer-associated transcript fusions. *Nucleic Acids Res*. 2018;46(D1):1144-1149. <https://doi.org/10.1093/nar/gkx1018>
 33. Gao Q. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. *Cell Rep*. 2018;23(1):227-238 3. <https://doi.org/10.1016/j.celrep.2018.03.050>
 34. Jia W. SOAPfuse: An algorithm for identifying fusion transcripts from paired-end RNA-Seq data. *Genome Biol*. 2013;14(2). <https://doi.org/10.1186/gb-2013-14-2-r12>
 35. Gorohovski A, Tagore S, Palande V, Malka A, Raviv-Shay D, Frenkel-Morgenstern M. ChiTaRS-3.1-the enhanced chimeric transcripts and RNA-seq database matched with protein-protein interactions. *Nucleic Acids Res*.

- 2017;45(D1):790–795,. <https://doi.org/10.1093/nar/gkw1127>.
36. Clapham D. Calcium Signaling. *Cell*. 2007;131(6):1047–1058,. <https://doi.org/10.1016/j.cell.2007.11.028>.
37. Laplante M, Sabatini D. MTOR signaling in growth control and disease. *Cell*. 2012;149(2):274–293,. <https://doi.org/10.1016/j.cell.2012.03.017>.
38. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer*. 2007;7(4):233–245,. <https://doi.org/10.1038/nrc2091>.
39. Singh D. Transforming fusions of *FGFR* and *TACC* genes in human glioblastoma. *Science*. 2012;337(6099):1231–1235,. <https://doi.org/10.1126/science.1220834>.
40. Zhang J. INTEGRATE: Gene fusion discovery using whole genome and transcriptome data. *Genome Res*. 2016;26(1):108–118,. <https://doi.org/10.1101/gr.186114.114>.
41. Davidson N, Majewski I, Oshlack A. JAFFA: High sensitivity transcriptome-focused fusion gene detection. *Genome Med*. 2015;7(1):1–12,. <https://doi.org/10.1186/s13073-015-0167-x>.
42. Berger M, Mardis E. The emerging clinical relevance of genomics in cancer medicine. *Nat. Rev. Clin. Oncol*. 2018;15(6):353–365,. <https://doi.org/10.1038/s41571-018-0002-6>.
43. Zhang C, Leibowitz M, Pellman D. Chromothripsis and beyond: Rapid genome evolution from complex chromosomal rearrangements. *Genes Dev*. 2013;27(23):2513–2530,. <https://doi.org/10.1101/gad.229559.113>.
44. Carrara M. State-of-the-art fusion-finder algorithms sensitivity and specificity. *Biomed Res. Int*. 2013;2013. <https://doi.org/10.1155/2013/340620>.
45. Haas B, Dobin A, Li B, Stransky N, Pochet N, Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol*. 2019;20(1):1–16,. <https://doi.org/10.1186/s13059-019-1842-9>.
46. Campbell P. Pan-cancer analysis of whole genomes. *Nature*. 2020;578(7793):82–93,. <https://doi.org/10.1038/s41586-020-1969-6>.



This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.