BMC Cancer

RESEARCH Open Access

Development and application of a whole transcriptome sequencing assay for the detection of gene fusions in clinical cancer specimens

Songchen Zhao[1], Xinhua Du[2], Yan Zhang[2], Jing Bai[2,3], Lu Meng[2], Xuefei Li[4], Jiaxin Ma[2], HeYu Sheng[2], Xiaorui Fu[2], Yanfang Guan[2], Yuting Yi[2], Ling Yang[2], Xuefeng Xia[2], Xin Yi[2], Xinxin Tan[5,6*] and Caicun Zhou[7*]

Abstract

Background Gene fusions are an important driver of cancer and require rapid and accurate detection to guide clinical decisions. However, the performance characteristics of whole transcriptome sequencing (WTS) for the detection of gene fusions have not been thoroughly investigated.

Methods We developed a novel WTS-based assay for the detection of gene fusions, *MET* exon 14 skipping and *EGFR* VIII alterations in clinical samples.

Results We defined a DV200 value ≥ 30% as the threshold for RNA degradation, RNA input, fusion expression and number of mapped reads greater than 100 ng, 40 copies/ng and 80 Mb for optimal sensitivity of the WTS assay. Our assay successfully identified 62 out of 63 known gene fusions, achieving a sensitivity of 98.4%. The specificity of the assay was 100%, as no fusions were detected in the 21 fusion-negative samples. Good repeatability and reproducibility were observed in replicates, except for the *TPM3::NTRK1* fusion, which was expressed below the threshold. Of all fusions identified in 101 NSCLC samples, 68.9% (20/29) were potentially actionable, compared to 20% in pan-cancer samples. In addition to actionable fusions, we also identified many fusions with potential diagnostic and prognostic value in pan-cancer.

Conclusions We have developed a novel WTS assay with high sensitivity, specificity, repeatability and reproducibility. This assay can identify potentially actionable gene fusions and provides valuable insights into the fusion landscape in various cancers, which may help guide treatment decisions and aid in diagnosis and prognosis.

Keywords Gene fusions, WTS assay, Actionable fusions, Cancer specimens

*Correspondence:
Xinxin Tan
tanxx@geneplus.org.cn
Caicun Zhou
caicunzhoudr@163.com
[1]Department of Medical Oncology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China
[2]Geneplus-Beijing Institute, Beijing, China
[3]College of Future Technology, Peking University, Beijing, China
[4]Department of Lung Cancer and Immunology, Tongji University Affiliated Shanghai Pulmonary Hospital, Shanghai, China
[5]Geneplus-Shenzhen Clinical Laboratory, Shenzhen, Guangdong, China
[6]Xiangya School of Pharmaceutical Sciences, Central South University, Changsha, Hunan, China
[7]Department of Medical Oncology, Shanghai East Hospital, Shanghai, China

Introduction

Gene fusions are a hallmark of some cancers. They result from genomic rearrangements, including chromosomal translocations, inversions, interstitial deletions, or duplications, which lead to the continued activation of proto-oncogenes and promote tumor initiation and development [1, 2]. Available data indicate that gene fusions occur in all malignancies and account for 20% of human cancer morbidity [3].

Lung cancer has received increasing attention in recent years, with the highest incidence and mortality rates of any cancer [4]. As the most common histological type of lung cancer, non-small cell lung cancer (NSCLC) accounts for about 85% of all lung cancers [5]. Most patients with NSCLC have actionable fusions (and splicing variants), including gene fusions in *ALK*, *ROS1*, *RET*, and *NTRK*, detected in approximately 5%, 2%, 1%, and 0.1%, respectively [6, 7]. Although the incidence of fusion genes is not as high in NSCLC, they can effectively guide clinical treatment to benefit more patients, especially with targeted therapies [6, 8–10]. Exon 14 skipping caused by mutations near the exon 14 splice sequence of *MET* has attracted widespread attention as an emerging therapeutic target in NSCLC [11]. The frequency of *MET* exon14 (*MET* EX14) skipping is approximately 4% in lung adenocarcinoma and as high as 22% in lung sarcomatoid carcinoma [12]. *MET* EX14 skipping may increase the stability of *MET* and cause activation of downstream signaling pathways such as RAS-RAF-MEK-ERK, and PI3K/AKT to promote the growth and proliferation of tumor cells [13]. Patients with NSCLC harboring *MET* EX14 skipping had better responses to *MET* inhibitors in clinical trials including crizotinib and cabozantinib [11, 12].

Beyond actionable fusions, other fusions hold clinical significance as well. Gene fusions can serve as biomarkers to distinguish various cancer types [14, 15]. For example, *RELA* fusion-positive ependymomas are now a distinct entity in the WHO classification of CNS tumors, as this subset of ependymomas tends to be much more aggressive than other supratentorial ependymomas [16–18]. Similarly, the *EVT6*::*NTRK3* fusion is defined as a diagnostic biomarker, found in 92% of human secretory breast cancers [19]. In addition, specific fusions can diagnose sarcomas. For example, the *SS18*::*SSX* fusion gene is a characteristic marker of synovial sarcoma [20], and *COL1A1*::*PDGFB* is specific to dermatofibrosarcoma protuberans [21]. Hence, fusion genes are therefore becoming increasingly important as diagnostic and therapeutic targets in cancer. Rapid and accurate identification of fusion genes and RNAs is crucial for patients.

Molecular techniques, such as fluorescence in situ hybridization (FISH), and reverse transcription PCR (RT-PCR) have been predominantly used for fusion gene diagnosis [22–24]. However, traditional molecular techniques require the genes involved to be known or suspected, are unable to identify novel fusion gene partners or resolve complex structural rearrangements, and usually only a few prioritized assays are performed, due to cost and time constraints [22–25]. The next-generation sequencing (NGS) approaches, including DNA and RNA sequencing, have proven effective in the detection of fusion genes [25]. Several studies have shown that RNA-based template NGS is significantly superior to DNA template NGS in the detection of fusions/rearrangements and translocations [26, 27]. RNA sequencing, also known as whole transcriptome sequencing (WTS), could provide global and unbiased detection of fusion genes, identify known and novel fusions within any expressed gene [14], and provide additional clinical benefits for tumor patients [28].

To date, more than 20 algorithms for fusion detection by RNA-sequencing have been published, but the identification of fusions by RNA-sequencing remains challenging and a high rate of false positives is common [29–32], which hinders its clinical application. Poor-quality RNA will decrease the number and quality of reads, thereby reducing the likelihood of detecting fusion transcripts. Therefore, careful assessment of the quality of RNA quality and fusion calls, as well as appropriate filtering strategies, are required to enable reliable application of this technology in diagnostics. However, few studies have evaluated the performance characteristics of WTS for the detection of gene fusions in a clinical laboratory setting. Furthermore, while many studies have filtered the reported extent of gene fusion [1] to improve accuracy, few have included information on splicing mutations, such as *MET* EX14 skipping. Timely and accurate detection of such skipped exons is critical for informing clinical treatments. In this study, we evaluated the performance characteristics of a designed WTS assay capable of detecting gene fusions, and exon-skipping variants, providing valuable insights for clinical application.

Materials and methods

Cohort selection for WTS assay development and validation

The majority of the development and validation work was performed using cohort 1, which included four fusion-positive commercial samples, one fusion-negative commercial sample, one *NTRK* fusion-positive transfected cell line, 42 formalin-fixed paraffin-embedded (FFPE) tumor samples, and 10 normal tissue samples (Table S1 and S2). The 42 tumor samples, consisted of 28 NSCLC, eight soft tissue sarcomas, three gliomas, two cases of diffuse large B-cell lymphoma (DLBCL), and one gastric cancer (Table S1 and S2).

A retrospective cohort (cohort 2) of **191** FFPE samples was used to assess RNA integrity in this study (Table S1).

In addition, we also included 156 clinical pan-cancer samples (cohort 3), including 45 soft tissue sarcomas, 25 colorectal cancers, 19 lung cancers, 12 breast cancers, 12 liver cancers, 22 uterine sarcomas, and 21 others (Table S1). These 156 samples were used to validate the newly developed WTS assay to further evaluate the clinical QC performance, which included RNA yield, degradation level, microbiological contamination, and bioinformatic analysis QC parameters. All experiments in this study were conducted in accordance with the World Medical Association's Code of Ethics (Declaration of Helsinki) and were approved by the Ethics Committee of Shanghai Pulmonary Hospital.

RNA extraction, library preparation and sequencing

A previous study demonstrated that compared with the newly prepared FFPE samples, the RNA extracted from FFPE samples stored at room temperature or 37℃ for one year was severely degraded whereas the quality of RNA extracted from FFPE samples stored at 4℃ for one year was relatively complete [33]. In this study, total RNA was extracted from FFPE samples stored at 4℃ for less than one year using RNeasy FFPE Kit (Qiagen, Germany). Furthermore, all FFPE samples should meet these criteria: 10 sections of a 5×5 mm² tissue piece, with tumor content exceeding 20%.

As the initial step, an accurate assessment of RNA quality determines the type of library preparation and sequencing parameters are required [34]. Total RNA was assessed and quantified using the NanoDrop 8000 (Thermo Fisher Scientific, Germany), Qubit 3.0 (life Invitrogen, USA), and the Agilent 2100 Bioanalyzer system (Agilent Technologies, USA). Ribosomal RNA was then removed using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (NEB, USA).

For RNA samples with a DV200 (percentage of RNA fragments > 200 nucleotides) of 50% or less, the fragmentation step was skipped, and the samples proceeded directly to library preparation. After rRNA depletion and optional fragmentation, cDNA synthesis and NGS library preparation were performed using the NEBNext ® Ultra™ II Directional RNA Library Prep Kit (NEB, USA) according to the manufacturer's protocol, except for substituted adaptor and index primers generated by Gene + seq 2000 (Geneplus-Suzhou, China).

The library was quantified using Qubit 3.0 (Life Technologies, USA), and quality was assessed using the LabChip GX Touch (PerkinElmer, USA). Sequencing was performed on the Gene + seq 2000 instrument, with each sample generating an average of approximately 25 gigabases (Gb) of data, consisting of 100 bp paired-end reads.

Reportable range

To assist in the filtering of gene fusions of potential diagnostic, prognostic, or therapeutic value, we have developed a list of genes that are considered reportable in this assay. Initially, we searched the NCCN clinical practice guidelines, oncology expert consensus, databases such as FusionGDB [35], ChimerDB [36], and the Geneplus-Beijing Internal Database, and the biomedical literature for gene fusions identified in both solid tumors and hematological cancers. The selection criteria for reportable fusions were: (i) fusions detected with high frequency (≥ 20 occurrences); (ii) all known oncogenes, where new activating fusions might be discovered; (iii) tumor suppressor genes valuable for diagnosis or prognosis (e.g., *TP53*). Applying these criteria, we narrowed down the list of reportable genes to 553, reducing the number of mRNA-encoding genes analyzed in the transcriptome from approximately 22,000 to 553 (Table S3). This list includes two gene splicing variants (*MET* exon 14 skipping and *EGFR* VIII) and 21 kinase domain duplications (KDDs) (Table S3). In addition to the above filters, we restricted fusion partners for most genes to refine the scope of the report. Among the 553 screened genes, fusion partners were limited for all except 63 (Table S3), with specific details in Table S4. This simplified the bioinformatics analyses, reduced false positives and results of uncertain clinical significance, and improved detection performance.

In the WTS assay, a fusion is classified as a "hotspot fusion" if it has been previously reported in the literature, NCCN clinical practice guidelines or clinical trials. The potentially actionable fusion defined in this WTS assay is functional and supported by A-D level clinical and experimental evidence as grouped by the Association for Molecular Pathology.

Bioinformatic analysis of fusion calling based on the WTS assay

We used Workflow Description Language (WDL) in conjunction with Docker to construct and manage the entire bioinformatics analysis workflow. The first step in the bioinformatics process was to filter the raw data to obtain clean data. This included the removal of terminal adaptor sequences and poor quality data using fastp (version: 0.19.5) [37], followed by the elimination of rRNA reads by aligning the reads to the rRNA database (downloaded from NCBI) using Bowtie2 (version:2.2.8) [38]. Clean reads were then aligned to the reference human genome (hg19) using STAR (version:2.7.8a) [39].

Haas et al. demonstrated that Arriba was one of the most accurate pipelines among 23 methods for the detection of fusion transcripts [29]. In this study, fusions were detected using Arriba 1.1.0 [40]. We updated Arriba's blacklist and adjusted the filter parameters for "known

fusions" to ensure the detection of specific fusions. We performed fusion identification in two separate runs with different parameter settings: In the first run, we used the default parameter for "-E MAX_E_VALUE" to identify fusions under standard conditions. In the second run, we increased "-E MAX_E_VALUE" from its default value of 0.3 to 1 to retrieve low-frequency variants that might have been missed in the initial analysis. For fusion annotation, we selected the most frequently reported or defined fusion transcript in the NCBI database (Table S5) as the primary transcript when multiple potential fusion transcripts were identified for the same fusion pattern.

Skipping variations were detected using an in-house software tool, FoRNA (Chinese Patent No. CN112164423). This tool can analyze the target rearrangement regions in bam files, specifically for *EGFR* (NM_005228.3, Exon1:Exon8) and *MET* (NM_000245.2, Exon13:Exon15), to directly count the split reads and spanning reads for each region. The final assessment was then based on a threshold of split reads + spanning reads ≥ 10 and spanning reads ≠ 0.

Quality control (QC) is critical to ensure reliable test results. Before generating the final clinical report, we calculated several QC parameters to assess both RNA quality and fusion calling. We use RNA-SeQC assessment (version:2.3.4) [41] to comprehensively assess the quality of RNA-seq results to ensure the reliability of subsequent analyses. The QC metrics include total bases, alignment and duplication rates; GC bias, rRNA content, alignment regions (exon, intron and intragenic), continuity of coverage, among others. The cross-contamination rate of RNA samples was assessed using Conta software (version:0.9.23, https://github.com/grailbio/conta/) to ensure the purity and accuracy of the RNA sample. In addition, microbiological analysis was performed using Kraken [42] to identify microorganisms in the samples and to exclude microbial contamination.

Cohort selection for clinical implementation of the WTS assay

We conducted a multi-center retrospective cohort (cohort 4) to detail the distribution of gene fusions across different cancer types (Table S1). This cohort includes 75 cases of gene fusion-positive soft tissue sarcomas, 63 cases of fusion-positive uterine cancers, 29 cases of fusion-positive lung cancers, 20 cases of fusion-positive central nervous system (CNS)/brain cancers, and 41 cases of fusion-positive tumor samples from nine other cancer types. Additionally, the cohort includes 70 cases of fusion-negative lung cancer patients, which helps to accurately depict the positive rate of gene fusions in lung cancer.

Statistical analysis

Statistical analysis was performed using R (version 4.1.2). Positive predictive value (PPV), positive percentage agreement (PPA), negative percentage agreement (NPA), and F1 score were calculated as:

$$positive\ predictive\ value\ (PPV) = 100 \times \frac{TP}{TP+FP} \quad (1)$$

$$positive\ percentage\ agreement\ (PPA) = 100 \times \frac{TP}{TP+FN} \quad (2)$$

$$negative\ percentage\ agreement\ (NPA) = 100 \times \frac{TN}{TN+FP} \quad (3)$$

$$F1\ score = 2 \times \frac{\mathrm{PPV}*\mathrm{PPA}}{\mathrm{PPV}+\mathrm{PPA}} \quad (4)$$

where TP = true positive, TN = true negative, FP = false positive, and FN = false negative.

Results

The WTS assay is an accurate fusion calling workflow

We developed an analysis pipeline for the detection of gene fusions, as described in the materials and methods (Fig. 1A). To assess the accuracy of fusion detection, all samples from cohort 1 were analyzed using our bioinformatics pipeline. Table S6 shows the fusions detected by our bioinformatics pipeline for the WTS assay, together with the fusions detected by the reference method, their fusion positions and the number of supporting reads for each fusion. The results from our bioinformatics pipeline provide detailed annotated results for each fusion, as well as an IGV visualization image link to help identify and check for alignment artifacts (Table S7). The fusion visualizations also show the fusion partners, orientation, retained exons in the fusion transcript, and an excerpt of the sequence around the breakpoint (Fig. 1B).

Figure 1A shows the diagram of the bioinformatics workflow for fusions calling in WTS assay. Figure 1B is the example of fusion visualization, which includes the fusion partners, orientation, retained exons in the fusion transcript, and an excerpt of the sequence around the breakpoint.

The results showed that 62 out of 63 known fusions were identified, giving a PPA (or sensitivity) of 98.4% (95% CI, 91.5-99.9%, Table S6). No false-positive fusions were detected, resulting in a PPV (or precision) of 100% (62/62, 95% CI, 94.2-100%). Thus, the F1 score for our WTS assay is 99.2%. One false-negative result (p-31, an *EML4::ALK* fusion) was observed in a patient with lung cancer, which had been detected by DNA sequencing (Table S6). RNA targeted sequencing was subsequently used to verify the fusion, but it was still not identified. This failure may be due to the expression of the fusion being below the detection threshold of this assay or due to non-productive DNA rearrangements that escape transcription and translation.

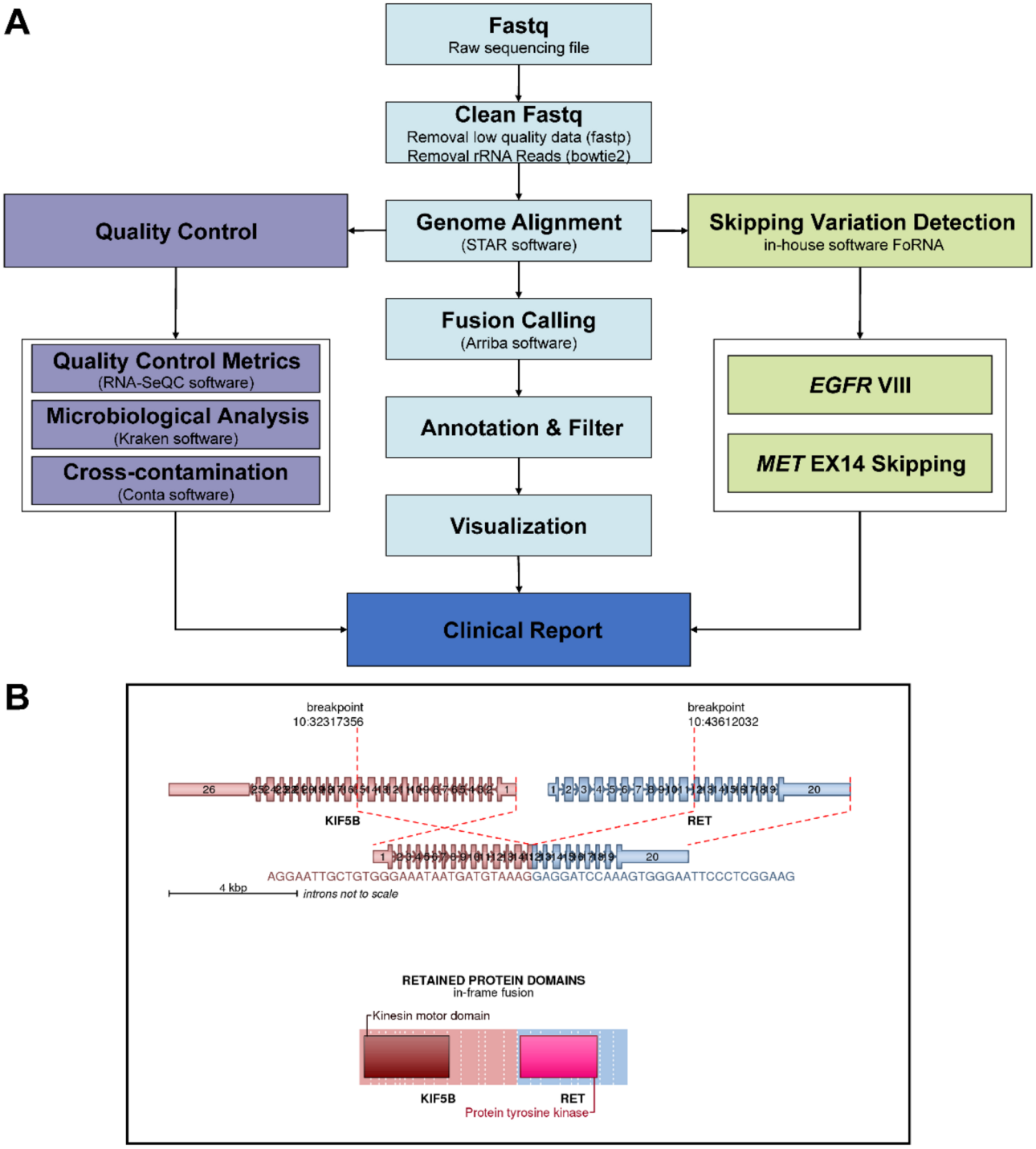

Fig. 1 Workflow of the WTS assay

To assess the NPA (or specificity) of the WTS assay, 10 normal tissue samples and 11 tumor samples with no fusion identified by other methods were analyzed (Table S6). All of these 21 negative control samples passed our QC criteria, confirming that the lack of detected fusions was not due to sample quality issues. Thus, the NPA value of the assay was 100% (95% CI, 84.5–100%) (Table S6). These samples were also used to determine the limit of blank (LoB) for this study (Table S2 and Table S6). The number of false-positive fusions was calculated for different supporting read thresholds, and the LoB of the assay was defined as the supporting reads at which the number of false-positive fusions across all samples was zero. Based on these results, the supporting read threshold for

considering a non-hotspot fusion positive was conservatively set at ten or more supporting reads. For all known fusions listed in Table S6, we found that no false positives were introduced when two supporting reads were used as the positivity threshold. Therefore, we set this value as the threshold for hotspot fusions with prior information.

To determine the precision of fusion detection, including repeatability or reproducibility, one commercial reference sample with four known fusions, four fusion-positive clinical samples, and one fusion-negative clinical sample were evaluated (Table S2 and Table S8). For each sample, we evaluated the repeatability between intra-run aliquots (samples run on the same plate under the same conditions) and the reproducibility across inter-run aliquots (samples run on different plates under different conditions), resulting in a total of $3+1+1=5$ replicates. The results showed that in the fusion-negative clinical sample, no fusions were detected in any of the replicates, whereas in the four fusion-positive clinical samples, the expected fusions were detected in all replicates. For the reference standard sample (s-01), the *TPM3::NTRK1* fusion was not detected in an inter-run aliquot due to its low expression level. The positive call rate for all eight fusions analyzed in the four fusion-positive clinical samples and one commercial reference sample was 39/40 (97.5%, 95% CI 87.1%-99.9%) (Table S8).

DV200 correlates positively with mapping rate

The storage time and conditions of FFPE samples can affect RNA quality and, consequently, the detection of gene fusions. RNA integrity, which affects QC pass rate and fusion gene detection, was assessed by calculating both the RNA Integrity Number (RIN) and the DV200 score. Retrospective analysis of internal sequencing data from 191 FFPE samples of varying quality showed that DV200 was positively correlated with the mapping rate (Spearman $\rho=0.48$, $P<0.0001$, Fig. 2A), whereas RIN value showed no correlation with the mapping rate (Spearman $\rho=0.07$, $P=0.34$, Fig. 2B). These results suggest that DV200 is more reliable than RIN for assessing the degree of RNA degradation in FFPE samples. When DV200 ≥ 30%, the qualification rate was 95%. Therefore, DV200 ≥ 30% was set as the threshold for acceptable levels of RNA degradation (Fig. 2C).

Figure 2A, 2B, and 2C illustrate the correlation between QC parameters and RNA integrity using 191 samples. Figure 2A shows the correlation between the mapping rate and DV200. Figure 2B displays the correlation between the mapping rate and the RIN value. Figure 2C presents the qualification rate for samples with varying RNA integrity, as indicated by different DV200 levels. Figure 2D illustrates the sensitivity of fusion detection at different RNA expression levels. Figure 2E and F show the support reads for each fusion (*MET* EX14 skipping, *CCDC6::RET*, *ETV6::NTRK3*, *KIF5B::RET*, *SLC34A2::ROS1*, and *FGFR3::TACC3*) across ten replicates, with varying thresholds for mapped reads. Figure 2G demonstrates fusion detection sensitivity at different levels of mapped reads. Figure 2H, 2I, and 2J present the clinical QC performance of the WTS assay, evaluated using 156 clinical pan-cancer samples. Figure 2H shows the distribution of DV200 values across the 156 clinical samples. Figure 2I depicts the QC qualification rate for samples with varying RNA integrity. Figure 2J shows the fusion-positive rate for different cancer types.

RNA input and fusion expression assessment

Minimum RNA requirements were established by measuring assay performance at different input levels (10–100 ng) in representative samples. One commercial sample and two FFPE clinical samples were evaluated at four different RNA inputs (10, 20, 50 and 100 ng) with five replicates per sample (Table S2 and Table S9). Sequencing failure rates and positive call rates were evaluated for each RNA input. A summary of the analytical performance for each RNA input range is shown in Table S9. The sequencing success rate was 100% for all input groups, except for the 10 ng RNA input. For the 10 ng RNA input, the sequencing failure rate was 13.3% (2/15), and the fusion positive call rate was 63.3% (19/30). Fusion positive call rates increased with higher RNA input: 73.3% (22/30) for 20 ng, 93.3% (28/30) for 50 ng, and 96.7% (29/30) for 100 ng. Based on these results, the standard RNA input should be ≥ 100 ng. Although lower amounts are acceptable, they result in lower sensitivity for fusion detection and a higher failure rate.

The level of fusion expression also affects the sensitivity of fusion detection. Two commercial samples with 11 fusions were evaluated fusion expression with quantitative PCR (qPCR) before (Table S2 and Table S10). We assessed each fusion with 10 replicates. The results showed that the positive call rate for fusion detection increased as the expression level increased (Table S10). The positive call rate was 10% (1/10) when the expression level was below 20 copies/ng (Fig. 2D and Table S10). For expression levels between 20 and 40 copies/ng, the positive call rate increased to 85% (17/20) (Fig. 2D and Table S10). When the expression level was between 40 and 60 copies/ng, the positive call rate increased to 95% (19/20) (Fig. 2D and Table S10), meeting the requirements for clinical testing. Overall, for all fusions with expression levels above 40 copies/ng, the detection sensitivity reached 98.8% (79/80) (Table S10). Therefore, 40 copies/ng was used as the minimum threshold for fusion detection sensitivity.

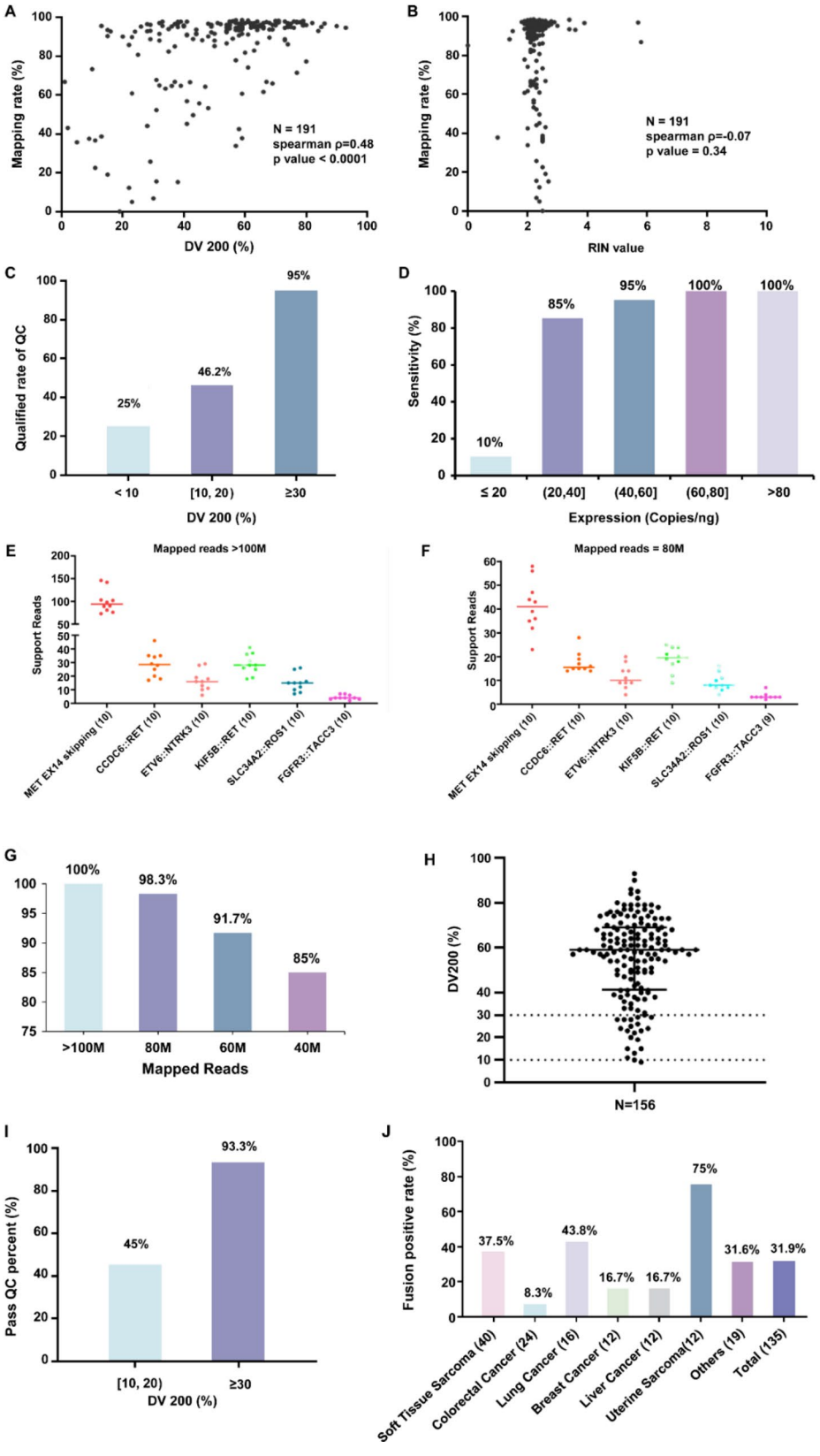

Fig. 2 QC threshold in the WTS assay

Fusion detection sensitivity is also affected by mapped reads

Post-analytical QC is essential to ensure reliable test results. For high quality libraries, mapping rates should exceed 80% [43]. Additionally, a low percentage of mapped junction reads—reads that span the junction of two exons—may indicate excessive RNA degradation or DNA contamination. A threshold of ≥ 10 million junction reads per sample has been recommended [1]. However, few studies have specifically evaluated the exact number of mapped reads required for accurate fusion detection. In this study, the mapped reads of a reference standard (s-02) were down-sampled to 100 Mb, 80 Mb, 60 Mb, and 40 Mb, respectively, for ten replicates each (Fig. 2E, F and G). The sensitivity of fusion detection was evaluated for fusions expressed above 40 copies/ng. The results showed that the sensitivity of fusion detection increased as the number of mapped reads increased (Fig. 2G). When the number of mapped reads exceeded 100 Mb, the sensitivity reached 100% (60/60) (Fig. 2E). However, when mapped reads were reduced to 80 Mb, one fusion, *FGFR3::TACC3*, was missed, resulting in a sensitivity of 98.3% (59/60) (Fig. 2F). These results suggest that maintaining mapped reads above 80 Mb is optimal for the WTS assay.

Clinical QC of the WTS assay

To further validate the clinical utility of the WTS assay, we evaluated its clinical QC performance using 156 pan-cancer clinical samples (cohort 3). The RNA yield (median: 100.565 ng/μl and range 5.76–926.2 ng/μl) for all samples met the QC criteria. The main causes of invalid results were RNA degradation and microbiological contamination. The quality of the RNA extracted from this cohort was variable, with DV200 values ranging from 9 to 93% (Fig. 2H). The RNA extraction qualification rate for all samples was 86.5% (135/156) due to 21 samples with a DV200 value below 30%. Besides the highly degraded sample with a DV200 value below 10%, the other 20 samples (DV200 between 10% and 30%) with unqualified extraction were still tested and sequenced. The results showed that 45% (9/20) of these unqualified extraction samples also passed QC, while the QC pass rate for the other 135 qualified samples was 93.3% (126/135) (Fig. 2I). Thus, for the entire cohort (cohort 3), the overall QC pass rate for all sequenced samples was 87% (135/155). The fusion positive rate varied across different cancers (Fig. 2J), with uterine sarcoma having the highest positive rate at 75%, while lung cancer had a 43.8% positive rate. Since some of the 156 samples were confirmed to be fusion-positive before analysis based on the WTS assay, the positive rate of fusion-positive rate of some tumors in this cohort was significantly higher.

Actionable fusions in NSCLC

Targeted therapies for NSCLC patients with actionable fusions have significantly improved clinical outcomes, offering better survival rates compared to standard chemotherapy [44, 45]. Identifying patients who may benefit from targeted therapy is therefore clinically important. In the evaluable NSCLC cohort of 101 patients (from cohort 4), the WTS assay detected fusions in 28.7% of cases ($n = 29$), including 19.8% ($n = 20$) samples with actionable fusions and 8.9% ($n = 9$) samples with other non-actionable fusions (Fig. 3A). The remaining 71.3% (72/101) of samples showed no fusions detected (Fig. 3A). Among the identified fusions, *ALK* fusions accounted for 9.8% ($n = 10$), including 9 actionable *EML4::ALK* fusions and 1 *GPC1::ALK* fusion (Fig. 3B, C and D). *ROS1* fusions and *RET* fusions were found in 3.92% ($n = 4$), and 2.94% ($n = 3$) of the samples, respectively (Fig. 3B and C). Additionally, *NRG1* fusion were identified in 0.98% ($n = 1$) of cases in this study (Fig. 3B and C). Although *NRG1* fusions are rare oncogenic drivers detected in NSCLC, clinical trials such as eNRGy with zenocutuzumab may improve patient survival [46]. In addition to common actionable fusions, the WTS assay can also detect rare fusions such as *MET* EX14 skipping (Fig. 3C), which indicates a better prognosis response to *MET* inhibitors [11, 12].

A total of 101 NSCLC samples from cohort 4 were used to evaluate the prevalence of fusions in NSCLC, with a particular focus on actionable fusions. Figure 3A illustrates the proportion of fusion-negative samples in NSCLC and the prevalence of both actionable and other non-actionable fusions. The specific numbers in parentheses after character "N" represent the number of samples contained in each group. Figure 3B presents the prevalence of fusions involving the genes *ALK*, *ROS1*, *RET*, *BRD4*, *MET*, *NRG1*, and *BRAF*. Figure 3C shows the patient counts and prevalence for all identified fusions in these 101 NSCLC samples. Figure 3D provides a fusion visualization of *EML4::ALK*, the most prevalent fusion in NSCLC, detailing the fusion partners, their orientation, the retained exons in the fusion transcript, and the breakpoint.

Pan-cancer fusions landscape

In addition to evaluating the application of the WTS assay in NSCLC, we further explored its clinical utility in a pan-cancer cohort (cohort 4). This cohort included 13 cancer types with a total of 228 fusion-positive patients, representing 173 unique fusions, of which 35 were actionable (Fig. 4A). Besides the common *ALK* and *ROS1* fusions in lung cancer [9, 47], actionable fusions also occur in other tumors. In particular, actionable fusions involving *ALK* were also detected in uterine cancer and soft tissue sarcoma (Fig. 4A **and** Fig. 4B), although the fusion partners differed among these cancers. *ALK* inhibitors are

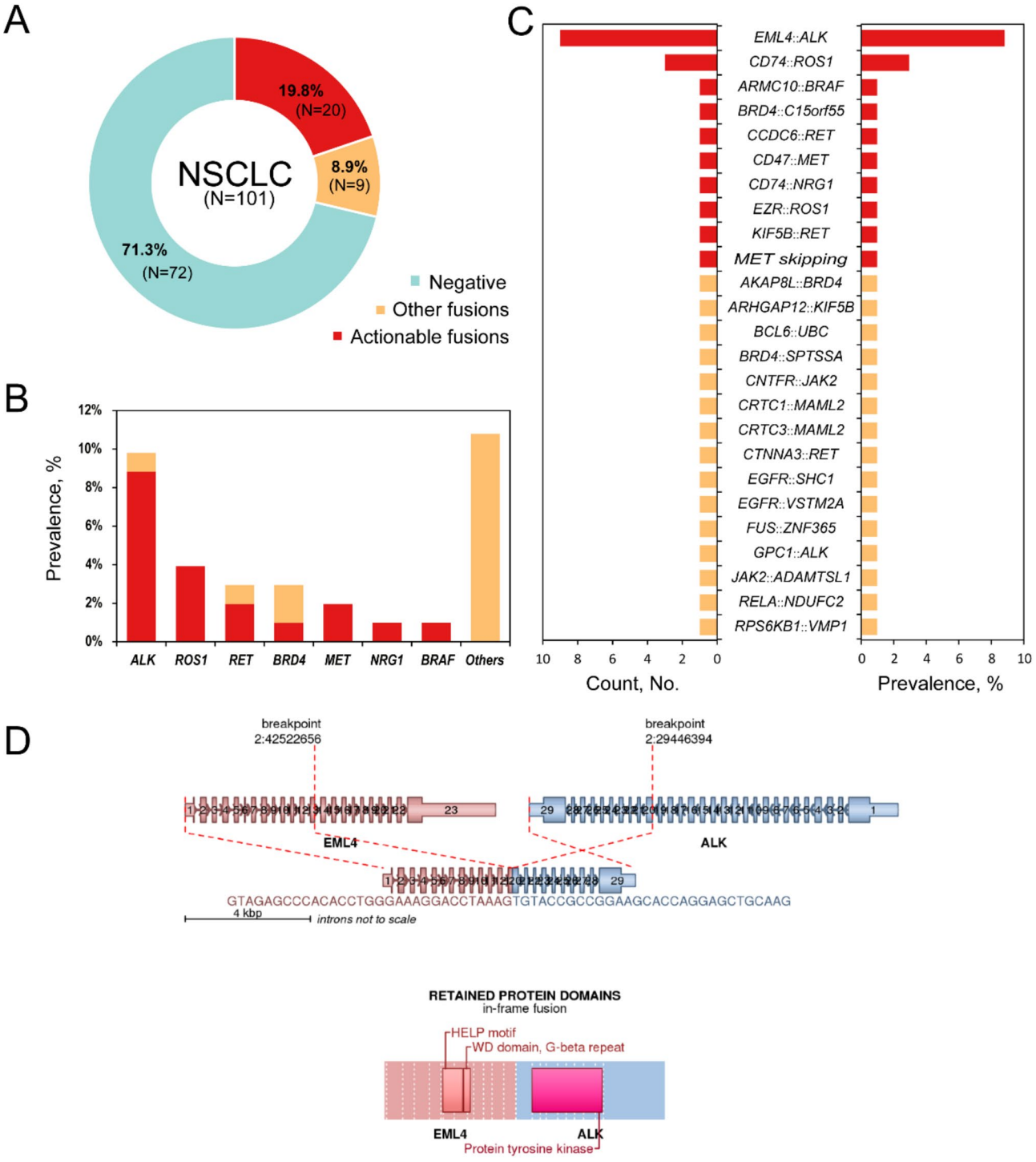

Fig. 3 Prevalence of fusions in NSCLC

effective in *ALK* fusion-positive uterine leiomyosarcoma patients who progressed on chemotherapy [48]. Recent studies have also investigated the use of *ALK* inhibitors in leiomyosarcoma, a type of soft tissue sarcoma, suggesting their potential in treating this disease [49]. Furthermore, though *NTRK* fusions such as *EML4::NTRK3* are rare in colorectal cancer [50] (Fig. 4A), they may serve as potential markers for targeted therapy [51].

In addition to 35 potentially actionable fusions, the remaining 138 other fusions don't currently have therapeutic potential or cannot be translated normally. However, they have other important clinical implications.

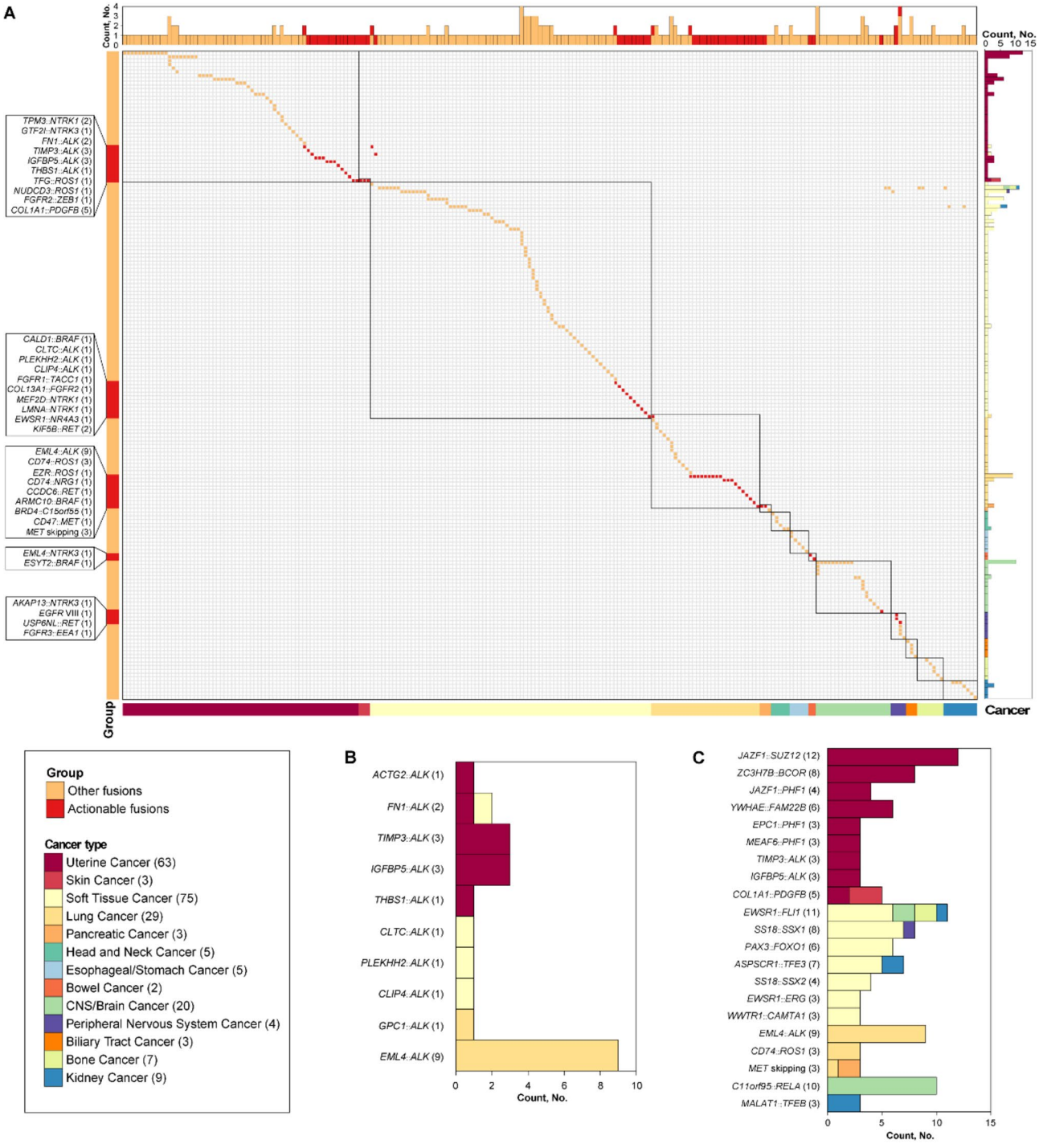

Fig. 4 Fusions landscape of pan-cancer

In particular, we observed significant heterogeneity in fusion gene profiles across different tumor types, highlighting the diagnostic potential of these fusions. The fusion *IGFBP5::ALK* has been reported as a diagnostic marker for uterine inflammatory myofibroblastic tumors [52]. Moreover, fusion genes with diagnostic relevance have been identified in several other cancer types. The *COL1A1::PDGFB* translocation, which is specific to dermatofibrosarcoma protuberans, has been found to play a key role in the diagnosis of this cancer [21]. This fusion is present in over 90% of dermatofibrosarcoma protuberans cases [53]. In our study, we identified *COL1A1::PDGFB* in three skin cancer samples, all of which were diagnosed as dermatofibrosarcoma protuberans (Fig. 4A).

Additionally, two cases of uterine sarcoma also showed this fusion (Fig. 4A). *COL1A1::PDGFB*-positive uterine sarcomas are rare and share certain clinicopathological features with soft tissue dermatofibrosarcoma protuberans [54]. In addition to their diagnostic value, fusions also have potential prognostic value. For instance, the *ZC3H7B::BCOR* fusion identified in uterine sarcoma (Fig. 4C) is associated with poor outcomes and high recurrence rates [55]. Similarly, *PAX3::FOXO1* fusion identified in soft tissue sarcoma (Fig. 4C) has been reported as an important prognostic molecular marker in rhabdomyosarcoma [56].

Figure 4A displays all 173 fusions identified in 228 fusion-positive patients across the pan-cancer cohort (cohort 4). These 173 fusions are categorized into two groups: actionable fusions ($n = 35$, marked in red on the left side of the heatmap) and other non-actionable fusions ($n = 138$, marked in orange). The cohort includes patients from 13 different cancer types, with the number of patients in each cancer type indicated at the bottom. Figure 4B shows all oncogenic fusions involving *ALK* and different fusion partners in this cohort. Figure 4C presents a list of fusions detected in more than three patients in the cohort.

Discussion

We have developed a WTS assay capable of detecting gene fusions, exon skipping events (e.g., *MET* exon 14 skipping), and *EGFR* VIII alterations. The assay demonstrated a sensitivity of 98.4%, and successfully detected 62 out of 63 known fusions identified by other methods, comparable to targeted RNA sequencing approaches [57, 58]. The specificity was 100%, with no false-positive fusions reported. Winters et al. [1] validated RNA-seq for fusion detection across 571 genes, achieving 93% sensitivity and 100% specificity. Another recent study reported an 83.3% sensitivity, detecting 10 out of 12 fusions initially identified by a DNA panel [59]. In contrast, our assay exhibited higher sensitivity and specificity, while also covering a wider range of fusion types, including unique alterations such as *MET* exon 14 skipping and *EGFR* VIII. The repeatability and reproducibility of the assay were excellent, consistently detecting the presence or absence of fusions as well as maintaining consistent supporting read counts and gene coverage. For all tested samples, the repeatability was 97.5% (39/40), due to the low-expression fusion *TPM3::NTRK1* not detected in an inter-run aliquot.

We reviewed the fusion results of 228 fusion-positive clinical pan-cancer samples analyzed using this WTS assay. Our analysis detected 35 potentially actionable fusions, such as *ALK* fusions in lung cancer [47], uterine cancer [48], soft tissue sarcoma [49], and *NTRK* fusions in colorectal cancer [51]. Although most of the fusions detected were not actionable, we found that they have potential diagnostic and prognostic value. For example, *IGFBP5::ALK* has been reported as a diagnostic marker for uterine inflammatory myofibroblastic tumors [52]. The *COL1A1::PDGFB* translocation plays a key role in the diagnosis of dermatofibrosarcoma protuberans [21]. Also, the *ZC3H7B::BCOR* fusion in uterine sarcoma is associated with poor outcomes and high recurrence rates [55]. Similarly, the *PAX3::FOXO1* fusion in soft tissue sarcoma is a key prognostic molecular marker in rhabdomyosarcoma [56]. These results suggest our WTS assay can detect a broader range of fusions, thereby assisting clinicians in the diagnosis, prognosis, and recommendation of targeted therapies for cancer patients.

Identification of fusions using RNA-seq remains challenging, particularly in older archival samples. Prolonged storage at room temperature can lead to degradation, affecting sequencing success and sensitivity of fusion detection [33]. The WTS assay may be less sensitive than targeted RNA sequencing for low expression fusions due to the enrichment capabilities of the latter [25]. Furthermore, substantial bacterial contamination is routinely observed in existing human-derived RNA-seq datasets, highlighting the need for stringent sequencing and analysis protocols [60]. In this study, we found that DV200, which showed a positive correlation with mapping rate, was more suitable for assessing RNA degradation in FFPE samples. We defined DV200 ≥ 30% as the threshold for qualifying RNA degradation. The standard RNA input for the assay should be greater than 100 ng, and fusions expression greater than 40 copies/ng were associated with higher sensitivity for fusion detection. To address contamination, we have incorporated cross-contamination and microbiological analyses into the WTS assay QC process. Additionally, mapped reads are valuable for the detection of non-human contamination, as previously noted [1, 60, 61]. The QC threshold for mapped reads was re-evaluated and showed that ≥ 80 Mb mapped reads were optimal for fusion detection in our WTS assay.

In addition to the aforementioned challenges, this study also has some limitations. First, in cohort 3 and cohort 4, some clinical samples were pre-tested for individual gene fusions during enrollment, leading to higher fusion positivity rates than other reports, which may not reflect the true fusion rates of the cancers. Second, for the false-negative result (p-31, an *EML4::ALK* fusion), sample deficiency prevents further experimental validation of the fusion's authenticity.

In summary, we have developed a new WTS assay that can detect gene fusions in 553 genes, providing valuable insights for the treatment of pan-cancers. The assay demonstrates high sensitivity, specificity, repeatability, and reproducibility, and can identify potentially actionable gene fusions that may be missed by other methods, such

as FISH, IHC, or DNA sequencing. Based on the WTS assay, we mapped the fusion landscape for pan-cancer patients. In NSCLC, 66.7% of the fusions identified were potentially actionable, compared to 20% in pan-cancer. Beyond potentially actionable fusions, we identified a large number of fusions with diagnostic and prognosis value, which can provide prognostic prediction, and help clinicians diagnose tumors.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-14186-w.

Supplementary Material 1

Acknowledgements
We thank the patients, their families, and caregivers for their contributions to this study.

Author contributions
S.C.Z. and X.X.T. wrote the main manuscript text, prepared all figures, and contributed to perform the bioinformatics analysis. C.C.Z., X.H.D., Y.Z., and J.B. designed the study. X.H.D., L.M., Y.F.G., and Y.T.Y. performed the experiments. Y.Z., J.B., X.F.L., J.X.M., H.Y.S., and X.R.F. contributed to perform the bioinformatics analysis. L.Y., X.F.X., X.Y., and C.C.Z. contributed to the preparation of the manuscript. All authors read and approved the final version.

Funding
This work was sponsored by National Natural Science Foundation of China (grant number 82141101), Shanghai Municipal Commission of Health (grant number 2020CXJQ02) and Shanghai pulmonary hospital (grant number fkcx2303).

Data availability
.The data sets used and/or analyzed in the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
All experiments in this study were conducted in accordance with the World Medical Association's Code of Ethics (Declaration of Helsinki) and were approved by the Ethics Committee of Shanghai Pulmonary Hospital (No. K20-288). The participants provided their written informed consent in this study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 27 December 2024 / Accepted: 18 April 2025
Published online: 08 May 2025

References

1. Winters JL, Davila JI, McDonald AM, et al. Development and verification of an RNA sequencing (RNA-Seq) assay for the detection of gene fusions in tumors. J Mol Diagn. 2018;20(4):495–511. https://doi.org/10.1016/j.jmoldx.2018.03.007.
2. Taniue K, Akimitsu N. Fusion genes and RNAs in Cancer development. Non-coding RNA. 2021;7(1). https://doi.org/10.3390/ncrna7010010.
3. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer. 2007;7(4):233–45. https://doi.org/10.1038/nrc2091.
4. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin. 2022;74(3):229–63. https://doi.org/10.3322/caac.21834.
5. Srivastava S, Mohanty A, Nam A, Singhal S, Salgia R, Chemokines. Emerging role in prognosis, heterogeneity, and therapeutics. Semin Cancer Biol. 2022;86(Pt 2):233–46. https://doi.org/10.1016/j.semcancer.2022.06.010.
6. Cho BC, Chiu C, Massarelli E, et al. Updated efficacy and safety of entrectinib in NTRK fusion-positive non-small cell lung cancer. Lung Cancer. 2024;188:107442. https://doi.org/10.1016/j.lungcan.2023.107442.
7. Damiola F, Alberti L, Mansuet-Lupo A, et al. Usefulness of an RNA extraction-free test for the multiplexed detection of ALK, ROS1, and RET gene fusions in real life FFPE specimens of Non-Small cell lung cancers. Expert Rev Mol Diagn. 2023;23(12):1283–91. https://doi.org/10.1080/14737159.2023.2277367.
8. Wu Y, Dziadziuszko R, Ahn JS, et al. Alectinib in resected ALK-Positive Non-Small-Cell lung Cancer. N Engl J Med. 2024;390(14):1265–76. https://doi.org/10.1056/NEJMoa2310532.
9. Shaw AT, Solomon BJ, Chiari R, et al. Lorlatinib in advanced ROS1-positive non-small-cell lung cancer: a multicentre, open-label, single-arm, phase 1–2 trial. Lancet Oncol. 2019;20(12):1691–701. https://doi.org/10.1016/S1470-2045(19)30655-2.
10. Drilon A, Subbiah V, Gautschi O, et al. Selpercatinib in patients with RET Fusion-Positive Non-Small-Cell lung cancer: updated safety and efficacy from the registrational LIBRETTO-001 phase I/II trial. J Clin Oncol. 2023;41(2):385–94. https://doi.org/10.1200/JCO.22.00393.
11. Reungwetwattana T, Liang Y, Zhu V, Ou SI. The race to target MET exon 14 skipping alterations in non-small cell lung cancer: the why, the how, the who, the unknown, and the inevitable. Lung Cancer. 2017;103:27–37. https://doi.org/10.1016/j.lungcan.2016.11.011.
12. Gong C, Xiong H, Qin K, et al. MET alterations in advanced pulmonary sarcomatoid carcinoma. Front Oncol. 2022;12:1017026. https://doi.org/10.3389/fonc.2022.1017026.
13. Kong-Beltran M, Seshagiri S, Zha J, et al. Somatic mutations lead to an oncogenic deletion of Met in lung cancer. Cancer Res. 2006;66(1):283–9. https://doi.org/10.1158/0008-5472.CAN-05-2749.
14. Maher CA, Kumar-Sinha C, Cao X, et al. Transcriptome sequencing to detect gene fusions in cancer. Nature. 2009;458(7234):97–101. https://doi.org/10.1038/nature07638.
15. Song Z, Lian S, Mak S, et al. Deep RNA sequencing revealed fusion junctional heterogeneity May predict Crizotinib treatment efficacy in ALK-Rearranged NSCLC. J Thorac Oncol. 2022;17(2):264–76. https://doi.org/10.1016/j.jtho.2021.09.016.
16. Parker M, Mohankumar KM, Punchihewa C, et al. C11orf95-RELA fusions drive oncogenic NF-kappaB signalling in ependymoma. Nature. 2014;506(7489):451–5. https://doi.org/10.1038/nature13109.
17. Hubner J, Kool M, Pfister SM, Pajtler KW. Epidemiology, molecular classification and WHO grading of ependymoma. J Neurosurg Sci. 2018;62(1):46–50. https://doi.org/10.23736/S0390-5616.17.04152-2.
18. Louis DN, Perry A, Reifenberger G, et al. The 2016 world health organization classification of tumors of the central nervous system: a summary. Acta Neuropathol. 2016;131(6):803–20. https://doi.org/10.1007/s00401-016-1545-1.
19. Tognon C, Knezevich SR, Huntsman D, et al. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. Cancer Cell. 2002;2(5):367–76. https://doi.org/10.1016/s1535-6108(02)00180-0.
20. Ren C, Liu J, Hornicek FJ, et al. Advances of SS18-SSX fusion gene in synovial sarcoma: emerging novel functions and therapeutic potentials. Biochim Et Biophys Acta Reviews cancer. 2024;1879(6):189215. https://doi.org/10.1016/j.bbcan.2024.189215.
21. Serra-Guillen C, Llombart B, Sanmartin O. Dermatofibrosarcoma protuberans. Actas Dermosifiliogr. 2012;103(9):762–77. https://doi.org/10.1016/j.ad.2011.10.007.
22. Guseva NV, Jaber O, Tanas MR, et al. Anchored multiplex PCR for targeted next-generation sequencing reveals recurrent and novel USP6 fusions and upregulation of USP6 expression in aneurysmal bone cyst. Genes Chromosomes Cancer. 2017;56(4):266–77. https://doi.org/10.1002/gcc.22432.
23. Sun L, McNulty SN, Evenson MJ, et al. Clinical implications of a targeted RNA-Sequencing panel in the detection of gene fusions in solid tumors. J Mol Diagn. 2021;23(12):1749–60. https://doi.org/10.1016/j.jmoldx.2021.08.009.
24. Cui C, Shu W, Li P. Fluorescence in situ hybridization: Cell-Based genetic diagnostic and research applications. Front Cell Dev Biol. 2016;4:89. https://doi.org/10.3389/fcell.2016.00089.

25. Heyer EE, Deveson IW, Wooi D, et al. Diagnosis of fusion genes using targeted RNA sequencing. Nat Commun. 2019;10(1):1388. https://doi.org/10.1038/s41467-019-09374-9.
26. Pan Y, Zhang Y, Ye T, et al. Detection of novel NRG1, EGFR, and MET fusions in lung adenocarcinomas in the Chinese population. J Thorac Oncol. 2019;14(11):2003–8. https://doi.org/10.1016/j.jtho.2019.07.022.
27. Benayed R, Offin M, Mullaney K, et al. High yield of RNA sequencing for targetable kinase fusions in lung adenocarcinomas with no mitogenic driver alteration detected by DNA sequencing and low tumor mutation burden. Clin Cancer Res. 2019;25(15):4712–22. https://doi.org/10.1158/1078-0432.CCR-19-0225.
28. Beaubier N, Bontrager M, Huether R, et al. Integrated genomic profiling expands clinical options for patients with cancer. Nat Biotechnol. 2019;37(11):1351–60. https://doi.org/10.1038/s41587-019-0259-z.
29. Haas BJ, Dobin A, Li B, et al. Accuracy assessment of fusion transcript detection via read-mapping and de Novo fusion transcript assembly-based methods. Genome Biol. 2019;20(1):213. https://doi.org/10.1186/s13059-019-1842-9.
30. Arindrarto W, Borras DM, de Groen RAL, et al. Comprehensive diagnostics of acute myeloid leukemia by whole transcriptome RNA sequencing. Leukemia. 2021;35(1):47–61. https://doi.org/10.1038/s41375-020-0762-8.
31. Liu S, Tsai W, Ding Y, 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. 2016;44(5):e47. https://doi.org/10.1093/nar/gkv1234.
32. Kerbs P, Vosberg S, Krebs S, et al. Fusion gene detection by RNA-sequencing complements diagnostics of acute myeloid leukemia and identifies recurring NRIP1-MIR99AHG rearrangements. Haematologica. 2022;107(1):100–11. https://doi.org/10.3324/haematol.2021.278436.
33. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA quality from FFPE samples. PLoS ONE. 2007;2(12):e1261. https://doi.org/10.1371/journal.pone.0001261.
34. Sheng Q, Vickers K, Zhao S, et al. Multi-perspective quality control of illumina RNA sequencing data analysis. Brief Funct Genomics. 2017;16(4):194–204. https://doi.org/10.1093/bfgp/elw035.
35. Kim P, Zhou X. FusionGDB: fusion gene annotation database. Nucleic Acids Res. 2019;47(D1):D994–1004. https://doi.org/10.1093/nar/gky1067.
36. Lee M, Lee K, Yu N, et al. ChimerDB 3.0: an enhanced database for fusion genes from cancer transcriptome and literature data mining. Nucleic Acids Res. 2017;45(D1):D784–9. https://doi.org/10.1093/nar/gkw1083.
37. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. https://doi.org/10.1093/bioinformatics/bty560.
38. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9(4):357–9. https://doi.org/10.1038/nmeth.1923.
39. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635.
40. Uhrig S, Ellermann J, Walther T, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res. 2021;31(3):448–60. https://doi.org/10.1101/gr.257246.119.
41. DeLuca DS, Levin JZ, Sivachenko A, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics. 2012;28(11):1530–2. https://doi.org/10.1093/bioinformatics/bts196.
42. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15(3):R46. https://doi.org/10.1186/gb-2014-15-3-r46.
43. Dobin A, Gingeras TR. Mapping RNA-seq reads with STAR. Curr Protoc Bioinf. 2015;51. 11.14.1-11.14.19.
44. Owen D, Ben-Shachar R, Feliciano J, et al. Actionable structural variant detection via RNA-NGS and DNA-NGS in patients with advanced Non-Small cell lung Cancer. JAMA Netw Open. 2024;7(11):e2442970. https://doi.org/10.1001/jamanetworkopen.2024.42970.
45. Thai AA, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung cancer. Lancet. 2021;398(10299):535–54. https://doi.org/10.1016/S0140-6736(21)00312-3.
46. Kim D, Schram AM, Hollebecque A, et al. The phase I/II eNRGy trial: zenocutuzumab in patients with cancers harboring*NRG1* gene fusions. Future Oncol. 2024;20(16):1057–67. https://doi.org/10.2217/fon-2023-0824.
47. Zhang SS, Nagasaka M, Zhu VW, Ou SI. Going beneath the tip of the iceberg. Identifying and Understanding EML4-ALK variants and TP53 mutations to optimize treatment of ALK fusion positive (ALK+) NSCLC. Lung Cancer. 2021;158:126–36. https://doi.org/10.1016/j.lungcan.2021.06.012.
48. Testa S, Million L, Longacre T, Bui N. Uterine leiomyosarcoma with FN1-Anaplastic lymphoma kinase fusion responsive to alectinib and lorlatinib. Case Rep Oncol. 2021;14(2):812–9. https://doi.org/10.1159/000516758.
49. Davis LE, Nusser KD, Przybyl J, et al. Discovery and characterization of recurrent, targetable ALK fusions in leiomyosarcoma. Mol cancer Research: MCR. 2019;17(3):676–85. https://doi.org/10.1158/1541-7786.MCR-18-1075.
50. Okamura R, Boichard A, Kato S et al. Analysis of NTRK alterations in Pan-Cancer adult and pediatric malignancies: implications for NTRK-Targeted therapeutics. JCO precision oncology. 2018;2018:PO.18.00183. https://doi.org/10.1200/PO.18.00183
51. Pagani F, Randon G, Guarini V, et al. The landscape of actionable gene fusions in colorectal Cancer. Int J Mol Sci. 2019;20(21). https://doi.org/10.3390/ijms20215319.
52. Mohammad N, Haimes JD, Mishkin S, et al. ALK is a specific diagnostic marker for inflammatory myofibroblastic tumor of the uterus. Am J Surg Pathol. 2018;42(10):1353–9. https://doi.org/10.1097/PAS.0000000000001120.
53. Hao X, Billings SD, Wu F, et al. Dermatofibrosarcoma protuberans: update on the diagnosis and treatment. J Clin Med. 2020;9(6). https://doi.org/10.3390/jcm9061752.
54. Lu Y, Chen X, Zeng W, et al. COL1A1::PDGFB fusion uterine sarcoma with a TERT promoter mutation. Genes Chromosomes Cancer. 2024;63(1):e23210. https://doi.org/10.1002/gcc.23210.
55. Fayoumi A. BCOR abnormalities in endometrial stromal sarcoma. Gynecologic Oncol Rep. 2025;57:101672. https://doi.org/10.1016/j.gore.2024.101672.
56. Missiaglia E, Williamson D, Chisholm J, et al. PAX3/FOXO1 fusion gene status is the key prognostic molecular marker in rhabdomyosarcoma and significantly improves current risk stratification. J Clin Oncology: Official J Am Soc Clin Oncol. 2012;30(14):1670–7. https://doi.org/10.1200/JCO.2011.38.5591.
57. Chang F, Lin F, Cao K, et al. Development and clinical validation of a large fusion gene panel for pediatric cancers. J Mol Diagn. 2019;21(5):873–83. https://doi.org/10.1016/j.jmoldx.2019.05.006.
58. Garcia R, Patel N, Uddin N, Park JY. Development and clinical validation of a multiplex gene fusion assay. Lab Med. 2020;51(5):512–8. https://doi.org/10.1093/labmed/lmz102.
59. Peng H, Huang R, Wang K, et al. Development and validation of an RNA sequencing assay for gene fusion detection in Formalin-Fixed, Paraffin-Embedded tumors. J Mol Diagn. 2021;23(2):223–33. https://doi.org/10.1016/j.jmoldx.2020.11.005.
60. Strong MJ, Xu G, Morici L, et al. Microbial contamination in next generation sequencing: implications for sequence-based analysis of clinical samples. PLoS Pathog. 2014;10(11):e1004437. https://doi.org/10.1371/journal.ppat.1004437.
61. Chew SK, Lu D, Campos LS, et al. Polygenic in vivo validation of cancer mutations using transposons. Genome Biol. 2014;15(9):455. https://doi.org/10.1186/s13059-014-0455-6

Publisher's note