npj | precision oncology

Article

Published in partnership with The Hormel Institute, University of Minnesota

https://doi.org/10.1038/s41698-026-01397-y

Comparison of gene fusion detection algorithms reveals frequently overlooked driver fusions in hematologic malignancies

Zen Tamura[1,2,10], Yuki Saito[1,3,10] ✉, Yasunori Kogure[1], Yuji Oshikawa-Kumade[1,4], Suguru Fukuhara[5], Sumito Shingaki[1], Hirokazu Kariyazono[1,4], Yoshiya Kikukawa[1,4], Yuta Ito[1,6], Kota Mizuno[1,2], Yuichi Shiraishi[7], Kotoe Katayama[8], Seiya Imoto[8,9], Koji Izutsu[5], Koichi Murakami[1,2], Junji Koya[1,2] & Keisuke Kataoka[1,2] ✉

Accurate detection of driver gene fusions is essential for the diagnosis, treatment, and prognostic prediction of hematologic malignancies. Despite increasing reliance on RNA sequencing (RNA-seq) for fusion detection, its algorithms have not been sufficiently examined for their ability to detect clinically relevant driver fusions. We evaluated 12 algorithms using conventional RNA-seq from 170 cell lines and targeted RNA-seq from 26 cell lines and 165 clinical samples. The true positive rate, based on 61 and 24 driver fusion–cell line pairs for conventional and targeted RNA-seq, varied between 0.41–1 (median 0.81) and 0–1 (median 0.85), respectively. Many algorithms failed to detect fusions resulting from small deletions (including *STIL*::*TAL1* and *FIP1L1*::*PDGFRA*), lowly expressed fusions, and IGH fusions (*DUX4*::IGH and IGH::*NSD2*). Targeted RNA-seq more sensitively detected driver fusions than conventional RNA-seq, especially lowly expressed ones. One algorithm, Arriba, detected all driver fusions. These findings will inform algorithm selection in clinical settings.

Gene fusions arise from various types of structural variations or chromosomal rearrangements, including translocations, deletions, inversions, and tandem duplications, which juxtapose two different genes. Among them, pathogenic "driver" gene fusions are crucial for the development of various hematologic malignancies. Accumulating evidence suggests that driver gene fusions, such as *BCR*::*ABL1* and *KMT2A* fusions, are highly oncogenic, with a single fusion event capable of driving malignancy[1,2]. Clinically, driver gene fusions can be of diagnostic, prognostic, and therapeutic significance. In the World Health Organization (WHO) classification of haematolymphoid tumours, a variety of gene fusions are listed as defining genetic abnormalities, which are essential for not only diagnosis but also subtype classification[3]. Such gene fusions include *PML*::*RARA* and *RUNX1*::*RUNX1T1* in acute myeloid leukemia (AML), *BCR*::*ABL1* and *ETV6*::*RUNX1* in B-lymphoblastic leukemia/lymphomas (B-ALL/LBL), and *PDGFRA* and *PDGFRB* rearrangements (such as *FIP1L1*::*PDGFRA*) in myeloid/lymphoid neoplasms with eosinophilia[3]. These subtype-defining gene fusions also provide valuable information for predicting prognosis and determining the indication for allogenic hematopoietic stem cell transplantation[4,5]. More importantly, several driver gene fusions are direct therapeutic targets, including *BCR*::*ABL1* and other kinase fusions (targeted by tyrosine kinase inhibitors) and *KMT2A* (*MLL*) rearrangements (by menin inhibitors)[6–8].

Traditionally, metaphase karyotyping and/or fluorescence in situ hybridization (FISH) have been used for detecting chromosomal

[1]Division of Molecular Oncology, National Cancer Center Research Institute, Tokyo, Japan. [2]Division of Hematology, Department of Medicine, Keio University School of Medicine, Tokyo, Japan. [3]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan. [4]Diagnostic Division, Otsuka Pharmaceutical Co., Ltd., Tokyo, Japan. [5]Department of Hematology, National Cancer Center Hospital, Tokyo, Japan. [6]Division of Clinical Oncology and Hematology, Department of Internal Medicine, The Jikei University School of Medicine, Tokyo, Japan. [7]Division of Genome Analysis Platform Development, National Cancer Center Research Institute, Tokyo, Japan. [8]Laboratory of Sequence Analysis, Human Genome Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [9]Division of Health Medical Intelligence, Human Genome Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [10]These authors contributed equally: Zen Tamura, Yuki Saito. ✉e-mail: yuki.saito@keio.jp; kkataoka-tky@umin.ac.jp

THE HORMEL INSTITUTE
UNIVERSITY OF MINNESOTA

rearrangements, while reverse transcriptase polymerase chain reaction (RT-PCR) has been employed to identify fusion transcripts in clinical settings. As the catalog of driver gene fusions has expanded[9–12], these methods have been increasingly supplemented by RNA-seq. Due to the relatively low resolution and/or limited number of target genes of traditional methods, certain gene fusions have remained difficult to detect. These include cryptic gene fusions generated by small intrachromosomal deletions, such as *STIL*::*TAL1* found in T-ALL/LBL[9,10], and promiscuous driver gene fusions[11], in which driver genes, such as *KMT2A* and *ABL1*, form fusions with multiple partner genes. On the other hand, conventional RNA-seq enables whole-transcriptome analysis, allowing systematic discovery of gene fusions. Furthermore, targeted RNA-seq, which enriches specific mRNA sequences of interest using hybridization or PCR amplification, provides a more cost-effective approach for known fusions[13–15]. To search for evidence of gene fusions from RNA-seq data, numerous computer algorithms have been developed. Most of these algorithms are alignment-based, which first align paired-end RNA-seq reads to a reference genome or transcriptome and identify junction reads and/or discordant read pairs. The remaining algorithms are assembly-based, which first assemble RNA-seq reads into a longer contiguous transcript sequence and then map this to the genome for detecting gene fusions (Supplementary Table 1).

Previous studies benchmarked these algorithms using many kinds of gene fusions, demonstrating that several algorithms (such as STAR-Fusion and Arriba) are the most accurate for fusion detection in cancer transcriptomes[16–20]. However, there are many "passenger" gene fusions in cancer. Among the gene fusions detected by RNA-seq, only 3% have been confirmed as recurrent[12,21], suggesting positive selection, consistent with a driver role. Even in hematologic malignancies, approximately 20% are recurrent[12]. The remaining gene fusions are passenger gene fusions, which are generated as by-products of the genomic instability characteristic of malignancies, highlighting the necessity to distinguish driver from passenger gene fusions. However, the ability of algorithms to detect driver gene fusions has not been fully assessed. Furthermore, previous assessments relied on conventional RNA-seq from cell lines or synthetic data, leaving the utility of targeted RNA-seq and the performance of these algorithms on clinical samples unexplored.

To address these issues, we evaluated the accuracy of 12 algorithms to detect driver gene fusions using conventional and targeted RNA-seq data. We show that there are differences in the overall sensitivity of each algorithm and demonstrate that restricting our analysis to driver gene fusions yields different results. Particularly, we identify certain classes of driver gene fusions missed by multiple algorithms, including intrachromosomal gene fusions, gene fusions with low expression, and IGH rearrangements.

Results

Comparative evaluation of detection algorithms for passenger and driver gene fusions using conventional RNA-seq data

We assessed the performance of 12 gene fusion detection algorithms, including 9 alignment-based algorithms (Arriba[22], EricScript[23], FusionCatcher[24], Genomon[25], InFusion[26], JAFFA-direct[27], Pizzly[28], STARChip[29], and STAR-Fusion[19]) and 3 assembly-based algorithms (TrinityFusion-C[19], TrinityFusion-D[19], and TrinityFusion-UC[19]) (Fig. 1). First, we used conventional RNA-seq data of 170 cell lines from hematologic malignancies (Supplementary Table 2). The number of uniquely mapped reads ranged from 4,871,256 to 155,023,272 (median = 71,443,037.5) (Supplementary Fig. 1). These cell lines encompassed a variety of hematologic malignancies, including chronic myeloid leukemia (CML; $n = 14$), AML ($n = 37$), B-ALL/LBL ($n = 15$), T-ALL/LBL ($n = 16$), diffuse large B-cell lymphoma (DLBCL; $n = 17$), Burkitt lymphoma (BL; $n = 10$), multiple myeloma (MM; $n = 30$), and Hodgkin lymphoma (HL; $n = 6$) (Supplementary Fig. 2a). In the conventional RNA-seq data, the median number of gene fusions per cell line detected by each algorithm varied between 2 and 3552, with Pizzly having the highest (median = 3552), followed by JAFFA-direct (median = 1908) and EricScript (median = 925) (Fig. 2a). A total of

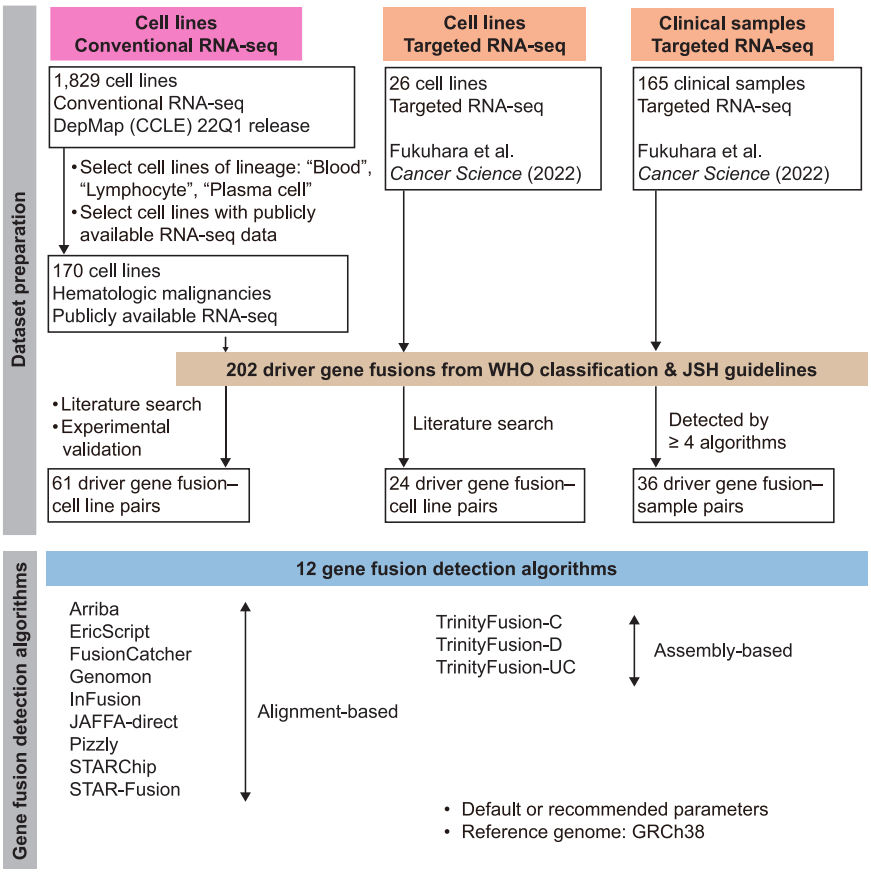


Fig. 1 | Overview of comparative analysis of gene fusion detection algorithms for hematologic malignancies. Dataset preparation steps for conventional RNA-seq data of cell lines, targeted RNA-seq data of cell lines, and targeted RNA-seq data of clinical samples, along with the selection of driver gene fusions in these datasets (top). Evaluation of 12 gene fusion detection algorithms using these datasets (bottom). Cancer Cell Line Encyclopedia (CCLE), Cancer Dependency Map Project DepMap, JSH Japanese Society of Hematology.

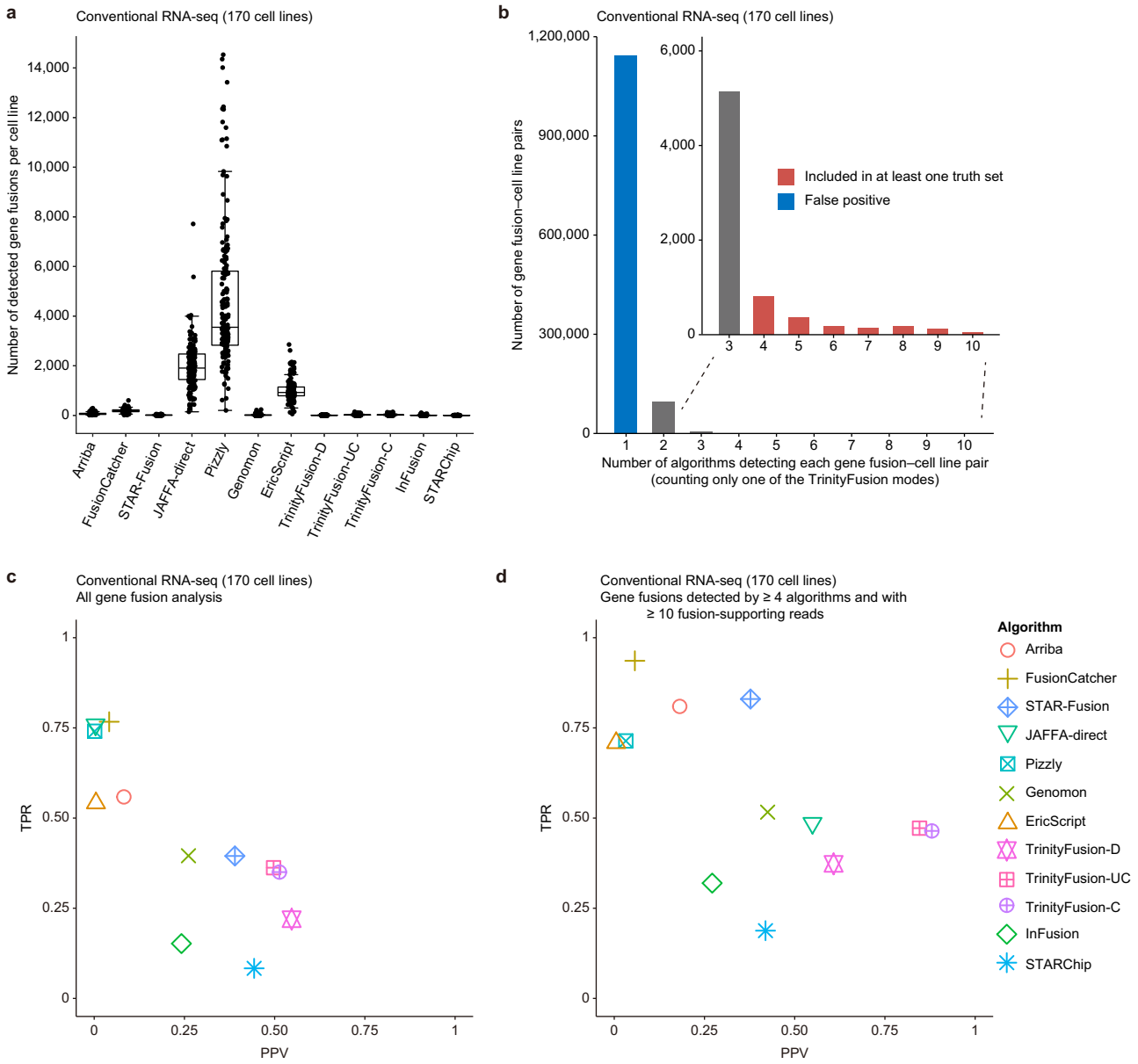


Fig. 2 | Assessment of 12 algorithms for detecting all gene fusions using conventional RNA-seq from cell lines. a Number of detected gene fusions per cell line (dots) for each algorithm in conventional RNA-seq data of 170 CCLE cell lines. Box plots show medians (lines), interquartile ranges (IQRs; boxes), and ± 1.5 × IQR (whiskers). **b** Distribution of the number of gene fusion–cell line pairs according to the number of algorithms detecting them in conventional RNA-seq data of 170 CCLE cell lines. The upper right panel shows an enlarged view of ≥ 3 algorithms. Pairs included in at least one truth set and those classified as false positives are colored red and blue, respectively. **c** PPV and TPR of 12 algorithms using conventional RNA-seq data from 170 CCLE cell lines. The truth set was defined as gene fusion–cell line pairs (including both passenger and driver fusions) detected by ≥ 4 algorithms (excluding the algorithm under evaluation). Different algorithms are shown in different shapes. **d** PPV and TPR of 12 algorithms using conventional RNA-seq data from 170 CCLE cell lines. The truth set was defined as gene fusion–cell line pairs detected by ≥ 4 algorithms (excluding the algorithm under evaluation) with fusion-supporting reads ≥ 10. Different algorithms are shown in different shapes.

1,245,741 gene fusion–cell line pairs were detected by at least one algorithm, with the vast majority (91.7%) detected by only one algorithm (Fig. 2b). Among these, 1840 (0.15%) pairs were detected by four or more algorithms (counting only one of the TrinityFusion modes; see "Methods").

We initially evaluated the ability of each algorithm to detect all gene fusions (both passenger and driver fusions) by calculating the positive predictive value (PPV) and the true positive rate (TPR; equivalent to sensitivity). Given the truth set defined as gene fusion–cell line pairs detected by at least four different algorithms (excluding the algorithm under evaluation), TrinityFusion-C, TrinityFusion-UC, and STAR-Fusion showed the highest F1 scores (harmonic mean of PPV and TPR) (Fig. 2c; Supplementary Table 3), which is almost consistent with previous reports[19,20]. The PPVs of each algorithm were generally low, ranging between 0.001 (Pizzly) and 0.55 (TrinityFusion-D). However, the PPVs increased when a stricter cut-off value of fusion-supporting reads was adopted (Fig. 2d). These findings suggest that most gene fusions are detected by only one algorithm and supported by few fusion-supporting reads, raising the possibility that many represent false positives.

Limiting analysis to driver gene fusions substantially alters the performance of detection algorithms

Given the particular clinical importance of detecting driver gene fusions, we next restricted our analysis to them. Among 1840 gene fusion–cell line pairs detected by four or more algorithms, only 59 (3.2%) involved driver gene fusions included in the list of 202 driver gene fusions (Supplementary Table 4). Among them, 58 driver gene fusion–cell line pairs in 58 cell lines were

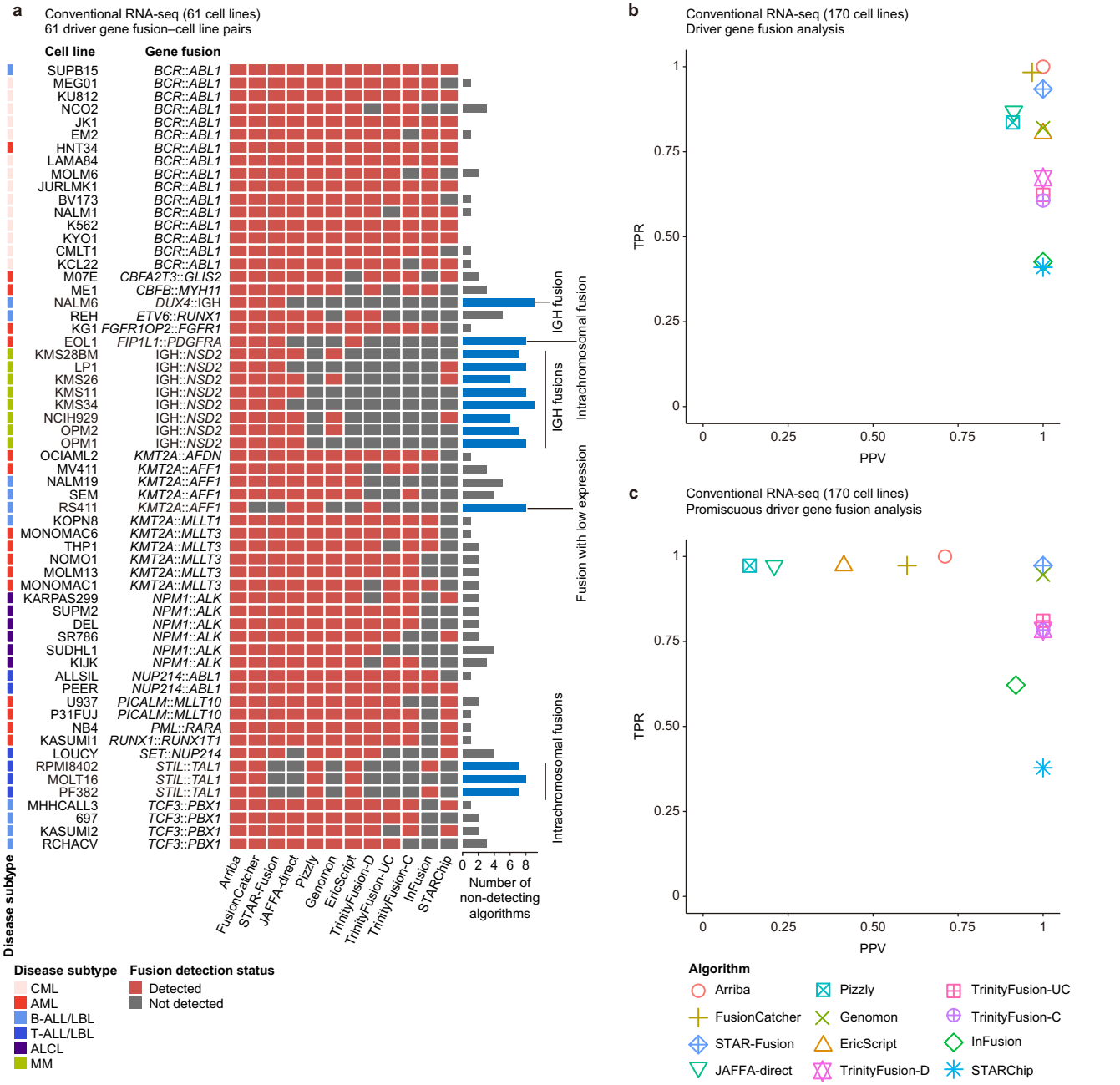


Fig. 3 | Assessment of 12 algorithms for detecting driver gene fusions using conventional RNA-seq from cell lines. a Detection status of 61 driver gene fusion–cell line pairs by 12 algorithms in conventional RNA-seq data of cell lines. Gene fusions, cell lines, and their disease subtypes are shown (left). Number of algorithms that failed to detect each fusion–cell line pair (right). Gene fusion–cell line pairs not detected by more than five algorithms are shown in blue. ALCL anaplastic large cell lymphoma. **b** PPV and TPR of 12 algorithms using conventional RNA-seq data from 170 CCLE cell lines. The truth set was defined as 61 driver gene fusion–cell line pairs. Different algorithms are shown in different shapes. **c** PPV and TPR of 12 algorithms for 37 promiscuous driver gene fusions detected by ≥ 4 algorithms. All promiscuous driver gene fusions involving 11 genes were evaluated.

previously reported in the literature. The remaining pair, *KMT2A*::*AFF1* detected in NALM19, was examined and validated by RT-PCR and Sanger sequencing (Supplementary Fig. 2b). We used 61 driver gene fusion–cell line pairs as a truth set, including these 59 pairs and 2 additional pairs (*DUX4*::IGH in NALM6 and IGH::*NSD2* in KMS34) which were detected by only 3 algorithms but previously reported in the literature (Supplementary Fig. 2c)[30,31].

The 12 algorithms showed a wide range of TPRs, from 0.41 (STARChip) to 1 (Arriba), based on 61 driver fusion–cell line pairs (Fig. 3a, b; Supplementary Table 5). Particularly, Arriba detected all driver gene fusions and reported the highest TPR, followed by FusionCatcher (Fig. 3b). On the other hand, InFusion and STARChip showed the lowest TPRs (less than 0.5). In addition, three assembly-based algorithms (TrinityFusion-C, -D, and -UC) had lower TPRs and thus considered to be less sensitive than alignment-based tools.

The number of false positives varied between 0 and 5 (median = 0), suggesting that restricting analysis to driver gene fusions substantially decreases false positives (Supplementary Fig. 2d), although the possibility that some of them represent true fusion events cannot be entirely excluded. Reflecting this point, the PPVs of most algorithms approached 1, which is a substantial increase compared to both our initial analysis (Fig. 2c) and a previous study[19], which assessed all gene fusions. On the other hand, Pizzly

and JAFFA-direct, which detected the largest number of gene fusions, reported the highest number of false positives and consequently relatively low PPV (Figs. 2a, 3b; Supplementary Fig. 2d).

Overall, Arriba and FusionCatcher showed the highest F1 score, followed by STAR-Fusion and JAFFA-direct (Fig. 3b), while InFusion and STARChip showed the lowest. These observations suggest that limiting the analysis to driver gene fusions yields substantially different performance results compared with analyses including all gene fusions (Fig. 2c). In this context, sensitivity (or TPR) is a more relevant metric compared to PPV mainly because such a procedure effectively excludes false positives. Therefore, the choice of detection algorithm should depend on performance assessments limited to driver gene fusions.

Next, we evaluated the performance of detection algorithms for all promiscuous driver gene fusions involving 11 genes which recurrently form driver gene fusions with multiple partner genes (e.g., *KMT2A*::*GeneX*; see "Methods"), using the conventional RNA-seq data of cell lines. Since considering promiscuous gene fusions did not yield any additional fusions that had been previously reported or were supported by at least four algorithms, the truth set for this analysis consisted of 37 gene fusions, all of which were included in the list of 202 driver gene fusions. When all promiscuous false positives (detected by a single algorithm) were considered, approximately half of the algorithms showed lower PPVs compared to the original results, with three algorithms (Pizzly, JAFFA-direct, and EricScript) reporting PPVs < 0.50 (Fig. 3c), suggesting that they may falsely report promiscuous gene fusions. On the other hand, relaxing the criteria for promiscuous false positives had almost no impact on the remaining algorithms, including STAR-Fusion and Genomon.

Multiple algorithms frequently overlook a specific spectrum of driver gene fusions

The relatively low TPR in multiple algorithms is mainly due to the failed detection of a specific spectrum of driver gene fusions (except for InFusion and STARChip). Among the unique 20 driver gene fusions, 5 driver gene fusions were not detected by at least six algorithms. These included *STIL*::*TAL1* in T-ALL/LBL cell lines (RPMI8402, MOLT16, and PF382)[32–34], *FIP1L1*::*PDGFRA* in an AML cell line (EOL1)[35], *KMT2A*::*AFF1* in a B-ALL/LBL cell line (RS411)[36], *DUX4*::IGH in a B-ALL/LBL cell line (NALM6)[30], and IGH::*NSD2* in MM cell lines (KMS28BM, LP1, KMS26, KMS11, KMS34, NCIH929, OPM2, and OPM1)[31,37–39] (Fig. 3a).

Among them, *STIL*::*TAL1* and *FIP1L1*::*PDGFRA* are generated by interstitial deletions of 90 and 800 kb in length within the same chromosome (at 1p33 and 4q12), respectively[32,40] (Fig. 4a). To assess detection performance for gene fusions arising from small interstitial chromosomal deletions, we generated 11 synthetic RNA-seq data, each containing a fusion transcript arising from an intrachromosomal deletion (<1.5 Mb), which contributes to the pathogenesis and/or is clinically actionable in solid and hematologic malignancies[3,41–47]. Although three algorithms (Arriba, FusionCatcher, and EricScript) successfully detected all 11 synthetic fusion transcripts, many synthetic transcripts, particularly those caused by shorter deletions, were overlooked by multiple algorithms (Fig. 4b). For *STIL*::*TAL1* and *FIP1L1*::*PDGFRA*, the detection status from synthetic RNA-seq data was largely consistent with that from actual RNA-seq data.

KMT2A rearrangements define a subtype of AML, B-ALL/LBL, and mixed-phenotype acute leukemia[3]. Among them, *KMT2A*::*AFF1* (*MLL*::*AF4*) arising from t(4;11)(q21;q23) translocation is the most frequent *KMT2A* fusion and typically associated with worse prognosis in pediatric and adult B-ALL/LBL[48]. Among 11 *KMT2A* fusions including *KMT2A*::*AFF1* in MV411, NALM19, SEM, and RS411[49,50], *KMT2A*::*AFF1* in RS411 had the lowest expression levels for both *KMT2A* (at 5′) and the 3′ partner gene as well as the fewest fusion-supporting reads (Fig. 4c). Using synthetic RNA-seq data, we confirmed that multiple algorithms, particularly assembly-based algorithms, failed to detect *KMT2A*::*AFF1* expressed at low levels (<1.0 TPM) (Fig. 4d). These data suggest that gene fusions with low expression levels can be missed even in cell lines, where tumor purity is generally higher than clinical samples.

DUX4::IGH, which defines a subtype of B-ALL/LBL, and IGH::*NSD2*, which is associated with an adverse prognosis and found in approximately 15% of MM[51,52], were also undetected by multiple algorithms. Although the main pathogenic mechanism of IGH translocations is the enhancer juxtaposition and subsequent aberrant expression of oncogenes without production of fusion transcripts, both fusions recurrently generate functional fusion transcripts (Fig. 4e, f)[30,53]. Among the 12 algorithms, only Arriba, FusionCatcher, and STAR-Fusion consistently detected both *DUX4*::IGH and IGH::*NSD2*, while JAFFA-direct, Genomon, and STARChip detected IGH::*NSD2* in some cell lines (Fig. 3a). As another *DUX4* fusion, *CIC*::*DUX4*, is occasionally found in primitive round cell sarcoma, we analyzed conventional RNA-seq data from three sarcoma samples harboring *CIC*::*DUX4*[54]. Arriba, EricScript, and STAR-Fusion detected *CIC*::*DUX4* fusions in all three samples, while other algorithms failed to detect them (Fig. 4g). These observations suggest that IGH rearrangements are difficult to detect by multiple algorithms, which can be exacerbated if the partner gene is also difficult to detect.

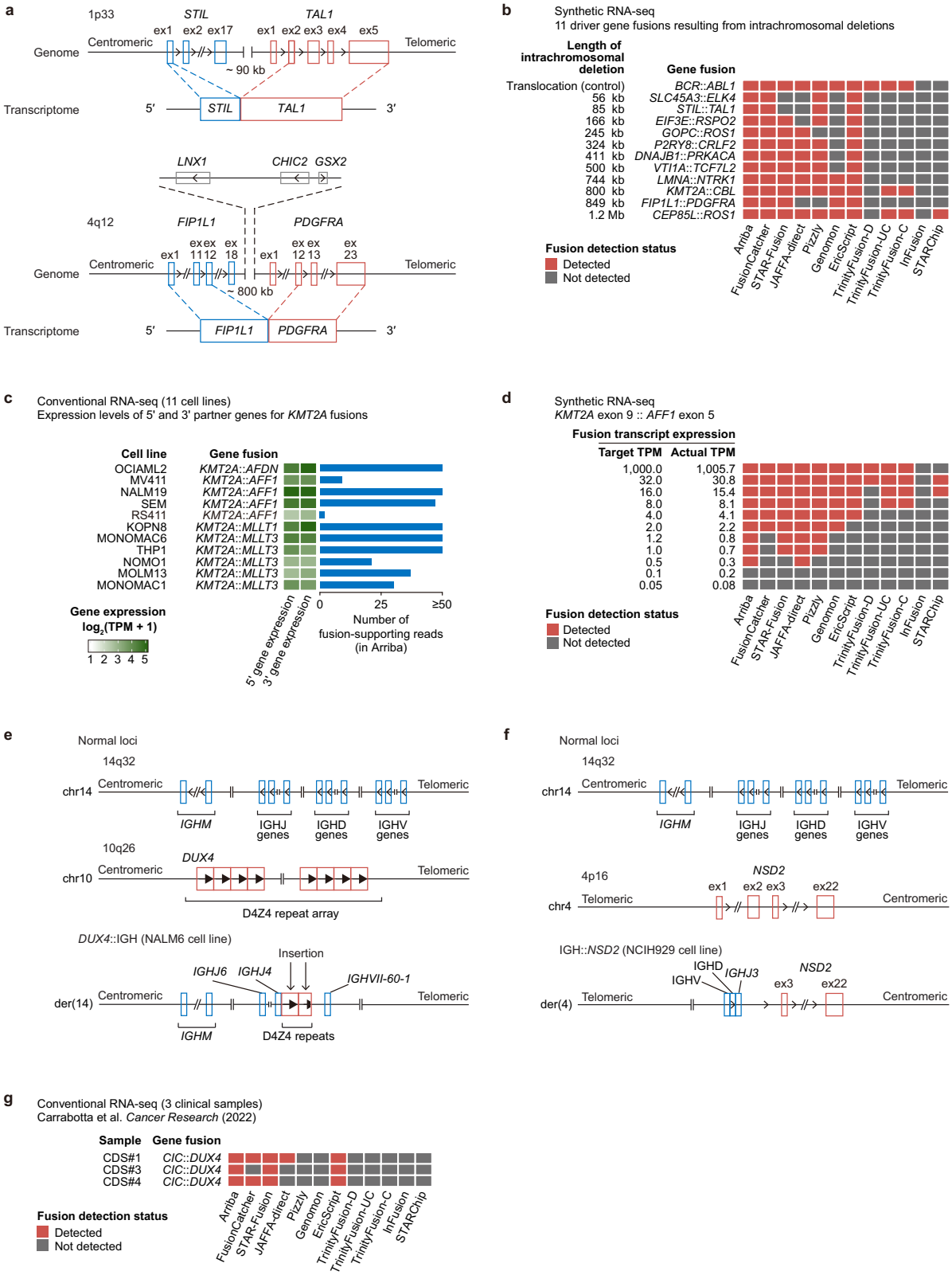
Finally, we assessed whether parameter adjustments can improve the detection of these difficult-to-detect driver fusions for STAR-Fusion, JAFFA-direct, and Genomon, which showed relatively high TPRs and PPVs but failed to detect a specific spectrum of driver gene fusions (Supplementary Fig. 3). *STIL*::*TAL1* in three cell lines was detected after adjusting parameters related to read-through transcript filtering in STAR-Fusion and JAFFA-direct. *KMT2A*::*AFF1* in RS411 was detectable upon relaxation of expression-related stringency in STAR-Fusion. Despite parameter adjustments, the detection status of IGH fusions remained largely unchanged in Genomon and JAFFA-direct.

Driver gene fusion detection using targeted RNA-seq

We next assessed the detection algorithms using targeted RNA-seq data from 26 cell lines, which were obtained from our previous study[15] (Supplementary Table 6). These mostly consisted of CML ($n = 2$), AML ($n = 11$), B-ALL/LBL ($n = 8$), and T-ALL/LBL ($n = 3$) (Supplementary Fig. 4a). The number of uniquely mapped reads ranged from 444,035 to 27,785,035 (median = 8,042,558) (Supplementary Fig. 1). The median number of gene fusions per cell line detected by each algorithm varied between 0 and 211, with JAFFA-direct having the highest (median = 211), followed by Pizzly (median = 121) and FusionCatcher (median = 105) (Supplementary Fig. 4b). A total of 10,755 gene fusion–cell line pairs were detected by at least one algorithm, with the majority (78.6%) detected by only one algorithm and considered to be false positives (Supplementary Fig. 4c). Among these, 267 (2.5%) gene fusions were detected by four or more algorithms (counting only one of the TrinityFusion modes; see "Methods"). When the ability to detect all gene fusions (both passenger and driver fusions) was evaluated, STAR-Fusion showed the highest F1 score, which is consistent with conventional RNA-seq (Supplementary Fig. 4d; Supplementary Table 3).

We next focused on driver gene fusions that were included in the list of 202 driver gene fusions (Supplementary Table 4). The median number of driver gene fusions per cell line detected by each algorithm ranged from 0 to 1 (Supplementary Fig. 4e). A total of 36 driver gene fusion–cell line pairs were detected by at least one algorithm, of which 23 pairs were detected by four or more algorithms (Supplementary Fig. 4f). Among them, 22 pairs were previously reported in the literature and thus included in the truth set. The remaining pair (*ETV6*::*RUNX1* in FKH1) was not previously reported and found to be negative by RT-PCR (Supplementary Fig. 4g, h); therefore, it was not included in the truth set. In addition to the above-mentioned 22 pairs, 2 additional pairs (*FIP1L1*::*PDGFRA* in EOL1 and *DUX4*::IGH in NALM6) were included in the truth set, despite being detected by only 3 and 1 algorithms, respectively, in targeted RNA-seq, because they were included in the conventional RNA-seq analysis. Consequently, the truth set consisted of 24 driver gene fusion–cell line pairs.

The TPRs of the 12 algorithms ranged between 0 and 1, based on 24 driver fusion–cell line pairs (Fig. 5a). Most algorithms had similar TPRs and PPVs to those in conventional RNA-seq analysis (Supplementary Fig. 4i). The number of false positives varied between 0 and 3 (median = 0)



(Supplementary Fig. 4j). Among them, Arriba detected all driver gene fusions and reported the highest F1 score. The two least sensitive algorithms in conventional RNA-seq analysis (STARChip and InFusion) also had the lowest TPRs in targeted RNA-seq. Particularly, STARChip did not detect any driver gene fusions in the default parameter setting. Although a small number of driver fusions could be detected with certain parameter settings (Supplementary Fig. 5), these algorithms may not be appropriate for analyzing targeted RNA-seq data. The frequently overlooked gene fusions in the conventional RNA-seq analysis (*FIP1L1*::*PDGFRA*, *STIL*::*TAL1*, and *DUX4*::IGH) were also not detected by multiple algorithms in the targeted RNA-seq analysis, although Arriba successfully detected them (Fig. 5a).

Fig. 4 | Characterization of frequently undetected driver gene fusions. a Genomic (upper) and transcriptomic (lower) structures of *STIL*::*TAL1* and *FIP1L1*::*PDGFRA* gene fusions resulting from small intrachromosomal deletions. Blue and red boxes represent exons and grey boxes represent genes. Diagonal and vertical double lines represent intragenic and intergenic regions, respectively. Arrows indicate gene orientation. ex exon. **b** Detection status of 11 driver gene fusions resulting from small intrachromosomal deletions and *BCR*::*ABL1* (control) by 12 algorithms in synthetic RNA-seq data. Gene fusions and the lengths of their causative intrachromosomal deletions are shown (left). **c** Gene expression levels for 5′ and 3′ partner genes of 11 *KMT2A* fusions (left). Gene expression levels are presented as $\log_2(\text{TPM} + 1)$. The number of fusion-supporting reads reported by Arriba (right). TPM transcripts per million. **d** Detection status of *KMT2A*::*AFF1* fusions with varying levels of expression ($0.05 \leq$ Target TPM ≤ 1000) in synthetic RNA-seq data. Target and actual TPM values are shown. **e**, **f** Genomic structures (top) and resultant fusions (bottom) of *DUX4*::IGH in NALM6 cell line[30,63] (**e**) and IGH::*NSD2* in NCH929 cell line[37] (**f**). Each box represents an exon or a gene. Vertical double lines represent intergenic regions. Arrows indicate gene orientation. **g** Detection status of *CIC*::*DUX4* in conventional RNA-seq data from three clinical samples known to harbor *CIC*::*DUX4*[54]. Sample IDs (used in the original publication) and gene fusions are shown.

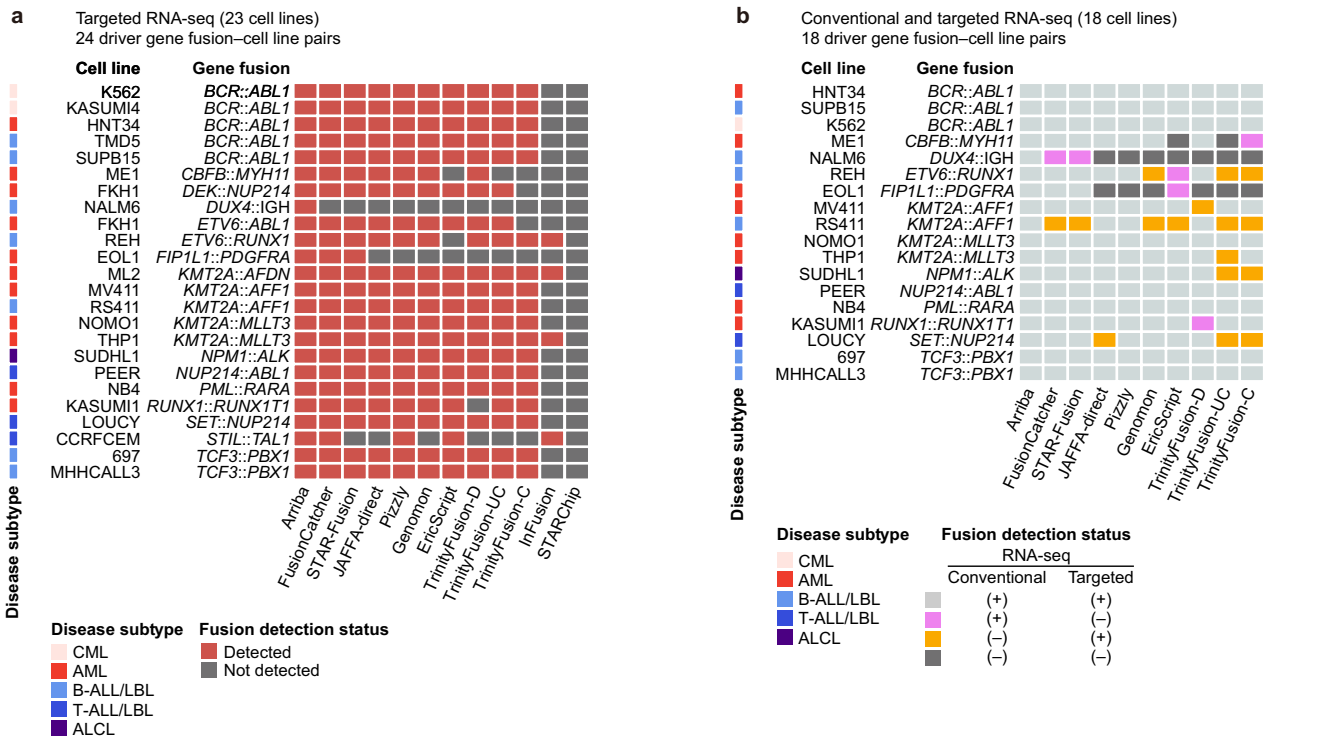


Fig. 5 | Comparison of the ability to detect driver gene fusions between conventional and targeted RNA-seq. a Detection status of 24 driver gene fusion–cell line pairs by 12 algorithms in targeted RNA-seq data from cell lines. Gene fusions, cell lines, and their disease subtypes are shown (left). **b** Comparison of the capacity to detect 18 driver gene fusion–cell line pairs between conventional and targeted RNA-seq across 10 algorithms. The detection status for conventional and targeted RNA-seq are colored in each algorithm. Gene fusions, cell lines, and their disease subtypes are shown (left).

Targeted RNA-seq improves sensitivity to detect driver gene fusions compared to conventional RNA-seq

After excluding InFusion and STARChip, which failed to detect most driver fusions in targeted RNA-seq, we next compared the results of 18 driver gene fusions–cell line pairs for which both conventional and targeted RNA-seq data were available (Fig. 5b). Among 180 combinations (18 driver gene fusion–cell line pairs × 10 algorithms), 16 (8.9%) were detected only in targeted RNA-seq, while 6 (3.3%) were detected only in conventional RNA-seq (Fig. 5b). Notably, *KMT2A*::*AFF1* in RS411, which had the lowest expression among *KMT2A* fusions, was only detected using targeted RNA-seq in six algorithms. In addition, assembly-based algorithms (especially TrinityFusion-UC) detected much more driver gene fusions. Collectively, our results suggest that targeted RNA-seq is an effective and sensitive method for driver gene fusion detection compared to conventional RNA-seq, particularly when analyzing gene fusions with low expression levels.

Targeted RNA-seq shows comparable performance in clinical samples

Finally, to examine the utility of targeted RNA-seq in clinical samples, we executed the algorithms using targeted RNA-seq data from 165 clinical samples consisting of various types of hematologic malignancies, which were obtained from our previous study[15] (Supplementary Fig. 6a; Supplementary Table 7). The number of uniquely mapped reads ranged from 751,207 to 11,737,671 (median = 7,186,721) (Supplementary Fig. 1). Across the dataset, the median number of gene fusions per sample detected by each algorithm varied between 0 and 535 (Supplementary Fig. 6b). A total of 139,505 gene fusion–sample pairs were detected by at least one algorithm, and 1617 (1.2%) were detected by four or more algorithms (Supplementary Fig. 6c). Among them, 36 (2.2%) pairs, consisting of 23 unique gene fusions, were present in the list of 202 driver gene fusions. These included several rare driver gene fusions: *TRAF1*::*ALK*, *XPO1*::*ROS1*, a novel fusion that we first reported[15], and *TRIM24*::*FGFR1*. IGH::*NSD2* was undetected by multiple algorithms, which is consistent with the results from conventional RNA-seq analysis of cell lines. The algorithms with high sensitivity in the analysis of cell lines (Arriba, FusionCatcher, and STAR-Fusion) successfully detected all driver gene fusions in clinical samples (Fig. 6).

Discussion

By evaluating 12 gene fusion detection algorithms using cell lines of hematopoietic origin, we identified three classes of driver gene fusions that are frequently overlooked by multiple algorithms: (1) gene fusions arising from small intrachromosomal deletions, (2) gene fusions expressed at low

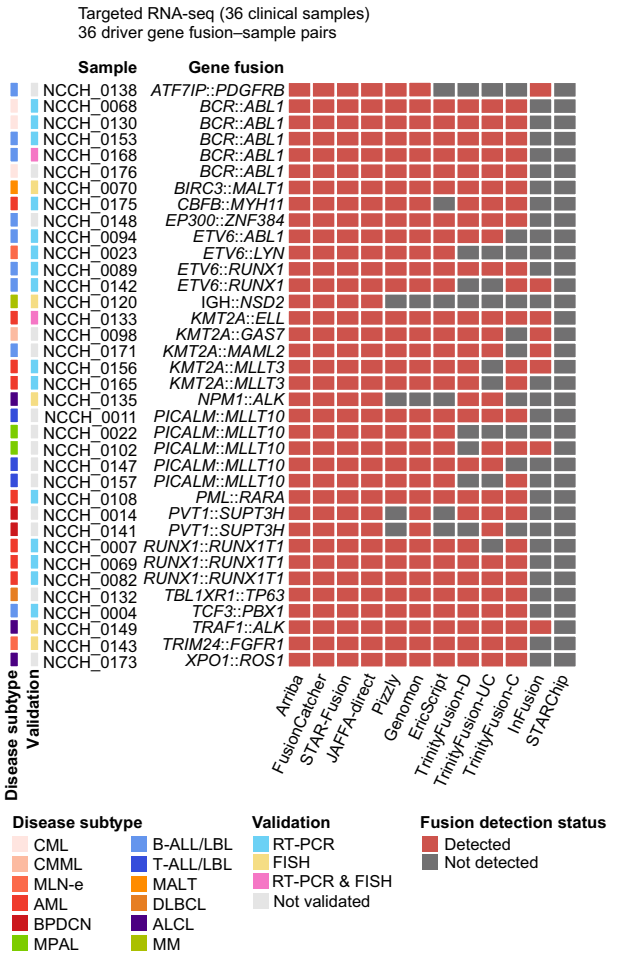


Fig. 6 | Detection of driver fusions using targeted RNA-seq from clinical samples. Detection status of 36 driver gene fusion–sample pairs by 12 algorithms in targeted RNA-seq data from clinical samples. Gene fusions, sample IDs, and their disease subtypes and validation status are shown (left). BPDCN blastic plasmacytoid dendritic cell neoplasm, CMML chronic myelomonocytic leukemia, MALT extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue, MLN-e myeloid/lymphoid neoplasms with eosinophilia and gene rearrangement, MPAL mixed-phenotype acute leukemia.

levels, and (3) IGH rearrangements, although parameter adjustments can improve the detection of these difficult-to-detect fusions. These fusions can play biologically important roles, even when expressed at low levels, as exemplified by *KMT2A*::*AFF1* in RS411[55,56]. More critically, they include clinically relevant ones with diagnostic, therapeutic, and prognostic relevance. Even large-scale genomic studies used algorithms that failed to detect these fusions (e.g., STAR-Fusion and EricScript in the Pan-Cancer Analysis of Whole Genomes[21] and/or The Cancer Genome Atlas[57]). In addition, *STIL*::*TAL1* in RPMI8402, MOLT16, and PF382 cell lines, and *DUX4*::IGH in NALM6 cell line were missed in the 22Q1 version of DepMap. Therefore, choosing the appropriate algorithms is necessary for detecting frequently overlooked driver gene fusions. This is especially relevant for *DUX4*::IGH, because detecting it with conventional cytogenetic approaches (such as FISH) is challenging due to the repetitive nature of the *DUX4* locus[3,58,59]. Additionally, while FISH is currently the preferred method for detecting IGH::*NSD2*, RNA-seq may become clinically necessary to determine the exact location of the *NSD2* breakpoint, because of recent advances detailing prognostic differences between their location[53].

For identifying novel gene fusions and comprehensively characterizing cancer transcriptomes in an exploratory manner, fast and accurate algorithms with high TPR and PPV are suitable. Recent comparative studies ranked Arriba and STAR-Fusion as top performers based on simulated and cell line RNA-seq data, which mostly consisted of passenger gene fusions[19,20]. In addition to these alignment-based algorithms, assembly-based algorithms are also useful for reconstructing fusion isoforms and tumor viruses in cancer research.

On the other hand, our results show that the relative performance of these algorithms differed when restricting to driver gene fusions, which constitute only approximately 3% of gene fusions detected by multiple algorithms. As such a procedure removes the great majority of false positives, the most sensitive algorithms (such as Arriba and FusionCatcher) showed the best performance. In addition, despite the high performance in overall gene fusion detection, several algorithms (such as STAR-Fusion) missed certain classes of driver gene fusions as described above. Moreover, several algorithms tend to falsely report promiscuous gene fusions, as reflected by decreased PPVs when promiscuous driver gene fusions were evaluated. These results suggest that careful consideration is required when interpreting gene fusions involving a well-known promiscuous gene and a rare partner gene. Taken together, in the clinical setting, where detecting known clinically relevant driver fusions is important, the performance assessments focusing on driver gene fusions should guide the choice of algorithms, although further improvement of the framework for benchmarking algorithms for driver gene fusion detection is warranted.

Another important observation of our study is that targeted RNA-seq improves the sensitivity of driver fusion detection for multiple algorithms compared to conventional RNA-seq, especially for gene fusions expressed at low levels. In addition, we confirm that top-performing algorithms in cell lines (such as Arriba) can perform well in clinical samples. These findings complement previous studies evaluating the performance of targeted RNA-seq using a limited number of driver gene fusions in cell lines[60,61].

Collectively, our study demonstrates that an assessment limited to driver gene fusions yields clinically relevant insights which can influence the choice of algorithms. Particularly, certain classes of frequently overlooked gene fusions, including intrachromosomal gene fusions, gene fusions with low expression, and IGH rearrangements, require close attention in clinical settings.

Methods

Conventional RNA-seq

We obtained conventional RNA-seq data from 170 hematologic malignancy cell lines registered in CCLE (Supplementary Table 2). From 1829 cell lines listed in the 22Q1 version of DepMap (https://depmap.org/portal/), we extracted 281 cell lines of hematologic malignancies (cell lineage of "Blood", "Lymphocyte", or "Plasma cell"). Among these 281 cell lines, we used all 170 cell lines for which RNA-seq data were available (Supplementary Table 2). FASTQ files for 11 cell lines were downloaded from the Sequence Read Archive (SRA) (Supplementary Table 8). BAM files for the other cell lines were downloaded from the National Cancer Institute's Genomic Data Commons Legacy Archive[62] and converted to FASTQ format using biobambam's bam2fastq (version 0.0.191). TPM-normalized gene expression data for these cell lines were obtained from DepMap.

To assess whether the presence of multiple copies of *DUX4* gene throughout the human genome affects the performance of fusion detection, conventional RNA-seq data of three sarcoma samples with experimentally validated *CIC*::*DUX4* fusion were additionally downloaded from SRA in FASTQ format (Supplementary Table 8)[54].

Targeted RNA-seq

Targeted RNA-seq data from 26 cell lines (Supplementary Table 6) and 165 clinical samples (Supplementary Table 7) were obtained from our previous study, which was approved by the institutional ethics committee of National Cancer Center (reference number: 2018-366) and was conducted in accordance with the Declaration of Helsinki[15]. All patients provided written informed consent. Duplicated samples from the same patient were removed. Tumor content was assessed based on histopathological review (Supplementary Table 7). The custom hybridization capture-based RNA

profiling assay for targeted RNA-seq has been described previously[15]. The genes and junctions covered by the targeted RNA-seq assay have been reported previously[15]. Briefly, the custom bait library was designed to detect (i) entire coding sequences (CDSs) of 44 genes targeted by recurrent fusions, (ii) CDSs and untranslated regions of 32 genes targeted by recurrent activating structural variations, (iii) immunoglobulin and T cell receptor genes, and (iv) fused exons, their adjacent exons, and junction sequences of 134 well-known fusions[15].

Detection of gene fusions by each algorithm

Twelve algorithms (Arriba[22], EricScript[23], FusionCatcher[24], Genomon[25], InFusion[26], JAFFA-direct[27], Pizzly[28], STARChip[29], STAR-Fusion[19], TrinityFusion-C[19], TrinityFusion-D[19], TrinityFusion-UC[19]) were executed with default or recommended parameters using GRCh38 as the reference genome (Supplementary Tables 1 and 9), unless otherwise specified. If an algorithm outputs a filtered and unfiltered list of gene fusions, only the filtered list was used. As some algorithms use outdated gene names, we cross-referenced gene names with all known alias symbols and previous symbols registered in the HUGO Gene Nomenclature Committee at the European Bioinformatics Institute (https://www.genenames.org/). As many copies of the nearly identical *DUX4* open reading frame are present in the human genome and annotated as *DUX4* pseudogenes, gene fusions involving such pseudogenes are also considered as *DUX4* rearrangements. Gene names from the immunoglobulin heavy chain locus were condensed to IGH. Fusion-supporting reads were calculated for each algorithm as in Supplementary Table 10. If there were multiple entries of the same gene fusion (e.g. because of different breakpoints), the entry with the largest number of supporting reads was chosen. In principle, we assessed not only the combination of genes but also their orientation to accurately identify a gene fusion. Specifically, for *GeneA*::*GeneB* fusion, where *GeneA* is the 5′ fusion partner and *GeneB* is the 3′ fusion partner, we distinguished *GeneA*::*GeneB* (correct orientation) from *GeneB*::*GeneA* (opposite orientation), except for Pizzly, which does not seem to distinguish 5′ and 3′ partners.

Definition of driver gene fusions

We manually curated a list of 202 driver gene fusions implicated in hematologic malignancies (Supplementary Table 4), consisting of almost all gene fusions listed in the 5th edition of the WHO classification[3] and the Japanese Society of Hematology Guideline for Genomic Testing in Hematologic Malignancies (2021 release).

Comparison of detection algorithms

For the analysis of all gene fusions in cell lines, the truth set was defined as gene fusion–cell line pairs detected by at least four different algorithms, excluding the algorithm under evaluation; its size ranged from 1211 to 1823 and 113 to 267 pairs per algorithm for conventional and targeted RNA-seq data, respectively. Within the truth set, orthogonal validation was unavailable except for driver gene fusion–cell line pairs. Based on this truth set, true positives and false negatives were tabulated. False positives were defined as gene fusions detected by one algorithm as previously described[19]. Similar to a previous study[19], TrinityFusion contributed only one vote, regardless of the number of modes (TrinityFusion-C, -D, and -UC) detecting a particular gene fusion. TPR (also known as sensitivity and recall), PPV (also known as precision), and F1 accuracy measure (the harmonic mean of TPR and PPV) were computed according to standards: TPR = true positives / (true positives + false negatives); PPV = true positives / (true positives + false positives); and F1 = 2 × (TPR × PPV) / (TPR + PPV).

For the analysis of driver gene fusions in cell lines, we selected gene fusion–cell line pairs which were (i) detected by four or more algorithms, (ii) included in the list of 202 driver gene fusions, and (iii) reported in the literature or experimentally validated by RT-PCR and Sanger sequencing. For conventional RNA-seq analysis, we added two pairs (*DUX4*::IGH in NALM6 and IGH::*NSD2* in KMS34) detected by 3 algorithms but reported in the literature[30,31]. For targeted RNA-seq analysis, we added two pairs (*DUX4*::IGH in NALM6 and *FIP1L1*::*PDGFRA* in EOL1), included in conventional RNA-seq analysis. In total, the truth set included 61 driver gene fusion–cell line pairs (20 unique gene fusions) for the conventional RNA-seq data ($n = 170$, of which 61 cell lines harbored driver gene fusions) and 24 driver gene fusion–cell line pairs (17 unique gene fusions) for targeted RNA-seq data ($n = 26$, of which 23 cell lines harbored driver gene fusions), all of which were orthogonally validated (Supplementary Tables 2 and 6). False positives were defined as gene fusions included in the driver gene fusion list and detected by one algorithm.

For the analysis of driver gene fusions in clinical samples, we selected gene fusions which were (i) detected by four or more algorithms and (ii) included in the driver gene fusion list. The truth set included 36 gene fusion–sample pairs (23 unique gene fusions), of which 22 pairs were validated using FISH or RT-PCR, as described previously[15].

We also evaluated the effect of (i) the number of fusion-supporting reads and (ii) inclusion of promiscuous driver gene fusions, defined as fusions in which given driver genes form fusions with multiple partner genes, on the accuracy of detection algorithms, using the conventional RNA-seq data from cell lines. For the former analysis, output files generated with default parameters were filtered according to the number of fusion-supporting reads (Supplementary Table 10). For the latter analysis, we focused on gene fusions involving 11 genes which recurrently form driver gene fusions with multiple partner genes: *KMT2A*, *ABL1*, *ABL2*, *ALK*, *FGFR1*, *PDGFRA*, *PDGFRB*, *NTRK1*, *NTRK2*, *NTRK3*, and *JAK2* (as the 3′ partner gene except for *KMT2A*).

Finally, we evaluated how parameter adjustments influence the detection performance of detection algorithms. Specifically, we systematically examined the detection of difficult-to-detect gene fusions (*STIL*::*TAL1*, *FIP1L1*::*PDGFRA*, *KMT2A*::*AFF1* [in RS411], and IGH rearrangements) by evaluating 76, 49, and 96 parameter settings for Genomon, JAFFA-direct, and STAR-Fusion, respectively. In addition, we analyzed gene fusion detection using targeted RNA-seq by evaluating 71 and 94 parameter settings for STARChip and InFusion, respectively. We modified one parameter at a time from its default setting. For parameters with binary or a limited number of discrete values, we evaluated most available parameter settings. For parameters with integer or continuous values, we evaluated parameter settings spanning one to two orders of magnitude above and below the default settings, whenever feasible.

Analysis using synthetic RNA-seq data

We used FusionSimulatorToolkit (https://github.com/FusionSimulatorToolkit/FusionSimulatorToolkit) with several modifications to generate synthetic RNA-seq data. We manually constructed each fusion transcript (Supplementary Table 11), combined them with a reference cDNA FASTA file (GENCODE version 38), and executed the Trinity software package (util/align_and_estimate_abundance.pl) with the pseudorandom seed fixed for both Bowtie and RSEM (--bowtie_RSEM --all --best --strata -m 300 --chunkmbs 512 --seed <seed > --rsem_add_opts --seed <seed >). We then modified the TPM values in RSEM.isoforms.results file to adjust the expression level of fusion transcripts to 1000 TPM unless otherwise specified. Finally, we executed the rsem-simulate-reads command with the pseudorandom seed fixed (--seed <seed >) to generate 30 million paired-end reads per sample using conventional RNA-seq data from the SUPB15 cell line as the model fastq file. The generated data were then executed through 12 algorithms in the above-mentioned manner to further evaluate their performance in detecting frequently overlooked driver gene fusions.

Cell lines

NALM19, REH, and FKH1 cell lines were obtained from the JCRB cell bank, ATCC, and DSMZ, respectively. Cell lines were cultured according to the providers' instructions and routinely checked for mycoplasma infection.

RT-PCR and Sanger sequencing

Total RNA was extracted using RNeasy Mini kit (Qiagen), followed by complementary DNA (cDNA) synthesis using the SuperScript IV First-

strand synthesis system (Invitrogen) with an oligo dT primer. Then, cDNA was amplified by PCR using Q5 Hot Start High-Fidelity DNA polymerase (New England BioLabs). The samples were denatured at 98 °C for 30 s, subjected to 40 cycles of amplification (98 °C for 10 s, 68 °C for 15 s, and 72 °C for 30 s), followed by a final extension step at 72 °C for 2 min. PCR products were resolved by agarose gel electrophoresis. Subsequently, Sanger sequencing was performed for the gel-extracted PCR products (Eurofins Japan). PCR and sequencing primers are listed in Supplementary Table 12.

Data availability

The data that support the findings of this study were obtained either from publicly available sources, as described in the Methods, or from datasets that are not publicly available due to privacy protections but are available from the corresponding authors upon reasonable request.

Code availability

The algorithms and computational tools used in this study are publicly available (refer to Methods and Supplementary Tables for details).

Received: 30 June 2025; Accepted: 16 March 2026;
Published online: 04 April 2026

References

1. Kamizela, A. E. et al. Timing and trajectory of BCR::ABL1-driven chronic myeloid leukaemia. *Nature* **640**, 982–990 (2025).
2. Andersson, A. K. et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat. Genet.* **47**, 330–337 (2015).
3. WHO Classification of Tumours Editorial Board. *Haematolymphoid tumours*. 5th edn, (International Agency for Research on Cancer, 2024).
4. Dohner, H. et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood* **140**, 1345–1377 (2022).
5. Mrozek, K. et al. Outcome prediction by the 2022 European LeukemiaNet genetic-risk classification for adults with acute myeloid leukemia: an Alliance study. *Leukemia* **37**, 788–798 (2023).
6. Issa, G. C. et al. The menin inhibitor revumenib in KMT2A-rearranged or NPM1-mutant leukaemia. *Nature* **615**, 920–924 (2023).
7. Issa, G. C. et al. Menin Inhibition With Revumenib for KMT2A-Rearranged Relapsed or Refractory Acute Leukemia (AUGMENT-101). *J. Clin. Oncol.* **43**, 75–84 (2025).
8. Wang, E. S. et al. Ziftomenib in relapsed or refractory acute myeloid leukaemia (KOMET-001): a multicentre, open-label, multi-cohort, phase 1 trial. *Lancet Oncol.* **25**, 1310–1324 (2024).
9. Aifantis, I., Raetz, E. & Buonamici, S. Molecular pathogenesis of T-cell leukaemia and lymphoma. *Nat. Rev. Immunol.* **8**, 380–390 (2008).
10. Brown, L. et al. Site-specific recombination of the tal-1 gene is a common occurrence in human T cell leukemia. *EMBO J.* **9**, 3343–3351 (1990).
11. Horie, S. et al. Pan-cancer comparative and integrative analyses of driver alterations using japanese and international genomic databases. *Cancer Discov.* **14**, 786–803 (2024).
12. Mertens, F., Johansson, B., Fioretos, T. & Mitelman, F. The emerging complexity of gene fusions in cancer. *Nat. Rev. Cancer* **15**, 371–381 (2015).
13. He, J. et al. Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting. *Blood* **127**, 3004–3014 (2016).
14. Kohsaka, S. et al. Comprehensive assay for the molecular profiling of cancer by target enrichment from formalin-fixed paraffin-embedded specimens. *Cancer Sci.* **110**, 1464–1479 (2019).
15. Fukuhara, S. et al. Feasibility and clinical utility of comprehensive genomic profiling of hematological malignancies. *Cancer Sci.* **113**, 2763–2777 (2022).
16. Carrara, M. et al. State of art fusion-finder algorithms are suitable to detect transcription-induced chimeras in normal tissues?. *BMC Bioinform.* **14**, S2 (2013).
17. Kumar, S., Vo, A. D., Qin, F. & Li, H. Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci. Rep.* **6**, 21597 (2016).
18. Liu, S. et al. Comprehensive evaluation of fusion transcript detection algorithms and a meta-caller to combine top performing methods in paired-end RNA-seq data. *Nucleic Acids Res.* **44**, e47 (2016).
19. Haas, B. J. et al. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* **20**, 213 (2019).
20. Creason, A. et al. A community challenge to evaluate RNA-seq, fusion detection, and isoform quantification methods for cancer discovery. *Cell Syst.* **12**, 827–838. e825 (2021).
21. PCAWG Transcriptome Core Group et al. Genomic basis for RNA alterations in cancer. *Nature* **578**, 129–136 (2020).
22. Uhrig, S. et al. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res.* **31**, 448–460 (2021).
23. Benelli, M. et al. Discovering chimeric transcripts in paired-end RNA-seq data by using EricScript. *Bioinformatics* **28**, 3232–3239 (2012).
24. Nicorici, D. et al. FusionCatcher—a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*, 011650, https://doi.org/10.1101/011650 (2014).
25. Shiraishi, Y. et al. Integrated analysis of whole genome and transcriptome sequencing reveals diverse transcriptomic aberrations driven by somatic genomic changes in liver cancers. *PLoS One* **9**, e114263 (2014).
26. Okonechnikov, K. et al. InFusion: advancing discovery of fusion genes and chimeric transcripts from deep RNA-sequencing data. *PLoS One* **11**, e0167417 (2016).
27. Davidson, N. M., Majewski, I. J. & Oshlack, A. JAFFA: high sensitivity transcriptome-focused fusion gene detection. *Genome Med* **7**, 43 (2015).
28. Melsted, P. et al. Fusion detection and quantification by pseudoalignment. *bioRxiv*, 166322, https://doi.org/10.1101/166322 (2017).
29. Akers, N. K., Schadt, E. E. & Losic, B. STAR Chimeric Post for rapid detection of circular RNA and fusion transcripts. *Bioinformatics* **34**, 2364–2370 (2018).
30. Yasuda, T. et al. Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. *Nat. Genet.* **48**, 569–574 (2016).
31. Intini, D. et al. Identification of a novel IGH-MMSET fusion transcript in a human myeloma cell line with the t(4;14)(p16.3;q32) chromosomal translocation. *Br. J. Haematol.* **126**, 437–439 (2004).
32. Aplan, P. D. et al. Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity. *Science* **250**, 1426–1429 (1990).
33. Nagel, S. et al. Multiple mechanisms induce ectopic expression of LYL1 in subsets of T-ALL cell lines. *Leuk. Res* **34**, 521–528 (2010).
34. Kjeldsen, E. et al. Characterization of camptothecin-induced genomic changes in the camptothecin-resistant T-ALL-derived cell line CPT-K5. *Cancer Genomics Proteom.* **15**, 91–114 (2018).
35. Cools, J. et al. The EOL-1 cell line as an in vitro model for the study of FIP1L1-PDGFRA-positive chronic eosinophilic leukemia. *Blood* **103**, 2802–2805 (2004).
36. Stong, R. C., Korsmeyer, S. J., Parkin, J. L., Arthur, D. C. & Kersey, J. H. Human acute leukemia cell line with the t(4; 11) chromosomal rearrangement exhibits B lineage and monocytic characteristics. *Blood* **65**, 21–31 (1985).
37. Chesi, M. et al. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood* **92**, 3025–3034 (1998).
38. Malgeri, U. et al. Detection of t(4;14)(p16.3;q32) chromosomal translocation in multiple myeloma by reverse transcription-polymerase chain reaction analysis of IGH-MMSET fusion transcripts. *Cancer Res.* **60**, 4058–4061 (2000).

39. Kuipers, J. et al. Fluorescence in situ hybridization analysis shows the frequent occurrence of 14q32.3 rearrangements with involvement of immunoglobulin switch regions in myeloma cell lines. *Cancer Genet. Cytogenet.* **109**, 99–107 (1999).
40. Gotlib, J. & Cools, J. Five years since the discovery of FIP1L1-PDGFRA: what we have learned about the fusion and other molecularly defined eosinophilias. *Leukemia* **22**, 1999–2010 (2008).
41. Honeyman, J. N. et al. Detection of a recurrent DNAJB1-PRKACA chimeric transcript in fibrolamellar hepatocellular carcinoma. *Science* **343**, 1010–1014 (2014).
42. Kastenhuber, E. R. et al. DNAJB1-PRKACA fusion kinase interacts with beta-catenin and the liver regenerative response to drive fibrolamellar hepatocellular carcinoma. *Proc. Natl. Acad. Sci. USA* **114**, 13076–13084 (2017).
43. Zhang, Y. et al. Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. *Cancer Discov.* **2**, 598–607 (2012).
44. Seshagiri, S. et al. Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660–664 (2012).
45. Bass, A. J. et al. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nat. Genet.* **43**, 964–968 (2011).
46. Davare, M. A. et al. Rare but recurrent ROS1 fusions resulting from chromosome 6q22 microdeletions are targetable oncogenes in glioma. *Clin. Cancer Res.* **24**, 6471–6482 (2018).
47. Wiesner, T. et al. Kinase fusions are frequent in Spitz tumours and Spitzoid melanomas. *Nat. Commun.* **5**, 3116 (2014).
48. Pieters, R. et al. Outcome of infants younger than 1 year with acute lymphoblastic leukemia treated with the interfant-06 protocol: results from an international phase III randomized study. *J. Clin. Oncol.* **37**, 2246–2256 (2019).
49. Hilden, J. M., Chen, C. S., Moore, R., Frestedt, J. & Kersey, J. H. Heterogeneity in MLL/AF-4 fusion messenger RNA detected by the polymerase chain reaction in t(4;11) acute leukemia. *Cancer Res.* **53**, 3853–3856 (1993).
50. Drexler, H. G., Quentmeier, H. & MacLeod, R. A. Malignant hematopoietic cell lines: in vitro models for the study of MLL gene alterations. *Leukemia* **18**, 227–232 (2004).
51. Stewart, A. K. et al. Diagnostic evaluation of t(4;14) in multiple myeloma and evidence for clonal evolution. *Leukemia* **21**, 2358–2359 (2007).
52. Sonneveld, P. et al. Treatment of multiple myeloma with high-risk cytogenetics: a consensus of the International Myeloma Working Group. *Blood* **127**, 2955–2962 (2016).
53. Stong, N. et al. The location of the t(4;14) translocation breakpoint within the NSD2 gene identifies a subset of patients with high-risk NDMM. *Blood* **141**, 1574–1583 (2023).
54. Carrabotta, M. et al. Integrated molecular characterization of patient-derived models reveals therapeutic strategies for treating CIC-DUX4 sarcoma. *Cancer Res.* **82**, 708–720 (2022).
55. Thomas, M. et al. Targeting MLL-AF4 with short interfering RNAs inhibits clonogenicity and engraftment of t(4;11)-positive human leukemic cells. *Blood* **106**, 3559–3566 (2005).
56. Ragusa, D. et al. The RS4;11 cell line as a model for leukaemia with t(4;11)(q21;q23): Revised characterisation of cytogenetic features. *Cancer Rep.* **2**, e1207 (2019).
57. Gao, Q. et al. Driver fusions and their implications in the development and treatment of human cancers. *Cell Rep.* **23**, 227–238. e223 (2018).
58. Zhang, J. et al. Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nat. Genet.* **48**, 1481–1489 (2016).
59. Kao, Y. C. et al. ETV transcriptional upregulation is more reliable than RNA sequencing algorithms and FISH in diagnosing round cell sarcomas with CIC gene rearrangements. *Genes Chromosomes Cancer* **56**, 501–510 (2017).
60. Levin, J. Z. et al. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. *Genome Biol.* **10**, R115 (2009).
61. Heyer, E. E. et al. Diagnosis of fusion genes using targeted RNA sequencing. *Nat. Commun.* **10**, 1388 (2019).
62. Grossman, R. L. et al. Toward a shared vision for cancer genomic data. *N. Engl. J. Med.* **375**, 1109–1112 (2016).
63. Tian, L. et al. Long-read sequencing unveils IGH-DUX4 translocation into the silenced IGH allele in B-cell acute lymphoblastic leukemia. *Nat. Commun.* **10**, 2789 (2019).

Acknowledgements

We thank Yoshiko Ito, Fumie Ueki, and Yoko Hokama for technical assistance. We used high-performance computing (supercomputing) resources from the Human Genome Center, the Institute of Medical Science, The University of Tokyo, Japan. We would also like to thank the patients and their families for providing clinical samples. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers JP21H05051 and JP23K15313), Japan Science and Technology Agency Moonshot R&D Program (JPMJMS2022), the Keio University Medical Science Fund, the Takeda Science Foundation, and National Cancer Center Research and Development Funds (30-A-1).

Author contributions

Z.T.: Formal analysis, Funding acquisition, Investigation, Software, Visualization, Writing—original draft, Writing—review and editing. Y. Saito: Formal analysis, Investigation, Project administration, Supervision, Visualization, Writing—original draft, Writing—review and editing. Y. Kogure: Funding acquisition, Supervision, Writing—review and editing. Y.O.-K.: Data curation, Resources, Writing—review and editing. S.F.: Resources, Writing—review and editing. S.S.: Data curation, Writing—review and editing. H.K.: Data curation, Writing—review and editing. Y. Kikukawa: Data curation, Resources, Writing—review and editing. Y.I.: Data curation, Writing—review and editing. K. Mizuno: Data curation, Writing—review and editing. Y. Shiraishi: Software, Writing—review and editing. K. Katayama: Data curation, Writing—review and editing. S.I.: Data curation, Writing—review and editing. K.I.: Data curation, Resources, Writing—review and editing. K. Murakami: Supervision, Writing—review and editing. J.K.: Supervision, Writing—review and editing. K. Kataoka: Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing—original draft, Writing—review and editing.

Competing interests

Y. Kogure has received honoraria from Daiichi Sankyo, Nippon Shinyaku, Kyowa Kirin, and Takeda. Y.O.-K. is an employee of Otsuka Pharmaceutical Co., Ltd. S.F. has received honoraria from AstraZeneca, Ono Pharmaceutical, Chugai, Janssen, Otsuka Pharmaceutical Co., Ltd., AbbVie, Takeda, Genmab, Nihon Shinyaku, and Amgen; and has received research funding from Chugai, LOXO Oncology, AbbVie, and Mitsubishi Tanabe. H.K. is an employee of Otsuka Pharmaceutical Co., Ltd. Y. Kikukawa is an employee of Otsuka Pharmaceutical Co., Ltd. K.I. has received honoraria from Astra Zeneca, Ono Pharmaceutical, Eisai, Chugai, Janssen, Symbio, Bristol Myers Squibb, Daiichi Sankyo, Otsuka Pharmaceutical Co., Ltd., AbbVie, Takeda, Eli Lilly, Genmab, Kyowa Kirin, MSD, Pfizer, Meiji Seika Pharma, Novartis, Nihon Shinyaku, and Gilead/Kite Pharma; has received research funding from MSD, AstraZeneca, AbbVie, Incyte, Symbio, Bristol Myers Squibb, Bayer, Pfizer, Janssen, Yakult, Kyowa Kirin, Daiichi Sankyo, Chugai, Beigene, Genmab, LOXO Oncology, Otsuka Pharmaceutical Co., Ltd., Regeneron, and Gilead/Kite Pharma; and has received consultancy fees from Ono Pharmaceutical, Mitsubishi Tanabe Pharma, Eisai, Chugai, Symbio, Bristol Myers Squibb, Otsuka Pharmaceutical Co., Ltd., AbbVie, Takeda, Zenyaku, Genmab, Kyowa Kirin, MSD, Carna Biosciences, and Nihon Shinyaku. K. Kataoka has received

honoraria from Ono Pharmaceutical, Eisai, Astellas Pharma, Novartis, Chugai, AstraZeneca, Sumitomo Dainippon Pharma, Kyowa Kirin, Janssen, MSD, Takeda, Symbio, Bristol Myers Squibb, Pfizer, Nippon Shinyaku, Daiichi Sankyo, Alexion Pharmaceuticals, AbbVie, Meiji Seika Pharma, Sanofi, Sysmex, Mundipharma, Incyte Corporation, Kyorin Pharmaceutical, PharmaEssentia Japan, Gilead Sciences, Kowa, Amgen, Nippon Kayaku, and Asahi Kasei Pharma; has received research funding from Otsuka Pharmaceutical Co., Ltd., Kyowa Kirin, Takeda, Chugai, Chordia Therapeutics, Genmab, Bristol Myers Squibb, and Regeneron; has received consultancy fees from Meiji Seika Pharma and Otsuka Pharmaceutical Co., Ltd.; has received scholarship endowments from Asahi Kasei Pharma, Eisai, Otsuka Pharmaceutical Co., Ltd., Kyowa Kirin, Shionogi, Daiichi Sankyo, Takeda, Sumitomo Dainippon Pharma, Chugai, Teijin Pharma, Japan Blood Products Organization, Mochida Pharmaceutical, JCR Pharmaceuticals, Nippon Shinyaku, and Nippon Kayaku; holds a patent to genetic alterations as a biomarker in T-cell lymphomas and a patent to PD-L1 abnormalities as a predictive biomarker for immune checkpoint blockade therapy; and holds stock in Asahi Genomics. The other authors declare no competing financial interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41698-026-01397-y.

Correspondence and requests for materials should be addressed to Yuki Saito or Keisuke Kataoka.