



Enhanced Detection of EGFRvIII in Tumors



A Comparative Study of Split Read— and Read Depth—Based DNA Sequencing Approaches

Nan Chen,^{*,†‡} Jingjing Feng,^{*,†} Dong Wan,^{*,†} Dongbing Li,^{*,†} and Sheng Xiao[§]

From the Advanced Molecular Pathology Institute, Soochow University and SANO,* Suzhou, China; Suzhou Sano Precision Medicine Ltd.,[†] Suzhou, China; the Department of Pharmacy,[‡] Xi An Jiaotong-Liverpool University, Suzhou, China; and the Department of Pathology,[§] Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts

Accepted for publication
November 6, 2025.

Address correspondence to
Sheng Xiao, M.D., Department
of Pathology, Brigham and
Women's Hospital, Harvard
Medical School, 75 Francis St.,
Boston, MA 02115.
E-mail: sxiao2025@126.com
or sxiao@partners.org.

Epidermal growth factor receptor variant (EGFRv)-III, a common oncogenic variant in glioblastoma and other solid tumors, results from an in-frame deletion of exons 2 to 7 from the *EGFR* gene. Detection of EGFRvIII is crucial for understanding tumor biology, guiding targeted therapies, and developing personalized treatment strategies. In this study, two detection approaches based on DNA next-generation sequencing—read depth (RD)-based and split read (SR)-based detection—were compared to evaluate their sensitivity and accuracy in identifying EGFRvIII. Thirty-one tumor samples, including glioblastoma and pancreatic adenocarcinoma, were analyzed using both methods. The SR-based method detected EGFRvIII in 20 of 31 samples, while the RD-based method identified it in only 12 samples ($P < 0.001$), demonstrating that the SR-based method had significantly higher sensitivity. RNA next-generation sequencing confirmed EGFRvIII expression in most SR-positive cases. Additionally, the SR-based method identified multiple breakpoints in several tumors, revealing intratumor heterogeneity and the subclonal origins of EGFRvIII. The RD-based method was prone to false negatives, particularly in cases with high *EGFR* amplification or low tumor cell percentage, where copy number variations could be masked by background noise. The findings highlight the sensitivity and accuracy of SR-based detection in identifying EGFRvIII and capturing intratumor heterogeneity. SR-based analysis is recommended as the method for EGFRvIII detection in both clinical and research settings. (*J Mol Diagn* 2026, 28: 179–186; <https://doi.org/10.1016/j.jmoldx.2025.11.002>)

Epidermal growth factor receptor (EGFR) is a member of the receptor tyrosine kinase family, playing a crucial role in regulating cell processes essential for proliferation, differentiation, and survival.^{1,2} Aberrant activation of EGFR, through mechanisms such as overexpression, amplification, or mutation, can lead to the dysregulation of key signaling pathways, including the Kirsten rat sarcoma viral oncogene homolog/mitogen-activated protein kinase and phosphatidylinositol 3-kinase/AKT (Ras/MAPK and PI3K/AKT) pathways.³ This dysregulation promotes uncontrolled cell growth and survival, which contribute to the development and progression of various human cancers. *EGFR* alterations have been identified in a wide range of epithelial tumors, including breast, head and neck, gastric, colorectal,

esophageal, prostate, bladder, renal, pancreatic, and ovarian, and non-small cell lung cancers.^{3,4} In primary glioblastoma (GBM), the most aggressive brain cancer, *EGFR* amplification is observed in approximately 50% of cases, often accompanied by gene rearrangements, such as specific exon deletions.⁵ Among these rearrangements, the EGFR variant (EGFRv)-III, characterized by an in-frame deletion of exons 2 to 7, is the most frequently detected variant.⁶

The full-length EGFR protein consists of four extracellular domains: leucine-rich repeat 1 (L1), cysteine-rich domain 1 (CR1), leucine-rich repeat 2 (L2), and cysteine-rich domain 2; L1 and L2 mediate ligand binding.⁷ EGFRvIII lacks the entire L1 and CR1 domains, resulting

in the absence of a ligand-binding domain.⁷ Despite its significantly weaker kinase activity compared to ligand-activated full-length EGFR, the constitutive activity of EGFRvIII is still sufficient for promoting tumor growth.^{8,9} The expression of EGFRvIII has been associated with enhanced tumor cell proliferation, invasion, angiogenesis, stemness, and resistance to therapies.¹⁰ These characteristics contribute to tumor progression, making EGFRvIII an important biomarker and therapeutic target in cancer research.¹¹

Accurate detection of EGFRvIII is pivotal to understanding tumor biology, guiding targeted therapies, and developing personalized treatment strategies. In clinical practice, EGFRvIII is most commonly detected using RNA-based methods, particularly RT-PCR, which is a well-established technique in both clinical and research laboratories. Immunohistochemistry analysis is also frequently used, whereas DNA-based next-generation sequencing (NGS) is less commonly used. Nonetheless, DNA NGS offers distinct advantages, including high sensitivity and the ability to simultaneously identify multiple genomic alterations, such as *EGFR* amplification and mutations. Two primary approaches for detecting EGFRvIII using DNA NGS are based on read depth (RD) and split read (SR) analysis. The RD-based method relies on comparing read depth differences across specific genomic regions to infer the presence of deletions, while the SR-based method directly identifies genomic breakpoints by mapping reads that span the junctions. Although both methods have been used to detect EGFRvIII, a thorough and comprehensive comparison of their sensitivity and accuracy has not been explored.

This study aimed to evaluate and compare the sensitivity of RD- and SR-based approaches for detecting EGFRvIII using targeted DNA NGS. A total of 31 tumor samples were analyzed to determine which method provides more reliable and sensitive detection of EGFRvIII. The findings are expected to provide valuable insight for optimizing EGFRvIII detection strategies in both clinical and research settings.

Materials and Methods

Sample Collection and DNA Extraction

This project was reviewed and approved by the Institutional Review Board of Suzhou Dushu Lake Hospital (Suzhou, China; approval number 220082). Written informed consent was obtained from all participants for the use of leftover discard materials after diagnostic testing, in accordance with institutional guidelines and ethics standards. Tumor tissues were collected from both frozen samples and formalin-fixed paraffin-embedded (FFPE) samples. DNA extraction was performed using the AllPrep DNA/RNA FFPE kit (Qiagen, Hilden, Germany), following the manufacturer's protocols. The quality and quantity of extracted DNA were assessed using a NanoDrop spectrophotometer

(Thermo Fisher Scientific, Waltham, MA) and Qubit fluorometer (Thermo Fisher Scientific).

DNA Library Preparation and Sequencing

DNA libraries were prepared using the Rapid Plus DNA library preparation kit for Illumina platforms (ABclonal, Wuhan, China) and the NadPrep Hybrid Capture Reagents Box (Nanodigm, Nanjing, China). Briefly, genomic DNA was fragmented to a mean size of 200 to 300 bp using a Covaris sonicator (Wolburn, MA). The fragmented DNA was then end-repaired, dA-tailed, and ligated to Illumina sequencing adapters (San Diego, CA). An in-house-designed, targeted DNA NGS panel was used to capture tumor-related genes, including the entire coding region of *EGFR* ([Supplemental Table S1](#)). Hybrid capture was performed using biotinylated DNA probes, and enriched DNA fragments were amplified by PCR. The final libraries were quantified using a KAPA library quantification kit (KAPA Biosystems, Wilmington, MA) and sequenced on a NovaSeq 6000 platform (Illumina) with 2×150 bp paired-end reads.

Validation of EGFRvIII Breakpoints by Nested PCR

Genomic DNA was amplified using a nested PCR approach with two primer sets flanking the EGFRvIII breakpoint. Both rounds were conducted under identical thermal cycling conditions. Each 25- μ L reaction contained 2.5 μ L of $10\times$ Taq buffer (Takara, Kusatsu, Japan), 2 μ L of dNTP mixture (2.5 mmol/L), 1 μ L of each primer (10 μ mol/L), 0.2 μ L of Taq DNA polymerase (5 U/ μ L), and 100 to 200 ng of genomic DNA, with double-distilled H₂O added to reach the final volume. The cycling program consisted of an initial denaturation at 95°C for 5 minutes; 35 cycles of 95°C for 30 seconds, annealing at 56 to 60°C for 30 seconds, and extension at 72°C for 1 minute, followed by a final extension at 72°C for 10 minutes. The final PCR products were separated on a 3% agarose gel and confirmed by Sanger sequencing to verify the EGFRvIII-specific junction.

Targeted RNA NGS of EGFRvIII

For RNA analysis, total RNA was extracted using the AllPrep DNA/RNA FFPE kit (Qiagen). RNA integrity was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA), and only samples with an RNA integrity number of six or more were included in the study. Reverse transcription was performed using the SuperScript III Reverse Transcriptase kit (InvitroGen, Carlsbad, CA) according to the manufacturer's instructions. Targeted PCR enrichment of *EGFR* exon 8 was performed using nested gene-specific primers (first round, 5'-CCGTCTTCTCCATCTCATAGCTGTC-3'; second round, 5'-GGCCCCACAGGCTCGGAC-3'). The PCR products were sequenced on the NovaSeq 6000 platform (Illumina) with 2×150 bp paired-end reads.

Fluorescence *in Situ* Hybridization for *EGFR* Amplification

Fluorescence *in situ* hybridization was performed on FFPE tissue sections using the *EGFR* SpectrumOrange probe (7p11) and the *CEP7* SpectrumGreen probe (7p11.1-q11.1) as a chromosome 7 control (HealthCare, Wuhan, China). Standard protocols were followed for deparaffinization, hybridization, washing, and counterstaining according to the manufacturer's instructions. Fluorescence signals were visualized using the BX43 microscope (Olympus, Tokyo, Japan) with a 100× oil immersion objective and iMstar imaging software version 6.04.17.00 (IMSTAR, Paris, France). *EGFR* and *CEP7* signals were counted in 50 nuclei per region. Positive cases were defined by gene amplification present in $\geq 10\%$ of tumor cells, characterized by clusters of at least four *EGFR* signals; or $>40\%$ of cells displaying more than four copies of the *EGFR* signal; or an *EGFR*-to-*CEP7* ratio of >2 over all scored nuclei and calculated using the sum of *EGFR* divided by the sum of *CEP7* when the mean *CEP7* per cell is more than copies.

Bioinformatics Analysis

Data Preprocessing

Sequencing data were processed using Fastp version 0.23.2 (<https://github.com/OpenGene/fastp>) to remove low-quality reads and adapter sequences. Clean reads were aligned to the human reference genome (hg19) using the Burrows-Wheeler Aligner version 0.7.17-r1188 (<https://github.com/lh3/bwa>) with default parameters. The alignment results were sorted and indexed using SAMtools version 1.9 (<https://github.com/samtools/samtools>). RNA sequencing data were mapped to the reference genome using Spliced Transcripts Alignment to a Reference (STAR) version 2.7.8a (<https://github.com/alexdobin/STAR>) with parameters that included twopassMode Basic, outSAMtype BAM SortedByCoordinate, alignInsertionFlush Right and chimJunctionOverhangMin 8.

Copy Number Variation Analysis

Read depth at each position in the target region was calculated using BAMDST version 1.0.9 (<https://github.com/shiquan/bamdsl>) with a bin size of approximately 200 bp. Copy number variation analysis was performed using an in-house script that normalized read depth based on the total depth in the target region to account for variations in sequencing read counts across samples. Copy number variations were identified by comparing the standardized depth between tumor and normal samples. The *EGFR* copy number was adjusted to a 100% tumor cell background according to estimated tumor purity.

Estimation of Tumor Purity

Tumor purity was estimated using the deviation of B-allele frequency (BAF) at heterozygous single-nucleotide polymorphism loci. Specifically, in a diploid region of normal

cells with a heterozygous genotype (A and B), the expected BAF is approximately 0.5. In contrast, in tumor cells with allelic imbalance—such as a heterozygous deletion (loss of B)—the observed BAF shifts toward a value that reflects the mixture of tumor and normal cells.

For example, in a sample with 50% tumor purity, if tumor cells harbor a deletion of the B allele (resulting in an A genotype) and normal cells retain the AB genotype, the observed BAF becomes:

$$\text{Observed BAF} = \frac{0.5 * (B_{\text{normal}}) + 0.5 * 0(B_{\text{tumor}})}{0.5 * (A_{\text{normal}} + B_{\text{normal}}) + 0.5 * (A_{\text{tumor}} + B_{\text{tumor}})}$$

Assuming:

$$A_{\text{normal}} = B_{\text{normal}} = 1$$

$$A_{\text{tumor}} = 1, B_{\text{tumor}} = 0 \text{ (heterozygous deletion in tumor cells)}$$

$$\text{Observed BAF} = \frac{0.5 * 1 + 0.5 * 0}{0.5 * (1 + 1) + 0.5 * (1 + 0)} \approx 0.33$$

By modeling expected BAF shifts under different purity levels and comparing them with observed values at informative loci, tumor purity was inferred.

EGFRvIII Detection Methods

Two strategies were employed to detect EGFRvIII at the DNA level: the RD-based method and the SR-based method (Figure 1). In the RD-based approach, the deletion of *EGFR* exons 2 to 7 was evaluated using the *t*-test to compare read depths between the putative deletion region (exons 2 to 7) and the nondeleted downstream region (exons 8 to 28). A statistically significant decrease ($P < 0.05$) in read depth at exons 2 to 7 was considered indicative of EGFRvIII presence. In the SR-based approach, Seeksv version 1.2.3 (<https://github.com/qiukunlong/seeksv>) was used to extract reads partially mapped to *EGFR* introns 1 and 7 (getclip: $t = 0.5$) and call breakpoints (getsv: $b = 3$, $t = 0.5$, and $f = 0.00001$). These split reads directly indicated the presence of genomic breakpoints associated with EGFRvIII. A minimum of three supporting split reads was required to confidently call an EGFRvIII event (Supplemental Figure S1). At the RNA level, an in-house script was used to directly extract reads supporting the fusion between *EGFR* exons 1 and 8. The aligned sequences were then manually inspected using the Integrative Genomics Viewer (IGV; <https://igv.org>).

Results

Higher Sensitivity of Split Read Analysis in Detecting EGFRvIII

The performance of two methods, RD-based detection and SR-based detection, was compared for identifying EGFRvIII in 31 tumor samples with *EGFR* amplification by DNA NGS. The SR-based method detected EGFRvIII in 20 of 31

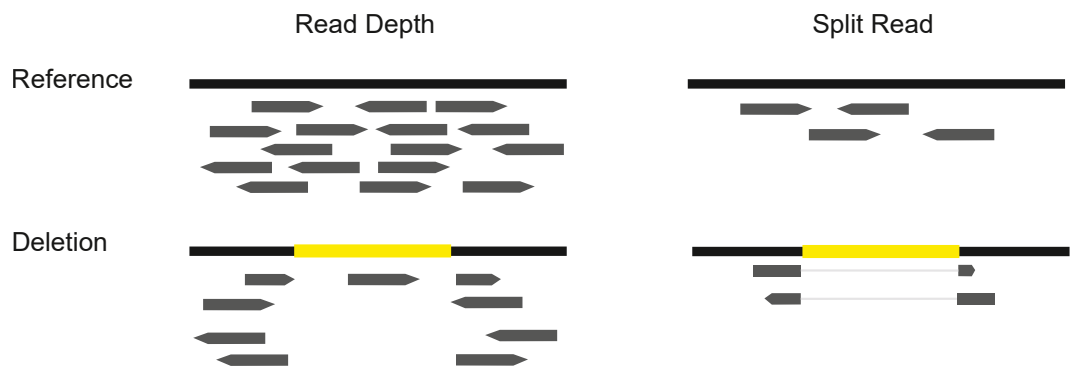


Figure 1. Schematic illustration of the read depth (RD) and split read (SR) methods for detecting EGFRvIII. The **yellow region** highlights exons 2 to 7 of *EGFR*.

samples, while the RD-based method identified it in only 12 of 31 samples ($P < 0.001$) (Table 1, Supplemental Tables S2 and S3, and Supplemental Figures S2-S5). This result indicates that the SR-based method is significantly more sensitive than the RD-based method for detecting EGFRvIII. RNA NGS confirmed the expression of EGFRvIII in

13 of 16 SR-positive cases (Supplemental Figures S6 and S7), with three false-negative results (cases 15, 20, and 21), each showing a limited number of supporting split reads (19, 3, and 7, respectively) (Supplemental Table S2). Notably, case 3, initially identified as EGFRvIII-negative by both the RD and SR methods, was later confirmed to

Table 1 Tumor Types, *EGFR* Copy Numbers, and EGFRvIII in 31 Tumors

Case no.	Cancer	<i>EGFR</i> copy number	EGFRvIII detection by RD	EGFRvIII detection by SR	RNA validation
Case 1	GBM	17.9	—	+	+
Case 2	GBM	45.5	+	+	+
Case 3	GBM	12.9*	—	—	+
Case 4	GBM	82	—	+	+
Case 5	GBM	24	+	+	NA
Case 6	PAAD	2	—	+	+
Case 7	GBM	37.8	+	+	+
Case 8	GBM	38.3	+	+	+
Case 9	GBM	23.4	+	+	+
Case 10	GBM	19.9	—	+	+
Case 11	GBM	36.6	+	+	+
Case 12	GBM	80.1	+	+	+
Case 13	GBM	80.8	+	+	+
Case 14	GBM	92.8	—	+	+
Case 15	GBM	77.7	—	+	—
Case 16	GBM	16.7	+	+	NA
Case 17	GBM	23.5	+	+	NA
Case 18	GBM	17.6	+	+	NA
Case 19	GBM	27.9	+	+	+
Case 20	GBM	48.9	—	+	—
Case 21	GBM	28	—	+	—
Case 22	GBM	29.1	—	—	—
Case 23	GBM	5.8	—	—	—
Case 24	GBM	108.8	—	—	—
Case 25	GBM	8.5	—	—	—
Case 26	GBM	25.3	—	—	—
Case 27	GBM	5.9	—	—	—
Case 28	GBM	8.5	—	—	—
Case 29	GBM	7	—	—	—
Case 30	GBM	64	—	—	—
Case 31	GBM	11.1	—	—	—

*The *EGFR* copy number in case 3 could not be accurately evaluated; the value shown represents the mean across all cells. In contrast, copy numbers in other cases reflect measurements in 100% tumor cells.

+, Positive; —, negative; GBM, glioblastoma; NA, not available; PAAD, pancreatic adenocarcinoma; RD, read depth; SR, split read.

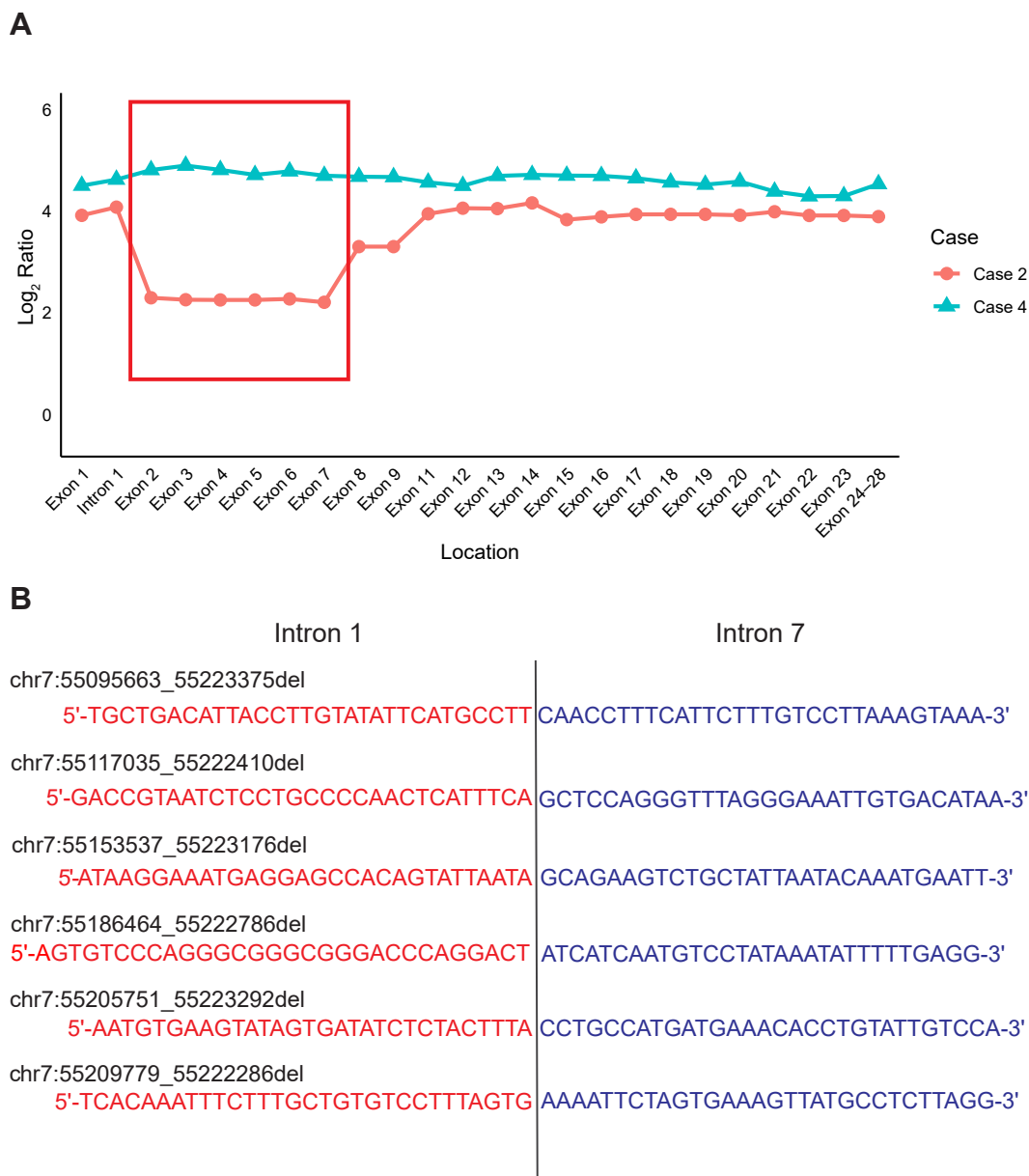


Figure 2 EGFRvIII detection in cases 2 and 4. **A:** Read depth (RD)—based detection of EGFRvIII. The y-axis represents the *EGFR* read depth $\log_2$ ratio in the tumor sample to the normal control. The **red boxed region** highlights exons 2 to 7. Case 2 (**orange line**) shows a clear depth drop in this region, indicating EGFRvIII positivity, while case 4 (**blue line**) shows no reduction. **B:** Six EGFRvIII breakpoints in case 4 detected by the split read (SR) method, with red and dark blue sequences located in introns 1 and 7, respectively.

express EGFRvIII by RNA NGS, suggesting that the DNA-based methods may miss low-abundance variants in some cases. Cases 22 to 31 showed no evidence of EGFRvIII by any of the three detection methods (RD, SR, and RNA NGS), indicating that these tumors were truly EGFRvIII-negative.

Limitations of RD—Based Detection

The RD-based method was prone to false-negative results, particularly in cases with high *EGFR* amplification or low

tumor cell percentage. For example, case 4 exhibited extreme *EGFR* amplification (copy number, 82) with a very high sequencing depth (mean depth, >40,000 reads) ([Supplemental Table S2](#)), but the RD method failed to detect EGFRvIII due to the masking effect of the high read depth ([Figure 2A](#)). In contrast, the SR method identified six distinct breakpoints in this case, confirming the presence of EGFRvIII ([Figure 2B](#)). Similarly, case 6, which lacked *EGFR* amplification but contained a low percentage of tumor cells with EGFRvIII, was missed by the RD method but detected by the SR method. These

findings highlight the limitations of RD-based detection in identifying subclonal alterations in heterogeneous tumor samples.

Breakpoint Analysis Reveals Multiple Subclones in Tumors

The SR-based method identified multiple breakpoints within the intronic regions of *EGFR* (introns 1 and 7) in nearly half of the SR-positive samples (9 of 20), suggesting the presence of multiple subclones within the tumors ([Supplemental Table S2](#)). Among these, five cases had two subclones, three cases had three subclones, and one case had six subclones. These breakpoints were further validated by nested PCR followed by Sanger sequencing, supporting the accuracy of the SR-based method ([Supplemental Figures S8-S10](#)). This finding underlines the ability of the SR method to capture intratumor heterogeneity and provides valuable insight into the clonal evolution of tumors.

EGFRvIII Detection in Nonamplified Tumors

In addition to GBM samples with *EGFR* amplification, EGFRvIII was identified in one pancreatic adenocarcinoma sample (case 6) without *EGFR* amplification ([Supplemental Table S2](#) and [Supplemental Figure S11](#)). This finding demonstrates that EGFRvIII can occur independently of *EGFR* gene amplification, highlighting the importance of comprehensive detection methods that are not solely reliant on copy number analysis.

EGFRvIII Validation at the Transcriptional Level

RNA validation was performed on 27 samples with sufficient RNA quality. Among these, 23 samples showed concordance between SR-based DNA detection and RNA expression, supporting the accuracy of the SR-based method. Three discordant cases were EGFRvIII positive by SR-based DNA detection but negative by RNA analysis.

To assess potential false positives in SR-based DNA detection, manual IGV inspection confirmed the presence of high-quality split reads ([Supplemental Table S2](#) and [Supplemental Figures S2-S4](#)). Furthermore, PCR assays followed up by Sanger sequencing confirmed the presence of EGFRvIII ([Supplemental Figures S8](#) and [S10](#)). The absence of the EGFRvIII transcript in these cases, despite detection at the DNA level, remains unclear. Possible reasons could include alternative splicing, promoter methylation, post-transcriptional regulation, or rare single-nucleotide polymorphisms at the primer binding site. Notably, EGFRvIII expression has been reported to be absent even in tumor cells harboring the underlying genetic alteration, potentially due to epigenetic suppression mechanisms.¹²

Discussion

Detecting EGFRvIII is important for understanding the biology of GBM and other solid tumors, as well as for guiding targeted therapies. This study provides a comprehensive comparison between two NGS approaches—RD-based and SR-based detection—in identifying EGFRvIII. The results highlight the significant advantages of the SR-based method over the RD-based method in terms of sensitivity, accuracy, and the ability to capture intratumor heterogeneity.

The SR-based method demonstrated a strikingly higher sensitivity in detecting EGFRvIII compared to the RD-based method. In this study, SR analysis identified EGFRvIII in 20 of 31 tumor samples, whereas the RD method detected it in only 12 samples. This superior sensitivity is attributed to the direct identification of genomic breakpoints by SR analysis, which relies on mapping reads that span the junctions of the deletion. This approach requires minimal supporting reads (as few as three in this study) to confidently call an EGFRvIII event, making it less susceptible to sequencing noise, tumor cell percentage, and heterogeneity. In contrast, RD-based detection relies on comparing read depth differences across specific genomic regions to infer deletions.¹³ This method is prone to false negatives, especially in cases with high *EGFR* amplification or low tumor cell percentage, where copy number variations can be masked by background noise.

The limitations of the RD-based method were particularly evident in cases with extreme *EGFR* amplification or low tumor cell percentage. For instance, case 4 exhibited extreme *EGFR* amplification (copy number, 82) with a very high sequencing depth (mean depth, >40,000 reads), yet the RD method failed to detect EGFRvIII due to the masking effect of the high read depth. Conversely, the SR method identified six distinct breakpoints in this case, confirming the presence of EGFRvIII. Similarly, case 6, which lacked *EGFR* amplification but contained a low percentage of tumor cells with EGFRvIII, was missed by the RD method but detected by the SR method. These findings underline the challenges of using RD-based detection in heterogeneous tumor samples, where subclonal alterations such as EGFRvIII are often present at low abundance.

One of the most significant advantages of the SR-based method is its ability to identify multiple breakpoints within a single tumor sample. In nearly half of samples (9 of 20), the SR method detected multiple breakpoints within the intronic regions of *EGFR* (introns 1 and 7), suggesting the presence of multiple subclones. For example, case 4 exhibited six distinct breakpoints, indicating the coexistence of multiple EGFRvIII subclones. This ability to capture intratumor heterogeneity is crucial for understanding the clonal evolution of tumors and has important implications for the detection of minimal residual disease and clone evolution.

RNA NGS provided further validation of the SR-based detection method. In 13 of 16 SR-positive cases, RNA analysis confirmed the expression of EGFRvIII fusion transcripts, supporting the accuracy of the SR method at the transcription level. The three exceptions (cases 15, 20, and 21) may be attributable to technical limitations in RNA analysis or suppression of EGFRvIII expression. Additionally, case 3, which was initially identified as EGFRvIII negative by both the RD and SR methods, was later confirmed to express EGFRvIII by RNA NGS. This finding highlights the potential for DNA-based methods to miss low-abundance variants in some cases, emphasizing the importance of multimodal validation approaches.

This study has important implications for both clinical and research settings. Given the superior sensitivity and accuracy of the SR-based method in detecting EGFRvIII, its use is strongly recommended as the preferred approach in targeted DNA NGS workflows. This method not only improved detection rates but also provided valuable insight into intratumor heterogeneity, which is crucial for understanding tumor biology and guiding personalized therapies. Additionally, the findings highlight the importance of comprehensive detection methods that are not solely reliant on copy number analysis, as EGFRvIII can occur independently of *EGFR* gene amplification, as demonstrated in the nonamplified pancreatic adenocarcinoma sample (case 6). Because of the increased sensitivity of detection, EGFRvIII may be identified in small subclones; nonetheless, there are several reasons why such findings remain clinically significant. First, EGFRvIII is a well-established oncogenic driver, and even small populations of EGFRvIII-positive cells can expand under therapeutic pressure, ultimately contributing to treatment resistance and recurrence.¹⁴ Second, a multicenter study showed that the presence of any detectable EGFRvIII, regardless of clonal fraction, was associated with inferior progression-free survival compared with *EGFR*-amplified, EGFRvIII-negative GBMs.¹⁰ Finally, although subclonal, EGFRvIII may still represent a therapeutically actionable target, similar to the treatment of subclonal *EGFR* mutations in lung cancer. Taken together, these data support the view that the detection of EGFRvIII, even at subclonal levels, is clinically meaningful and should be considered a positive finding in diagnostic assays.

Conclusion

This study compared two approaches—RD-based and SR-based detection—for their ability to identify the EGFRvIII oncogenic variant in tumor samples. The comprehensive analysis revealed that the SR-based method significantly outperformed the RD-based method in sensitivity, accuracy, and capability to capture intratumor heterogeneity. The SR-based detection method identified EGFRvIII in 20 of 31 tumor samples, whereas the RD-based method

detected it in only 12 samples. This superior sensitivity of the SR method is attributable to its direct identification of genomic breakpoints, which required minimal supporting reads and was less affected by sequencing noise, tumor cell percentage, or heterogeneity. In contrast, the RD-based method was prone to false-negative results, particularly in cases with high *EGFR* amplification or low tumor cell percentage, where copy number variations could be masked by background noise. Furthermore, the ability of the SR-based method to identify multiple breakpoints within a single tumor sample provided valuable insight into intratumor heterogeneity and subclonal evolution. This feature is crucial for understanding tumor biology and guiding personalized treatment strategies, including the detection of minimal residual disease. RNA validation further confirmed the accuracy of SR-based detection at the transcription level, supporting its reliability as a diagnostic tool. Given the superior performance of the SR-based method in detecting EGFRvIII, its adoption is strongly recommended as the preferred approach in clinical and research settings. This method not only enhanced detection rates but also provided a more comprehensive understanding of tumor genetics, which is essential for advancing targeted therapies and improving patient outcomes. Future studies should focus on optimizing SR-based detection workflows and exploring their potential applications in other cancer types and clinical scenarios.

Author Contributions

N.C. validated data and reviewed and edited the manuscript; J.F. and D.W. curated data and reviewed and edited the manuscript; D.L. wrote the original draft and reviewed and edited the manuscript; and S.X. conceptualized the study, developed methodology, and reviewed and edited the manuscript.

Disclosure Statement

None declared.

Supplemental Data

Supplemental material for this article can be found at <https://doi.org/10.1016/j.jmoldx.2025.11.002>.

References

1. van de Water JA, Bagci-Onder T, Agarwal AS, Wakimoto H, Roovers RC, Zhu Y, Kasmieh R, Bhore D, Van Bergen en Henegouwen PM, Shah K: Therapeutic stem cells expressing variants of EGFR-specific nanobodies have antitumor effects. *Proc Natl Acad Sci U S A* 2012, 109:16642–16647
2. Treda C, Popeda M, Ksiazkiewicz M, Grzela DP, Walczak MP, Banaszczyk M, Peciak J, Stoczynska-Fidelus E, Rieske P: EGFR

- activation leads to cell death independent of PI3K/AKT/mTOR in an AD293 cell line. *PLoS One* 2016, 11:e0155230
3. Yarden Y, Sliwkowski MX: Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2001, 2:127–137
 4. Arteaga CL: Epidermal growth factor receptor dependence in human tumors: more than just expression? *Oncologist* 2002, 7:31–39
 5. Zheng L, Luthra R, Alvarez HA, San Lucas FA, Duose DY, Wistuba II, Fuller GN, Ballester LY, Roy-Chowdhuri S, Sweeney KJ, Rashid A, Yang RK, Chen W, Liu A, Wu Y, Albarracin C, Patel KP, Routbort MJ, Sahin AA, Ding Q, Chen H: Intragenic EGFR::EGFR.E1E8 fusion (EGFRvIII) in 4331 solid tumors. *Cancers (Basel)* 2023, 16:6
 6. Rutkowska A, Stoczynska-Fidelus E, Janik K, Włodarczyk A, Rieske P: EGFR(vIII): an oncogene with ambiguous role. *J Oncol* 2019, 2019:1092587
 7. Gan HK, Cvrljevic AN, Johns TG: The epidermal growth factor receptor variant III (EGFRvIII): where wild things are altered. *FEBS J* 2013, 280:5350–5370
 8. Batra SK, Castelino-Prabhu S, Wikstrand CJ, Zhu X, Humphrey PA, Friedman HS, Bigner DD: Epidermal growth factor ligand-independent, unregulated, cell-transforming potential of a naturally occurring human mutant EGFRvIII gene. *Cell Growth Differ* 1995, 6: 1251–1259
 9. Moscatello DK, Montgomery RB, Sundareshan P, McDanel H, Wong MY, Wong AJ: Transformational and altered signal transduction by a naturally occurring mutant EGF receptor. *Oncogene* 1996, 13:85–96
 10. Felsberg J, Hentschel B, Kaulich K, Gramatzki D, Zacher A, Malzkorn B, Kamp M, Sabel M, Simon M, Westphal M, Schackert G, Tonn JC, Pietsch T, von Deimling A, Loeffler M, Reifenberger G, Weller M: Epidermal growth factor receptor variant III (EGFRvIII) positivity in EGFR-amplified glioblastomas: prognostic role and comparison between primary and recurrent tumors. *Clin Cancer Res* 2017, 23:6846–6855
 11. Sterner RC, Sterner RM: EGFRvIII and EGFR targeted chimeric antigen receptor T cell therapy in glioblastoma. *Front Oncol* 2024, 14: 1434495
 12. Del Vecchio CA, Giacomini CP, Vogel H, Jensen KC, Florio T, Merlo A, Pollack JR, Wong AJ: EGFRvIII gene rearrangement is an early event in glioblastoma tumorigenesis and expression defines a hierarchy modulated by epigenetic mechanisms. *Oncogene* 2013, 32: 2670–2681
 13. Zare F, Dow M, Monteleone N, Hosny A, Nabavi S: An evaluation of copy number variation detection tools for cancer using whole exome sequencing data. *BMC Bioinformatics* 2017, 18:286
 14. Weller M, Butowski N, Tran DD, Recht LD, Lim M, Hirte H, Ashby L, Mechtler L, Goldlust SA, Iwamoto F, Drappatz J, O'Rourke DM, Wong M, Hamilton MG, Finocchiaro G, Perry J, Wick W, Green J, He Y, Turner CD, Yellin MJ, Keler T, Davis TA, Stupp R, Sampson JH: Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. *Lancet Oncol* 2017, 18:1373–1385