Neuro-Oncology
26(1), 85–99, 2024 | https://doi.org/10.1093/neuonc/noad153 | Advance Access date 24 August 2023

A novel EGFR variant EGFRx maintains glioblastoma stem cells through STAT5

Wei Huang†, Jing Li†, Hongtao Zhu, Xuhui Qin, Chao Chen, Bing Wang, Jinxia Wei, Yanyang Song, Xia Lu, Zhongyu Li, Wenqing Xia, Aodi He, Lidong Cheng, Xingjiang Yu, Kai Shu*, Wei Wang*,

All author affiliations are listed at the end of the article

†These authors have contributed equally to this work.

*Corresponding Authors: Kai Shu, No. 1095, Jiefang Avenue, Wuhan, 430030, China (kshu@tjh.tjmu.edu.cn); Wei Wang, No. 13 Hangkong Road, Wuhan, 430030, China (wekle@hust.edu.cn)

Abstract

Background. Glioblastomas are universally lethal brain tumors containing tumor-propagating glioblastoma stem cells (GSCs). EGFR gene amplification or mutation is frequently detected in GBMs and is associated with poor prognosis. However, EGFR variants in GSCs and their role in the maintenance of GSCs and progression of GBM are unclear.

Methods. EGFR variants were detected through bioinformatic HISAT-StringTie-Ballgown pipeline and verified through 5′ RACE, RT-PCR, ribonuclease protection, and northern blotting assays. EGFRx function was investigated through neurosphere, cell viability, intracranial xenograft and RNA-seq assays. EGFRx-STAT5 signaling was investigated through western blotting, coimmunoprecipitation, immunofluorescence, luciferase reporter, RT-PCR and CUT&Tag assays.

Results. We identified a novel EGFR variant (EGFRx), that is specifically expressed in GSCs. Unlike the EGFRvIII variant, which lacks exons 2–7, EGFRx is characterized by the absence of exons 2–14, and encodes an EGFR protein that does not possess the entire extracellular ligand-binding domain. We observed that EGFRx exhibits significant glycosylation, is required for GSC self-renewal, proliferation, and tumorigenesis, and highly active in glioblastomas compared to normal brain tissue. Mechanistically, EGFRx constitutively and specifically activates STAT5 in GSCs through spontaneous asymmetric dimerization of the kinase domain.

Conclusions. EGFRx plays essential roles in the maintenance of the GSC phenotype through constitutive activation of STAT5 and promotes GBM progression, suggesting that EGFRx-STAT5 signaling represents a promising therapeutic target for GBM.

Key Points

1. EGFRx is an EGFR variant specifically expressed in glioblastoma stem cells.
2. EGFRx is required for GSC self-renewal, proliferation, and tumorigenesis.
3. STAT5 is a major downstream effector of EGFRx signaling.

Glioblastoma (GBM; World Health Organization grade IV glioma) is the most prevalent and malignant primary brain tumor.[1] Despite recent advances in cancer treatment, current therapies for GBM, including surgical resection, radiotherapy, and chemotherapy, remain largely ineffective. The estimated median survival of affected patients is less than a year,[2] mainly due to therapeutic resistance and rapid tumor recurrence.[3] GBM-derived stem cells (GSCs) represent a population of undifferentiated cells that are characterized by their self-renewal, proliferative, differentiation capacities and their in vivo tumorigenic capacity to propagate tumors that recapitulate the phenotype of the original tumor from which they are

Importance of the Study

GSCs promote GBM progression and confer high resistance to chemotherapy and radiotherapy. In this study, we discovered a mechanism for the maintenance of GSC population through a novel EGFR variant called EGFRx. EGFRx exhibits some unique characteristics compared to other known EGFR variants, such as the absence of the entire extracellular ligand-binding domain, its restricted expression in certain GSCs , and the specific regulation of downstream STAT5 signaling. Moreover, we found that a subset of GBM samples predominantly express EGFRx. These results highlight the heterogeneity of EGFR and have profound implications for understanding the underlying mechanisms driving GBM development.

isolated.[4] Cancer stem cells have been shown to confer high resistance to chemotherapy and radiotherapy.[5,6] Hence, targeting GSCs is an attractive strategy for treating GBM.

The EGFR receptor consists of an extracellular ligand-binding domain (ECD), a hydrophobic transmembrane domain (TMD), and an intracellular cytoplasmic domain (ICD) that is further divided into a tyrosine kinase domain and C-terminal tail. Ligand binding to the ECD triggers dimerization of the receptor and induces autophosphorylation of the C-terminal tail.[7] EGFR activation depends on the asymmetric dimerization of the kinase domain, in which one kinase domain allosterically activates the other.[8] EGFR activation leads to the activation of multiple downstream effectors, such as phosphoinositide 3-kinase (PI3K), mitogen-activated protein kinase (MAPK), and signal transducer and activator of transcription (STAT).[9]

Aberrant EGFR activation is frequently detected in GBMs and is associated with poor prognosis. Multiple mechanisms contribute to abnormal EGFR activation in GBMs, such as EGFR gene amplification and mutation.[10] *EGFR* gene amplification/mutation is associated with the TCGA-GBM classical subtype[11] and enriched in the astrocyte-like GBM cells.[12] As the most common EGFR mutant in GBM, EGFRvIII contains an in-frame deletion of exons 2–7 and is often expressed in GBMs with EGFR amplification.[13–15] EGFR vIII is constitutively phosphorylated and has been demonstrated to be tumorigenic.[16,17] Other EGFR variants in GBM include ECM truncation mutants (EGFR vI and vII),[13,18] several ECM point mutation variants,[19,20] and cytoplasmic tail deletion mutants (EGFRvIV and vV).[15] In addition, circular RNA in GSCs encodes a rolling-translated EGFR protein complex that persistently activates oncogenic EGFR signaling.[21] The presence of multiple EGFR variants highlights the GBM heterogeneity and may limit the effectiveness of EGFR-targeted therapy for GBMs.[22] Therefore, identifying other EGFR variants in GBM cells, especially in GSCs, is of great importance.

In this study, we identified a novel *EGFR* transcript (*EGFRx*) specifically expressed in GSCs. *EGFRx* lacks exons 2-14 and translates an in-frame EGFR truncation protein missing the entire ECD. EGFRx constitutively and specifically activates STAT5 through the asymmetric dimerization of its kinase domain. We further demonstrate that the EGFRx signaling is essential for the self-renewal and proliferation of GSCs and that its upregulation in GBMs may contribute to GBM progression.

Materials and Methods

Cell Culture

GSCs (456, 387, 4121, and 3691) were kind gifts from Dr. Jeremy N Rich (Department of Neurology, University of Pittsburgh, Pittsburgh, Pennsylvania), Dr. Shideng Bao (Department of Cancer Biology, Lerner Research Institute, Cleveland Clinic, Cleveland, USA), and Dr. Xingjiang Yu (Department of Histology and Embryology, Huazhong University of Science and Technology). Further details regarding additional validation assays for GSCs and cell culturing conditions can be found in the Supplementary Materials and Methods.

Plasmids and Reagents, Western blotting, Immunofluorescence, Coimmunoprecipitation (CO-IP), and Cytoplasm/Nuclear Fractionation

See the Supplementary Materials and Methods for details.

GBM Samples, CD133$^+$ Cell Sorting, RNA Isolation, and RT-PCR

The GBM tissues were obtained from patients who underwent surgical treatments in Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. This study was approved by the ethical committee under the Helsinki Criteria. Informed consent was obtained from all patients included in this study. The CD133$^+$ and CD133$^-$ cells were isolated from the GBM fresh samples through flow cytometry. The RNA was extracted with TRIzol, and cDNA was synthesized for RT-PCR detection. See the Supplementary Materials and Methods for details.

5′ RACE

The HiScript-TS 5′/3′ RACE Kit (Vazyme, RA101-01) was used to obtain the 5′ UTR sequence of the *EGFRx* transcript with the reverse primer E16-17R (5′ ATCTTAGGCCCATTCGT 3′). The purified 5′ UTR fragments were cloned into the pUC-19 vector, followed by sequencing and alignment with *EGFR* mRNA and genomic sequences.

Endonuclease Protection

The biotin-labeled probe listed in Table S3 was generated by PCR with biotin-16-dUTP (Roche, 11093070910) to target *EGFRwt*, *EGFRx,* or both transcripts. RNase protection experiments were conducted using the RPA III RNase Protection Assay Kit (Thermo Fisher Scientific, AM1414). RNA was applied to a 5% TBE/urea gel and transferred to a nylon membrane (Thermo Fisher Scientific, AM10100). The biotinylated probes bound to membranes were recognized by IRDye 680RD streptavidin (LICOR, 926-68079), and the fluorescence signal was captured by a Licor Odyssey-CLx machine.

Northern Blot

Total RNA was separated by formaldehyde gel electrophoresis. The negatively charged RNA was transferred onto a nylon membrane driven by capillary forces (overnight). Membranes were cross-linked to the RNA (60 °C, 2 h). Subsequently, the RNA fixed in the membrane was hybridized with the biotin-labeled probe (68 °C, 16 h). The sequences for the E26 and E1-15 probes were the same as the endonuclease protection probes, as listed in Supplementary Table S3. The hybridization signal was recognized by IRDye 680RD streptavidin and detected by a Licor Odyssey-CLx machine.

Stable Cell Establishment, Cell Viability, and Limiting Dilution Assay

The lentivirus packaging system was used to establish the knockdown and overexpressing stable cells. Cell viability was measured using the CellTiter-Glo Assay Kit (Promega, #G7571). GSCs were plated with diluted cell numbers and the sphere-forming frequency was calculated using extreme limiting dilution analysis.[23] See the Supplementary Materials and Methods for details.

Intracranial Xenografts

Healthy male nude mice at 4–6 weeks old were purchased from GemPharmatech and randomly grouped (6 mice per group). 456 cells (2×10^4/mouse, control vs. *EGFRx* knockdown) or U251 cells (5×10^5/mouse, control vs. EGFRx overexpression) were injected into the right cerebral cortex of nude mice at a depth of 3.5 mm. All mice were maintained in a 12-hour light/12-hour dark cycle with no more than 5 mice per cage. The mice were sacrificed on the same day when neurological signs were observed in either mouse group. Brains were harvested and fixed in 4% formaldehyde overnight. Hematoxylin and eosin (H&E) staining was performed to analyze tumor growth. In parallel, the survival status of the injected mice in each group was recorded and the Kaplan–Meier survival analysis was performed in GraphPad Prism.

HISAT-Stringtie-Ballgown Analysis

The RNA-seq reads were subjected to the HISAT-Stringtie-ballgown analysis, as described.[24] For further information, please refer to the Supplementary Materials and Methods. Details regarding the analyzed samples, along with the expression values of the EGFR transcripts, can be found in Supplementary Table S1.

Differential Gene Expression and Gene Function Enrichment Analysis

Differential gene expression analysis of the RNA-seq data was conducted using DESeq2 (v1.34.0), edgeR (v3.36.0), and limma (v3.50.3) software packages. Gene function enrichment analysis was performed using Gene Ontology (GO) and Gene Set Enrichment Analysis (GSEA). See the Supplementary Materials and Methods for details.

CUT&Tag Analysis

The CUT&Tag libraries were constructed using the Hyperactive Universal CUT&Tag Assay Kit for Illumina (Vazyme, TD903, China) and sequenced on the Novaseq PE150 platform. The CUT&Tag analysis was performed as previously described.[25] See the Supplementary Materials and Methods for details.

Statistical Analysis

Data are presented as the mean ± SD, and statistical tests (Student's *t* test, Wilcox test, and log-rank test) are indicated in the respective figure legends. Statistics were performed with GraphPad Prism or R packages. p values of <0.05 were assumed to be statistically significant. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

Results

Identification of the *EGFRx* Transcript Specifically Expressed in GSCs

To identify novel *EGFR* transcripts, we performed RNA-seq analysis on a GSC line (456) and a GBM cell line (U251). Exon coverage analysis revealed that *EGFR* reads ranging from exon 2 to exon 14 (E2–E14) were almost absent in 456 cells compared to those in U251 cells (Figure 1A), indicating the expression of a novel *EGFR* transcript (which we term *EGFRx*) lacking E2-E14 in 456 cells. This *EGFRx* transcript is distinct from EGFRvIII, which lacks exons E2-E7,[13–15] as confirmed by exon coverage analysis of an additional RNA-seq dataset (GSE51281) comprising three GSCs: BTSC68 and BTSC73 expressing EGFRvIII (EGFRvIII$^+$) and BTSC41 lacking EGFRvIII expression[26] (Supplementary Figure S1A).

The HISAT-StringTie-Ballgown pipeline allows the detection and quantification of novel transcripts from RNA-seq data.[24] Through this pipeline, we identified 12 *EGFR* transcripts, of which 10 transcripts matched the *EGFR* transcript ID in the Ensemble database (known transcript class code "="), while transcript 4 and 12 (novel class code "J") corresponded to the *EGFRvIII* and *EGFRx* transcript, respectively (Supplementary Figure S1B, Figure 1B).

The Ballgown quantification results revealed that the two *EGFRvIII*$^{+}$ GSCs (BTSC68 and BTSC73) predominantly expressed *EGFRvIII* (Figure 1B, Supplementary Figure S1C), whereas the 456 cells predominantly expressed *EGFRx* (Figure 1B–C).

To further investigate *EGFR* expression in GSCs, we analyzed a GEO RNA-seq dataset (GSE119834) (Supplementary Table S1). Among the examined 43 GSCs, we observed a positive correlation between *EGFRx* and *EGFRtotal* expression (Supplementary Figure S1D). The majority of cells exhibited higher expression levels of *EGFR* wild-type (*EGFRwt*) compared to *EGFRx* (Fig. 1D). However, we identified two GSCs (SRR7820965 and SRR7820944) expressed higher levels of *EGFRx* than *EGFRwt* (Figure 1D) and showed no detectable expression of *EGFRvIII* (Supplementary Figure S1E). In contrast, one GSC (SRR7820949) demonstrated elevated expression of *EGFRvIII* relative to *EGFRwt*, while exhibiting low levels of *EGFRx* expression (Supplementary Figure S1E–F).

We also analyzed the expression pattern in 9 neural stem cells (NSCs) from the GSE119834 dataset and 69 GBM cell lines (GBLs) from the Cancer Cell Line Encyclopedia (CCLE) (Supplementary Table S1). Our analysis revealed that all NSCs and GBLs exhibited low expression levels of *EGFRx* (Figure 1D). With the exception of DKMG cells, all NSCs and GBLs showed low expression of *EGFRvIII* (Supplementary Figure S1E–F). The high expression of *EGFRvIII* in DKMG cells corroborates previous findings.[27] Collectively, our findings identify a novel *EGFR* variant, *EGFRx*, distinct from EGFRvIII, which is specifically expressed in a subset of GSCs.

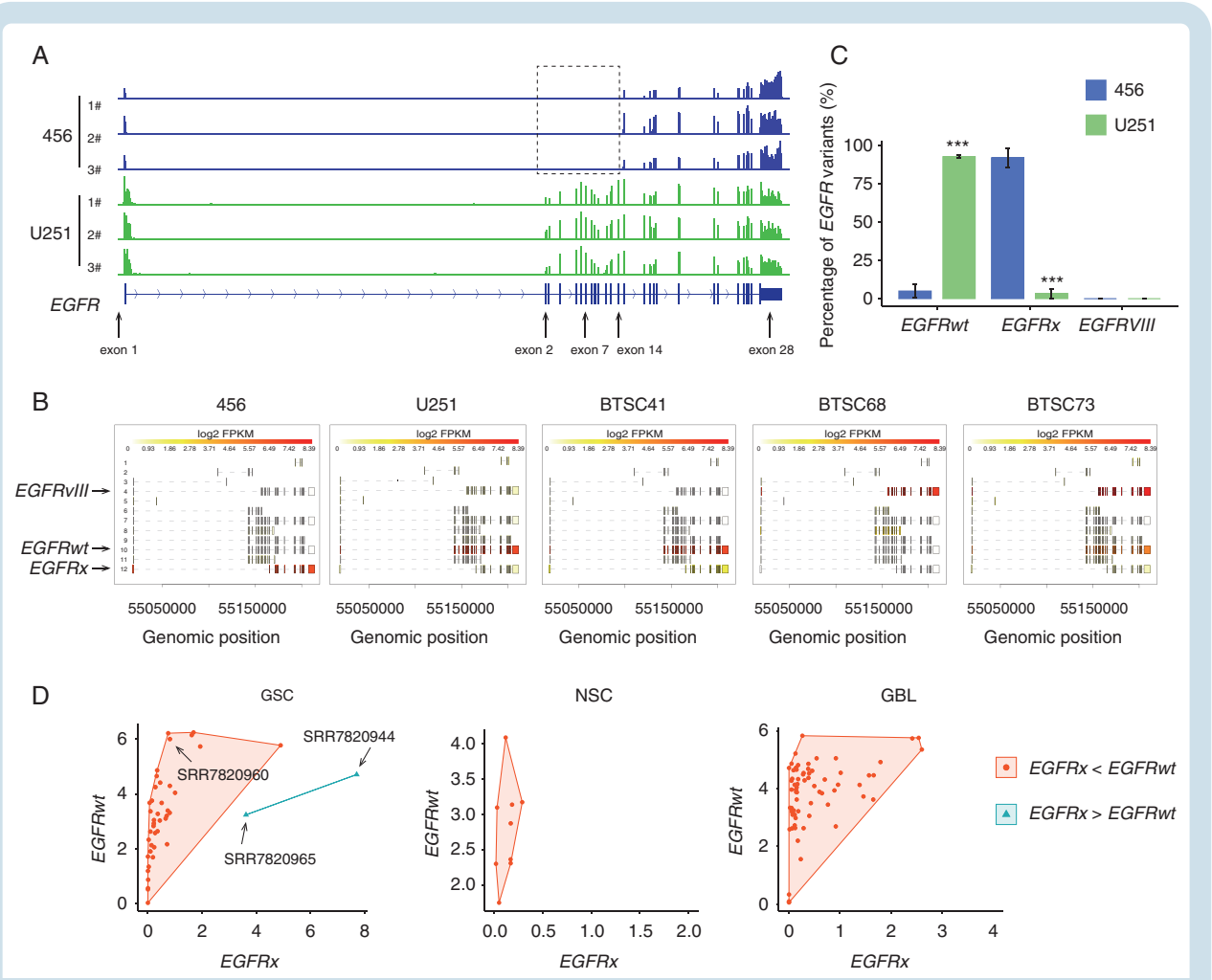

Figure 1 Identification of *EGFRx* transcripts in GSCs. (A) Visualization of the RNA-seq read profile through IGV. The vertical lines indicate the sequence reads. The boxes and horizontal lines indicate the *EGFR* exons and introns, respectively. The dashed box indicates the absence of exons 2-14 within the *EGFR* transcript in the 456 cells. Three replicates (1#, 2#, 3#) of RNA-seq reads derived from 456 and U251 cells are shown. (B) *EGFR* transcript detection in 456, U251, BTSC41, BTSC68, and BTSC73 cells. Twelve *EGFR* transcripts were identified through the HISAT-StringTie-Ballgown pipeline. The boxes and dotted horizontal lines indicate exons and introns in each transcript. Transcript 4, 10, and 12 represent the *EGFRvIII*, *EGFRwt*, and *EGFRx*, respectively. The expression of each transcript is presented as $\log_2$-transformed FPKM values. (C) Percentages *EGFR* transcripts in 456 and U251 cells. The percentages were calculated by dividing the transcript expression values by the sum of all *EGFR* transcript expression values. Data are presented as the mean ±SD from three replicate cells. *** $p < 0.001$ (Student's *t* test). (D) Grouping of GSC, GBL, and NSC based on the relative expression of *EGFRx and EGFRwt*. The expression of *EGFR* transcripts in the three indicated GSC lines were shown in Supplementary Figure S1E.

Validation of *EGFRx* Transcript in GSCs

The full-length EGFR protein is encoded by 28 exons, among which exons 1–15, exons 15–17, and exons 17–28 encode the ECD, TMD, and ICD, respectively[28] (Figure 2A). The above bioinformatic analysis indicates that *EGFRx* lacks exons 2-14 but contains the remaining exons (Figure 2A). We then performed the following experiments to validate the exon composition of *EGFRx* in 4 GSC lines (3691, 4121, 387, 456). These GSCs were derived from $CD133^+$ cells isolated from GBM surgical specimens or xenografts and their functional characterization has been previously described[5,29,30] and validated through the neurosphere-forming, differentiation, and tumor formation assays (Supplementary Figure S2), as described in the Supplementary Material and Methods.

First, we performed 5′ RACE with a reverse primer complementary to the exon 16-exon 17 junction (E16-17R) in 456 cells (Figure 2B). Through DNA sequencing of the 5′ RACE products, we detected a transcript fragment with the 5′ end sequence consisting of exon 1 and exons 15 but skipping exons 2-14 in 456 cells (Figure 2B).

Second, we assessed the presence of *EGFRx* through RT-PCR. The flanking primer sets (1#, 2#, 3#) are located within sequences common to *EGFRwt* and *EGFRx* and detect both isoforms simultaneously. RT-PCR showed that the small amplicon, indicative of *EGFRx,* was detected only in the GSCs, while the large amplicon, indicative of *EGFRwt,* was predominantly detected in U251 cells (Figure 2C–D). We designed *EGFRx* and *EGFRwt* isoform-specific primers (#4 and #5), which consisted of the same reverse primer E18R and different forward primers (E1-15F and E13F) (Figure 2C). The combination of E1-15F, E13F, and E18R (#4 + #5) allows the simultaneous detection of both isoforms. We found that the 454 bp amplicon, indicative of *EGFRx*, was specifically detected in GSCs, whereas the 662 bp amplicon, indicative of *EGFRwt*, was mainly detected in U251 cells (Figure 2D). Quantification of the isoform-specific amplicons further showed that *EGFRx* was specifically expressed in GSCs (Figure 2E).

Third, we performed ribonuclease protection assays using probes targeting exon 26 (E26), exon 12 (E12), and the exon 1-exon 15 junction (E1-15), respectively (Figure 2F). Binding of these probes protects the bound transcripts from ribonuclease degradation. We detected almost equal signals of the E26 probe, which indicates the presence of *EGFRtotal*, in these cells (Figure 2G). In contrast, our results showed that the E12 probe, which detects *EGFRwt*, yielded a strong signal in U251 cells, while the E1–15 probe, which detects *EGFRx*, produced a strong signal in GSCs (Figure 2G).

Finally, northern blotting revealed that the common E26 probe detected major large and small *EGFR* transcripts in U251 cells and GSCs, respectively, while the *EGFRx*-specific probe (E1–15) detected a major transcript in GSCs (Figure 2H). Taken together, these results demonstrate that *EGFRx* lacks exons 2–14 and is specifically expressed in GSCs.

Translation of the *EGFRx* Transcript

We detected moderate expression of the ~180 kDa EGFRwt in 3691 GSCs and the expression was hardly detected in other GSCs (4121, 387, and 456), while EGFRwt was highly expressed in U251 and other GBM cell lines (Figure 2I and Supplementary Figure S3A). Instead, smear binds with reduced molecular weights were specifically detected in GSCs by the EGFR (α-ICD) but not the α-ECD antibody (Fig. 2I). Immunofluorescent staining with the α-ICD antibody showed that EGFR was present in the cellular membranes of 456 and 387 GSCs (Figure 2J). These results suggest that EGFR proteins expressed in GSCs contain the ICD and TMD but not the ECD.

The smear binds may reflect posttranslational modification of EGFR in GSCs. Treatment with the deglycosylation enzyme PNGaseF reduced the molecular weight of the known glycosylated protein CD44 and shifted the smeared EGFR to a major 70 kDa bind (Figure 2K), while λ-phosphatase treatment did not alter EGFR migration (Supplementary Figure S3B), indicating that GSCs express a 70 kDa EGFR protein before glycosylation.

EGFRx transcript is predicted to encode a 70 kDa protein lacking the ECD, with a translation initiation codon in exon 15, equivalent to the 600^{th} Met (M600) within EGFRwt proteins (Figure 2L), which may generate the smear EGFR proteins observed in GSCs. To test this hypothesis, we designed an shRNA that targets E1–15 (Figure 2L). Strikingly, we observed significant silencing of both *EGFRx* mRNA and EGFR proteins, while *EGFRwt* expression remained unaffected in the sh-*EGFRx* expressing GSCs (Figure 2M–N). When we expressed the deduced CDS, we generated proteins with a similar migration and localization pattern to EGFR proteins in GSCs (Figure 2N, Supplementary Figure S3C–E). Importantly, we found that the M600 codon mutation (AUGmut) abolished the expression of exogenous EGFR proteins (Supplementary Figure S3C–D). These findings provide strong evidence that *EGFRx* encodes a heavily glycosylated truncation protein with initiation at M600.

Our analysis of RNA-seq data revealed that the 456 cells showed no detectable expression of *EGFRvIII* (Figure 1B–C). To further validate the presence of *EGFRvIII*, we performed RT-PCR using both the flanking (#3) and *EGFRvIII*-specific (#6) primer set, which yielded amplicons of 1148 bp and 893 bp, respectively, to indicate the presence of *EGFRvIII* (Supplementary Figure S3F). As anticipated, we identified the presence of *EGFRvIII* amplicons in the $EGFRvIII^+$ DKMG cells mentioned earlier (Supplementary Figure S1E–F) but not in any of the four examined GSCs (Supplementary Figure S3G). To further validate these findings, we employed an EGFRvIII-specific antibody (L8A4)[31] and verified its specificity in detecting the overexpressed EGFRvIII but not EGFRwt or EGFRx proteins (Supplementary Figure S3H). Consistent with the RT-PCR results, we detected EGFRvIII expression in DKMG cells, but not in the GSCs (Supplementary Figure S3I). Hence, EGFRvIII expression is absent in the four GSCs.

EGFRx is Required for GSC Self-renewal, Proliferation, and Tumorigenesis

We selected two GSC lines (456 and 4121) that predominantly express EGFRx, as demonstrated above to examine EGFRx function. EGFRx inhibition through either erlotinib

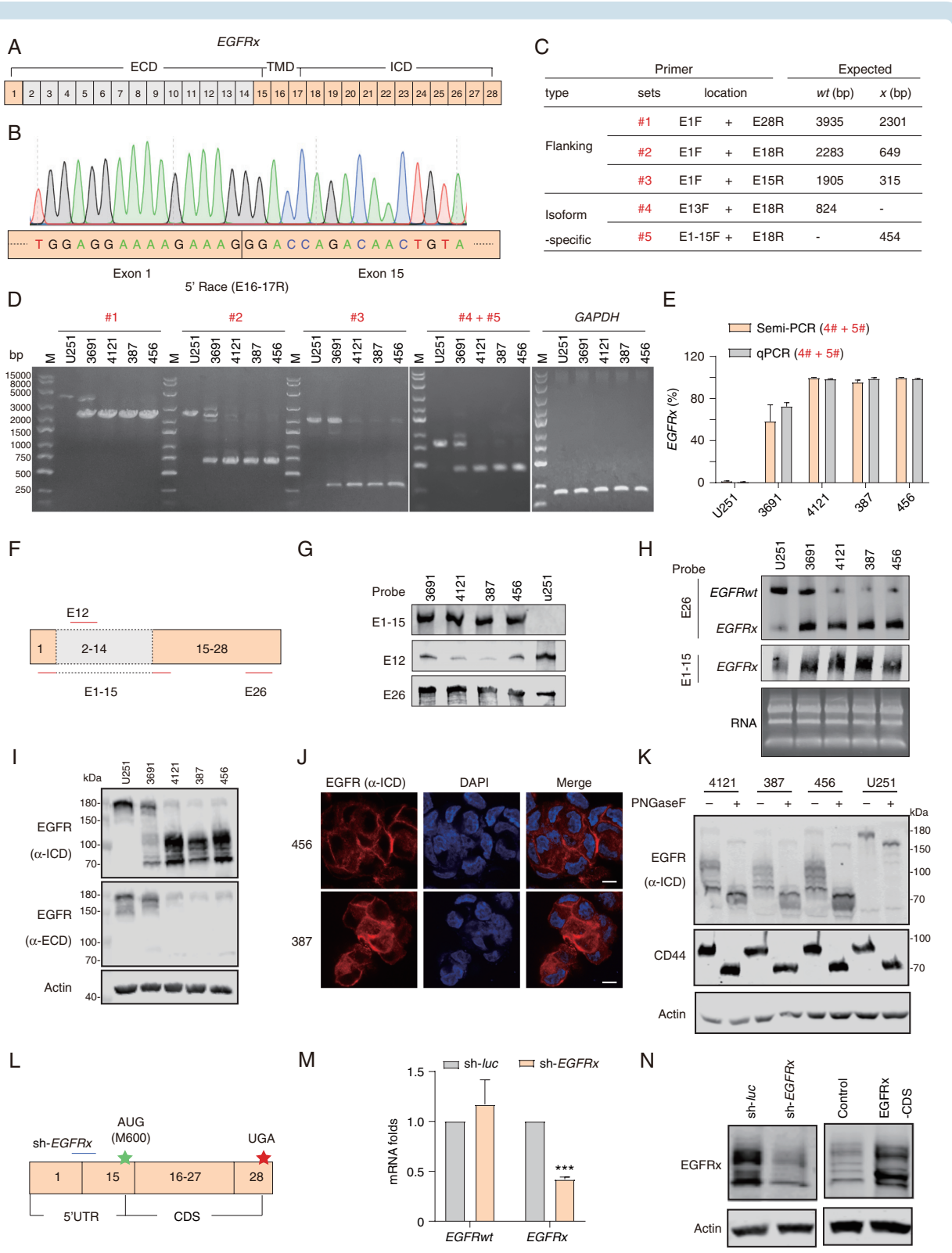

Figure 2 *EGFRx* transcript validation and the encoding proteins. (A) Diagram of *EGFR* gene structure. Exons 2–14 are absent in *EGFRx*, as analyzed in Figure 1. The exons encoding the ECD, TMD and ICD domains are shown above the exon boxes. (B) *EGFRx* transcript detection through 5' Race. A reverse primer complementary to the exon 16-exon 17 junction (E16-17R) was used. DNA sequencing of the products detected an *EGFR* transcript lacking exons 2–14. (C) The RT-PCR primer table used to detect *EGFR* transcripts. The location of each forward (F) and reverse (R) primer is indicated by the exon (E) number. For example, E1F indicates a forward primer located in exon 1. The expected sizes (bp, base pair) of amplicons for *EGFRwt* and *EGFRx* transcripts with each primer set are indicated; "-" indicates no expected amplicons. (D) *EGFR* transcript detection with primers in Panel (C). The DNA marker is shown on the left. (E) Quantification of *EGFRwt* and *EGFRx* transcripts.

Fig. 2 continued

The amplicon intensities through RT-PCR primers (#4 + #5) in Panel (D) were quantified. *EGFRwt* and *EGFRx* transcripts were also detected through q-PCR with primer sets #4 and #5, respectively. EGFRx percentages were calculated and presented as the means ± SD (n = 3). (F) Diagram of probes used for *EGFR* transcript detection in Panels (G) and (H). (G) Detection of *EGFR* transcript by ribonuclease protection assays. (H) Detection of *EGFR* transcript by northern blot. (I) Detection of EGFR proteins in GSCs. Cell lysates from U251 cells and GSC lines (3691, 4121, 387, and 456) were used for western blot detection with EGFR antibodies against the intracellular domain (α-ICD) or extracellular domain (α-ECD). Actin was used as a loading control. The protein markers and the respective molecular weights (kDa) are indicated on the left. (J) Subcellular localization of EGFR in GSCs. EGFR (α-ICD) was used for immunofluorescence staining in 456 and 387 cells. Scale bars, 10 μm. (K) Cell lysates were subjected to PNGaseF (25 units/μl) treatment, followed by western blot detection. (L) Diagram of *EGFRx* transcript and the deduced translation proteins. The predicted initiation and termination codons are marked by green and red pentagrams, respectively. The shRNA to knockdown *EGFRx* (sh-*EGFRx*) targeted the E1-E15 junction. (M) Examination of *EGFRx* knockdown efficiency by qPCR. Isoform-specific primer sets (#4 and #5) were used to detect *EGFRwt* and *EGFRx,* respectively. Data are presented as the mean ± SD ($n = 3$) from three independent experiments. *** $p < 0.001$ (Student's *t* test). (N) Detection of EGFR proteins in 456 cells expressing sh-*EGFRx* (left panel) or the deduced *EGFRx* CDS. The α-ICD EGFR antibody was used.

treatment or *EGFRx* silencing resulted in severe growth inhibition, as shown by a cell viability assay (Figure 3A–E). We also observed the growth inhibition of GSCs by another shRNA directed *EGFRx* silencing (Supplementary Figure S4A). The capacity to grow as neurosphere-like masses has been used widely as an in vitro indicator for defining the stemness of GSCs.[5] We found through a limiting dilution assay that EGFRx inhibition robustly prevented neurosphere formation (Figure 3F–H). These results demonstrate that EGFRx is required for GSC self-renewal and proliferation.

To assess the oncogenic role of EGFRx in vivo, we intracranially injected GSCs into immunocompromised mice. We found that the *EGFRx*-silenced group displayed a prolonged survival time (Figure 3I). In a parallel experiment, mice were sacrificed, and the brains were examined for tumor formation. H&E staining of the brain indicated that the tumor was markedly reduced in the *EGFRx*-silenced group (Figure 3I). When we overexpressed EGFRx in U251 cells, a GBM cell line that does not express EGFRx, we found that EGFRx overexpression reduced mouse survival and promoted tumor formation (Supplementary Figure S4B). These results demonstrate that EGFRx is oncogenic in vivo.

We then performed RNA-seq analysis to uncover the transcriptional regulation by EGFRx in GSCs. Principal component analysis (PCA) revealed a clear separation between the control and erlotinib-treated groups (Supplementary Figure S4C). We examined the DEGs by three differential expression methods (DESeq2, edgeR, and limma) and found that 1,852 and 517 genes were downregulated and upregulated, respectively, by EGFRx inhibition (Figure 3J, Supplementary Figure S4D). We selected 38 genes and verified that their expression was downregulated by EGFRx inhibition (Supplementary Figure S4E). The 1,852 downregulated genes were enriched in GO terms such as nuclear division and stem cell population maintenance (Figure 3K-3L, Supplementary Figure S4F), further underscoring the essential role of EGFRx in self-renewal and proliferation of GSCs.

EGFRx Functions Through STAT5

EGFR activates several downstream oncogenic effectors, such as PI3K-AKT, MAPK, and STAT signaling.[10] We observed that EGFRx inhibition by treatment with EGFR kinase inhibitors (erlotinib or gefitinib) abolished p-EGFR and p-STAT5 but did not significantly affect p-AKT, p-ERK1/2 or p-STAT3 (Figure 4A). We further demonstrated that p-STAT5 signaling was very sensitive to EGFRx inhibition, as shown by an inhibitor titration assay (Supplementary Figure S5A). Consistently, EGFRx silencing diminished p-STAT5 (Figure 4B), and EGFRx overexpression promoted p-STAT5 but not p-AKT, p-ERK1/2 or p-STAT3 (Figure 4C). The increase in p-STAT5 by EGFRx overexpression was prevented by erlotinib treatment (Figure 4C). In addition, we observed that treatment with a JAK kinase inhibitor (ruxolitinib or tofacitinib) abolished p-STAT3 but not p-STAT5 (Figure 4D). These results suggest that EGFRx promotes p-STAT5 in a manner dependent on its kinase activity and that STAT5 is a major downstream effector.

We observed that the knockdown effect by the shRNA in the two GSCs (456 and 4121) was specific to EGFRx, as EGFRwt expression was not altered (Figure 4B). To further elucidate the specific effects of EGFRx, we conducted knockdown experiments targeting EGFRwt and EGFRx individually in 3691 cells, which express EGFRx and relatively high levels of EGFRwt (Figure 2D–I). The results demonstrated that 3691 cells with *EGFRx* KD exhibited a diminished p-STAT5 signal (Supplementary Figure S5B), similar to the observations in 456 and 4121 cells (Figure 4B). However, *EGFRwt* KD in 3691 cells did not influence STAT5 signaling (Supplementary Figure S5B). Besides, we have verified the absence of EGFRvIII in these GSCs (Supplementary Figure S3F–I). These findings suggest that in these GSCs, EGFRx but not EGFRwt or EGFRvIII specifically promotes STAT5 signaling.

Active STAT5 moves into the nucleus for transcriptional activity upon phosphorylation.[32] We observed that EGFRx overexpression promoted the nuclear translocation of p-STAT5 in a kinase activity-dependent manner (Figure 4E, Supplementary Figure S5C). By using a STAT5-responsive reporter,[33] we found that EGFRx overexpression resulted in an ~13-fold increase in transcriptional activity (Figure 4F). GSEA analysis revealed a STAT5 target gene set (STAT5A_04) was suppressed in erlotinib-treated 456 cells (Figure 4G). To further evaluate the activity of STAT5 signaling, we conducted a CUT&Tag assay. Our results revealed that inhibition of EGFRx resulted in a decrease

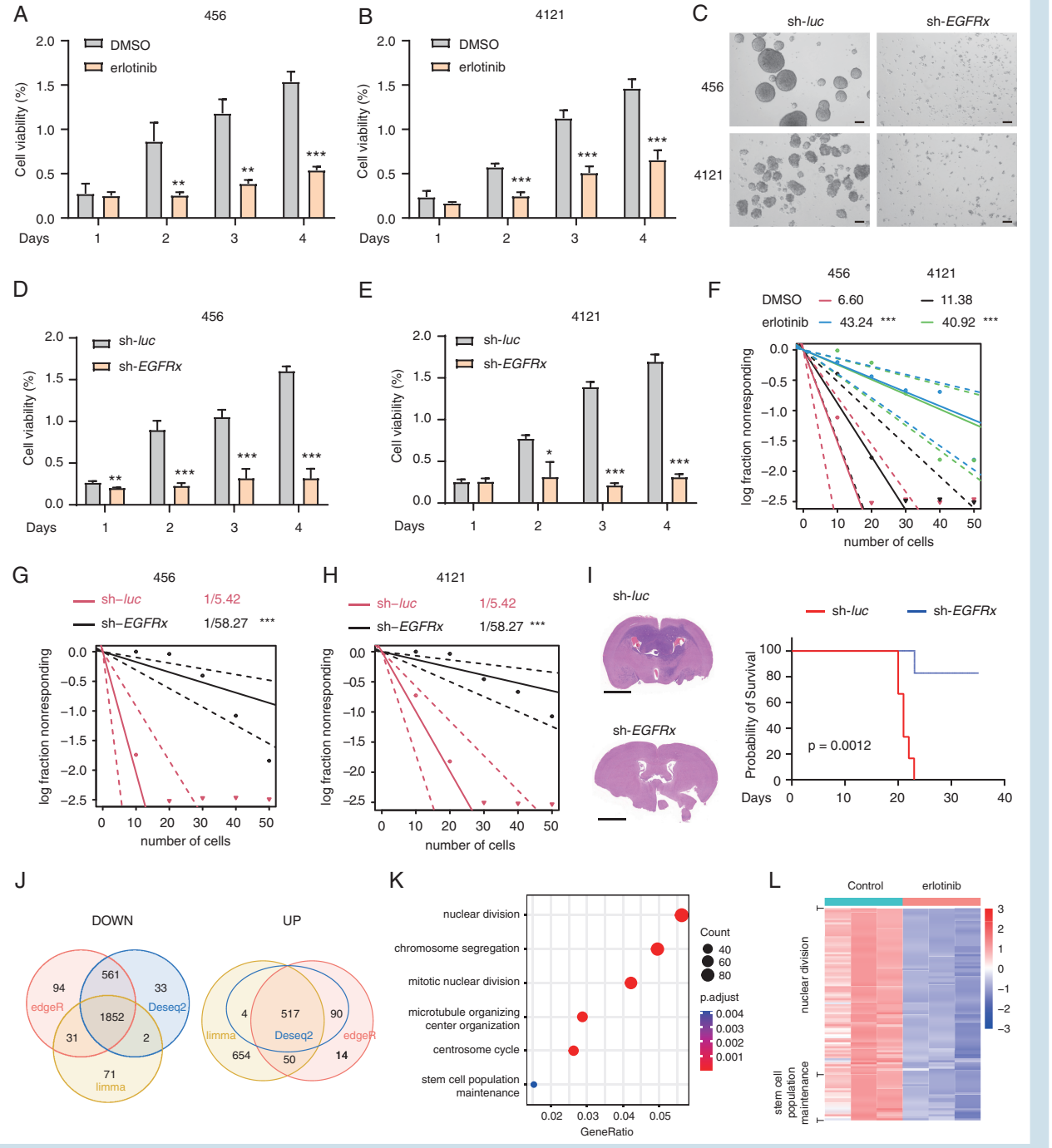

Figure 3 EGFRx is required for GSC self-renewal, proliferation and tumorigenesis (A–E) The effect of EGFRx inhibition on the viability of GSCs. 456 and 4121 cells were incubated with erlotinib (1 μM), and cell viability was measured on the indicated days (A-B). GSCs expressing sh-*luc* or sh-*EGFRx* were plated, and cell viability was measured on the indicated days (D–E). Cell images were acquired under a microscope (C). Scale bars (panel C), 20 μm. Data are presented as the mean ± SD from 3 replicates and are representative of 3 independent experiments. ** $p < 0.01$; *** $p < 0.001$ (Student's *t* test). (F–H) The effect of EGFRx inhibition on neurosphere formation of GSCs. In vitro limiting dilution assays were performed in GSCs treated with erlotinib (F) or GSCs expressing sh-*luc* or sh-*EGFRx* (G-H). *** $p < 0.001$ (Student's *t* test). (I) In vivo function of EGFRx. 456 cells expressing sh-*luc* or sh-*EGFRx* were intracranially injected into nude mice. Representative tumors are shown by H&E staining (left panel), and mouse survival status was monitored for Kaplan–Meier survival analysis (right panel). The *p* value is indicated (log-rank test). Scale bars, 200 μm. (J) Venn diagram plots showing the overlapping DEGs analyzed by limma, Deseq2, and edgeR. UP and DOWN indicate the upregulated and downregulated genes in erlotinib-treated 456 cells, respectively. (K) GO (biological process) analysis of the Down genes in Panel (J). (L) Heatmap depicting the DOWN genes enriched in nuclear division and stem cell population maintenance.

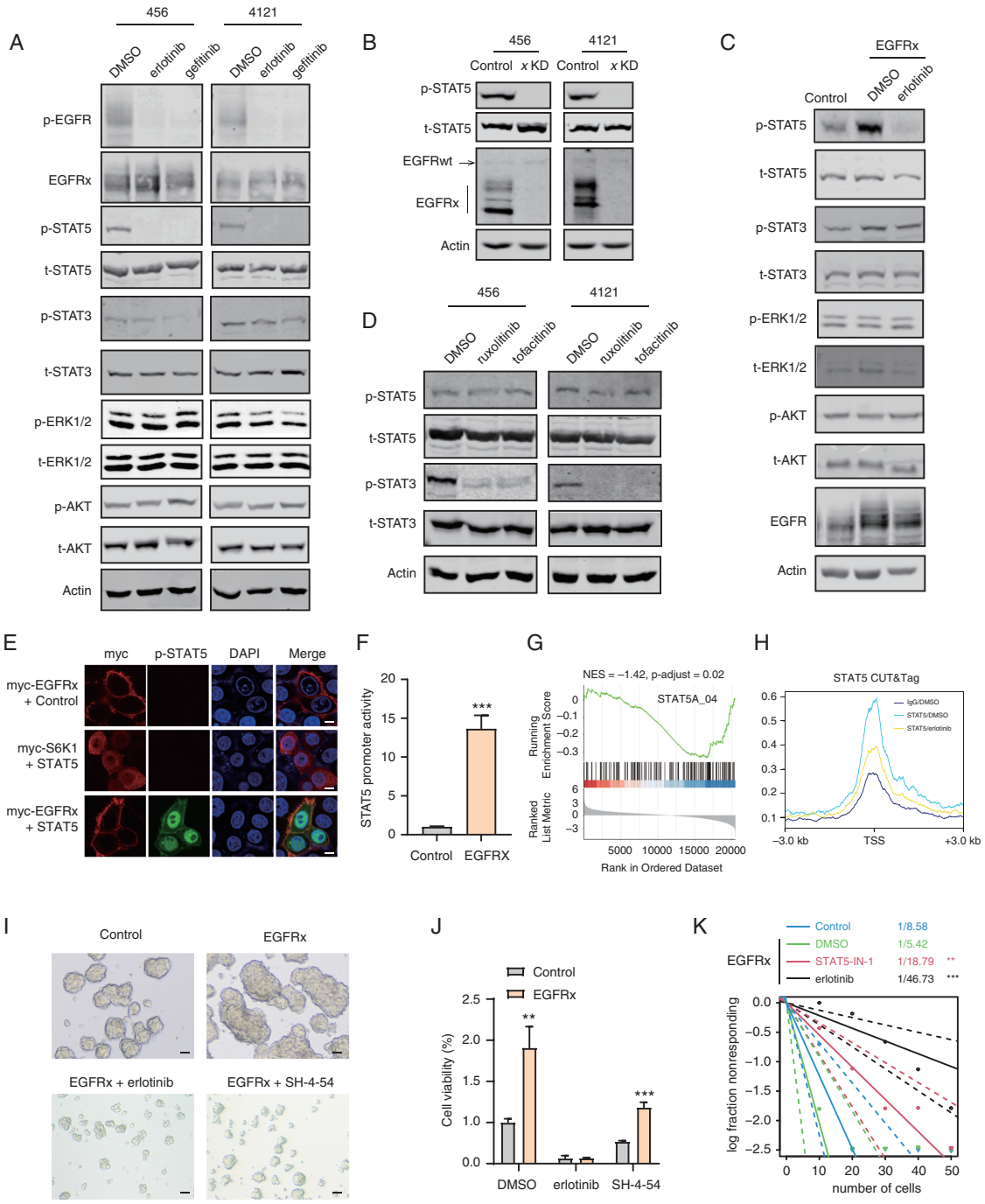

Figure 4 STAT5 is a major downstream effector of EGFRx signaling. (A) The effects of EGFR inhibitors on the signaling pathways. 456 and 4121 cells were treated with DMSO, erlotinib (5 μM, 1 hour) or gefitinib (5 μM, 1 hour). (B) p-STAT5 detection in control and *EGFRx* knockdown (*x* KD) GSCs. (C) Signaling pathways regulated by EGFRx overexpression in 456 cells. Cells were treated with DMSO or erlotinib (5 μM, 1 hour). (D) The effects of JAK inhibition on p-STAT3 and p-STAT5. Cells were treated with the JAK inhibitors ruxolitinib (5 μM, 1 hour) or tofacitinib (5 μM, 1 hour). (E)Nuclear translocation of p-STAT5 by EGFRx overexpression. HeLa cells were cotransfected with the myc-EGFRx and an empty control (Control) expressing constructs, myc-S6K1 and STAT5-FLAG, or myc-EGFRx and STAT5-FLAG expressing constructs, followed by costaining with myc and p-STAT5 antibodies. Scale bars, 10 μm. (F) The transcriptional activity of STAT5 by EGFRx overexpression. The β-casein promoter reporter and EGFRx overexpression constructs were cotransfected into HeLa cells, and luciferase activity was measured. Data are presented as the mean ± SD from three independent experiments. *** $p < 0.001$ (Student's *t* test). (G) A GSEA plot showing that the STAT5A_04 gene set was suppressed in the erlotinib treated 456 cells. (H) Average binding profiles of STAT5 around the TSS of the 1852 DOWN genes, as shown in Figure 3J within a ± 3-kb window in the control and erlotinib treated 456 cells. (I–K) EGFRx function depends on its kinase activity and STAT5 activity. 456 cells expressing control or EGFRx were treated with erlotinib (5 μM) or SH-4-54 (5 μM). Cell images were acquired under a microscope (I) and the cell viability (J) was measured 48 hours after the inhibitor treatment. Neurosphere number (K) were determined in the presence or absence of the inhibitors. Scale bars in panel (I), 20 μm.

in STAT5 binding in the vicinity of the transcription start sites (TSS) of the 1852 genes that showed downregulation upon erlotinib treatment (Figure 4H, Supplementary Figure S5D). These results collectively demonstrate that EGFRx activates STAT5.

We then examined whether EGFRx functions through STAT5 activation. EGFRx overexpression promoted cell viability and neurosphere formation of GSCs (Figure 4I–K). However, the enhancing effects were fully prevented by erlotinib treatment and partially reversed by treatment with the STAT5 inhibitor SH-4-54 (Figure 4I–K). These results suggest that the EGFRx function depends on its kinase activity as well as downstream STAT5 activity.

EGFRx Constitutively Phosphorylates STAT5 Through Spontaneous Asymmetric Dimerization of the Kinase Domain

The ECM domain has been shown to inhibit the formation of the activated kinase domain dimer prior to EGF binding (22, 23). The absence of ECM may render EGFRx constitutively active. We performed the following assays to investigate this hypothesis.

We investigated whether EGFRx was phosphorylated in response to EGF stimulation in GSCs. Our findings demonstrated that EGF promoted phosphorylation of EGFRwt but not EGFRx in the 456 cells (Figure 5A). Furthermore, treatment with erlotinib effectively inhibited the phosphorylation of both EGFRwt and EGFRx (Figure 5A). These results suggested that the lack of a ligand binding domain precludes the ability of ligand to activate EGFRx. To further explore the activation differences between EGFRwt and EGFRx, we overexpressed EGFRwt or EGFRx or coexpressed them in NIH3T3 cells, a cell line with low EGFR expression.[34] We found that EGFRx promoted p-STAT5 in an EGF-independent manner, as shown by the constitutive phosphorylation of EGFR and STAT5, while EGFRwt-STAT5 signaling depended on EGF availability, and coexpression had no additive effects (Figure 5B–C). Once EGFR is activated, STAT5 is recruited to the receptor.[32] Consistently, we observed a constitutive association of STAT5 with EGFRx but not EGFRwt (Figure 5D).

The above results suggested that EGFRx is constitutively activated and the activation does not depend on EGFRwt. To further demonstrate this, we utilized an EGFR kinase-dead (kd, D813N) mutant.[35] The equivalent EGFRx-kd mutant also lost the kinase activity and did not promote p-STAT5 (Figure 5E). EGFRwt overexpression failed to rescue the phosphorylation of EGFRx-kd and constitutive activation of STAT5 (Figure 5E). Likewise, EGFRx overexpression did not rescue the phosphorylation of EGFRwt-kd and EGF-induced p-STAT5 (Supplementary Figure S5E). These results suggest that EGFRx and EGFRwt act independently and that EGFRx does not depend on EGFRwt for STAT5 activation, which is distinct from the EGFRvIII mutant that has been reported to be a substrate for EGFRwt.[35]

Activation of the catalytic domain within EGFR requires allosteric and asymmetric dimerization mediated by the N-lobe and C-lobe faces of the two kinase domains.[8,9] Coexpressing the N-lobe mutant (I682Q) and C-lobe mutant (V924R) resulted in asymmetric dimerization and activation of EGFRwt-STAT5 signaling in response to EGF, while the expression alone did not (Supplementary Figure S5F), as previously reported.[8,9] We made equivalent N- and C-lobe EGFRx mutants, and their coexpression resulted in constitutive activation of EGFRx-STAT5 signaling, while expression alone did not (Figure 5F). These results demonstrate that EGFRx constitutively activates STAT5 through spontaneous asymmetric dimerization of the kinase domain.

EGFRx is Upregulated in GBM

We analyzed the expression of *EGFR* transcripts in a GBM dataset using the HISAT-StringTie-Ballgown pipeline. We observed elevated expression of *EGFRx*, *EGFRvIII*, *EGFRwt* and *EGFRtotal* compared to those in normal brain tissues (Figure 6A, Supplementary Figure S6A), consistent with previous findings that *EGFR* gene is frequently amplified in GBMs.[10] In general, *EGFRx* and *EGFRvIII* expression showed a positive correlation with that of *EGFRtotal* or *EGFRwt* (Figure 6B, Supplementary Figure S6B). We collected 59 GBM samples and examined *EGFRx* and *EGFRwt* expression through qPCR. Consistently, we detected *EGFRx* expression in GBMs, and its expression was positively correlated with that of *EGFRwt* (Figure 6C). *EGFRx* and *EGFRvIII* was enriched in the GBM classical subtype, similar to *EGFRwt* or *EGFRtotal*, as previously reported[11] (Supplementary Figure S6C–D).

Among the 81 GBM samples analyzed, we identified eleven samples (13.6%) that showed higher expression of *EGFRvIII* compared to *EGFRwt* (Supplementary Figure S6E). We only identified two samples that exhibited higher expression of *EGFRx* than *EGFRwt* (Figure 6D), suggesting that EGFRx is less frequently detected than EGFRvIII in GBM. The two GBMs with high levels of *EGFRx* demonstrated negative expression of *EGFRvIII* (Figure 6E), whereas the eleven GBMs with high levels of *EGFRvIII* exhibited low expression of *EGFRx* (Supplementary Figure S6F).

To further demonstrate the clinical significance, we examined cancer hallmark changes in GBMs with high *EGFRx* expression. To avoid the *EGFR* gene dose effect, we selected samples with high *EGFRtotal* expression and further divided them into *EGFRx*- high, median, and low groups (Supplementary Figure S6G). GSEA showed that patients with high *EGFRx* expression were enriched in several cancer hallmarks, such as HYPOXIA, epithelial-to-mesenchymal transition (EMT), and MYC_TARGETS_V1 (Figure 6F). These results further suggest that high EGFRx activity plays a role in maintaining GSC phenotype and contributes to the progression of GBM.

Our study has identified positive EGFRx expression in two GSC lines from the GSE119834 dataset, as well as in four GSCs (456, 3691, 387, 4121) (Figure 1, Figure 2). Given that CD133 is a well-known marker for GSCs[36] and all these cells were derived from the CD133$^+$ population,[5,29,30,37] we proceeded with flow cytometry to isolate CD133$^+$ cells from GBM samples. We confirmed that CD133 was highly expressed in the sorted CD133$^+$ cells (Figure 6G). We then performed RT-PCR using the flanking primer set (#3), which generated amplicons of distinct sizes to indicate the

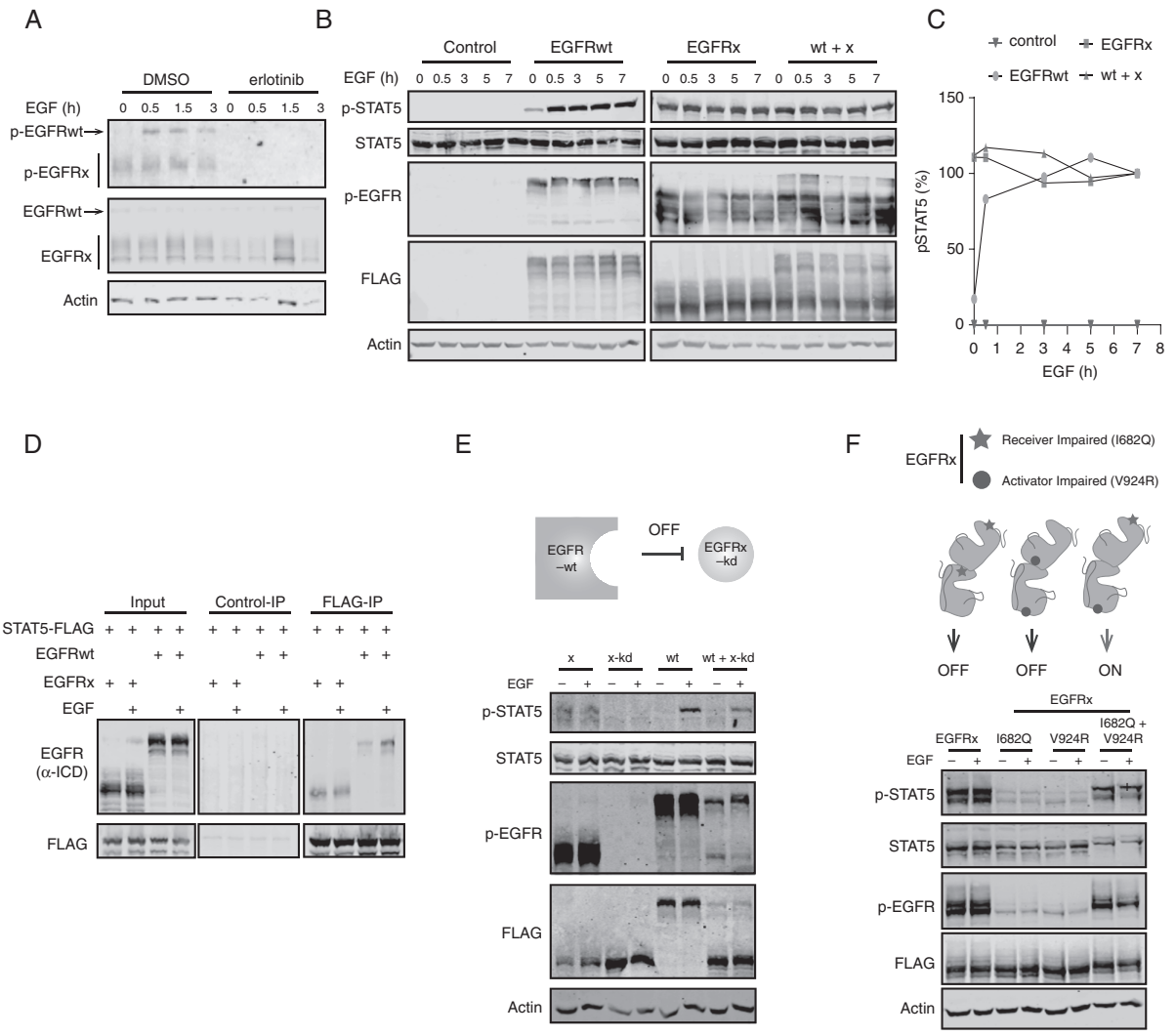

Figure 5 **EGFRx constitutively activates STAT5** (A) Detection of EGFR phosphorylation in GSCs. The 456 cells were starved in GSC growth media lacking EGF for 12 hours and stimulated with EGF (100 ng/ml) for the indicated times. (B, C) EGFRx promotes p-STAT5 independent of EGF availability. NIH3T3 cells transfected with the indicated expression constructs were serum starved and stimulated with EGF (100 ng/ml) for the indicated times. p-STAT5 signals relative to those of total STAT5 were calculated and plotted (C). The p-STAT5 value of the final time point (7 hours) in each group was set to 100%. (D) Interaction between STAT5 and EGFRx or EGFRwt. NIH3T3 cells transfected with the indicated constructs were immunoprecipitated (IP) with FLAG antibody or mouse IgG (Control), followed by EGFR detection. (E) The effect of EGFRwt overexpression on the EGFRx kinase dead mutant (x-kd). NIH3T3 cells transfected with the indicated constructs were serum starved and stimulated with EGF (100 ng/mL, 1 hour). The diagram (top panel) shows that EGFRwt overexpression did not rescue the activity of the EGFRx-kd mutant. (F) EGFRx-STAT5 signaling depends on the asymmetric dimerization of the kinase domain. NIH3T3 cells were transfected with the indicated EGFRx expression constructs. I682Q and V924R were the receiver-impaired and activator-impaired EGFRx mutants, respectively. Top panel: stars (I682Q) and circles (V924R) denote mutations in the N-lobe and C-lobe face of the dimer interface within EGFRx, respectively. Expression of either mutant alone does not activate EGFRx-STAT5 (OFF), while expression of both mutants does (ON).

presence of the *EGFRwt*, *EGFRx* and *EGFRvIII* transcripts (Figure 2C–D, Supplementary Figure S3F–G). The analysis revealed that CD133$^+$ cells from one GBM sample (GBM1) expressed higher levels of *EGFRx*, while the paired CD133$^-$ cells primarily expressed *EGFRwt*. Notably, both cell populations showed no expression of *EGFRvIII* (Figure 6G). In the other three samples (GBM2-4), we did not observe a significant change in *EGFRx* expression between the CD133$^+$ and CD133$^-$ cells (Figure 6G), which is consistent with our earlier findings that EGFRx is expressed in some, but not all, CD133$^+$ GSCs (Figure 1). These results further suggest that cells expressing EGFRx possess GSC features in vivo.

Discussion

EGFR gene alterations in GBMs have been detected in the ECD and the cytoplasmic tail. The known ECD truncation mutants include EGFRvI, EGFRvII, and EGFRvIII variants.[10]

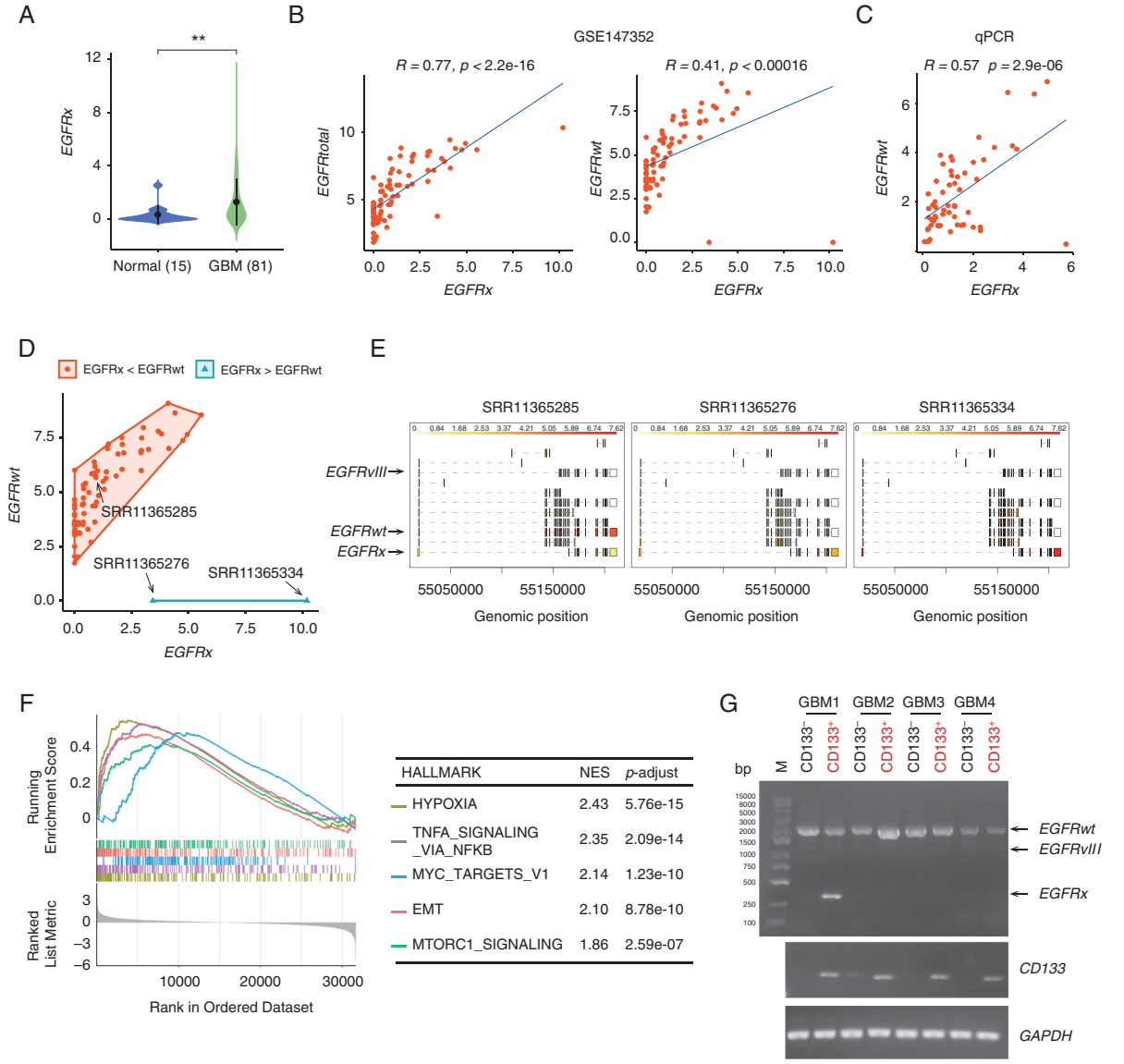

Figure 6 *EGFRx* expression in GBM. (A) Comparison of *EGFRx* expression in normal vs. GBM samples from the GSE147352 dataset. The number of samples analyzed is indicated within the brackets. Data are presented as the mean ± SD. ** $p < 0.01$; *** $p < 0.001$ (Wilcox test). (B) Correlation of *EGFRx* expression with that of *EGFRtotal* and *EGFRwt* in 81 GBMs from the GSE147352 dataset. Pearson correlation was used to calculate the R and *p* values. (C) Correlation of *EGFRx* expression with that of *EGFRwt* in 59 collected GBMs. The cDNAs of these samples were synthesized, and qPCR was performed to detect *EGFRwt* and *EGFRx* transcripts using primer sets #4 and #5, respectively. (D) Grouping of GBM samples based on the relative expression of *EGFRx and EGFRwt*. (E) Representative of *EGFRx* and *EGFRwt* transcript expression in samples from *EGFRx* high and low groups in panel (D). The boxes and dotted horizontal lines indicate exons and introns in each transcript. The expression of *EGFRx* and *EGFRwt* is presented as $\log_2$-transformed FPKM values. (F) GSEA plots revealed that several cancer hallmark signatures were enriched in the *EGFRx*-high GBM group compared to the *EGFRx*-low group as classified in Supplementary Figure S6G. (G) Detection of *EGFR* transcripts in $CD133^-$ and $CD133^+$ cells derived from 4 GBM samples (GBM1, GBM2, GBM3, GBM4). The flanking primer set (#3) was used to detect the *EGFRwt*, *EGFRx*, and *EGFRvIII* transcripts.

In this study, we used bioinformatic and experimental methods to identify EGFRx as another bona fide ECM mutant. Unlike EGFRvI, vII, and vIII with partial ECD deletion, EGFRx encodes a protein lacking the entire ECD. The absence of exons 2–14 suggests that EGFRx may be generated by aberrant splicing in GSCs. At this stage, it is not clear whether EGFRx is generated as a result of the cis splicing site mutation or trans alteration of splicing factor activities. Accumulating evidence has demonstrated that aberrant RNA splicing directs oncogenic gene expression in multiple types of cancers, including GBMs.[38–40] Future studies elucidating the EGFRx generation mechanism will provide novel insights into GBM pathogenesis.

Previous studies have reported the presence of EGFRvIII in GSCs.[26,41,42] In line with these findings, our HISAT-StringTie-Ballgown analysis revealed high levels

of EGFRvIII expression in certain GSCs. Interestingly, cells expressing high levels of EGFRvIII showed low levels of EGFRx expression, and vice versa. This expression pattern was also observed in GBM samples. In addition, among the four EGFRx-expressing GSCs (456, 3691, 387, and 4121), varying levels of EGFRwt expression were observed, with 3691 cells exhibiting relatively higher levels compared to the other three GSCs. Studies have demonstrated that EGFRvIII induces the expression and secretion of pigment epithelium-derived factor or its coreceptor, the cytokine receptor OSMR, via STAT3 activation, thereby promoting self-renewal and tumor progression of GSCs.[26,42] Our study provides compelling evidence that EGFRx functions independently of EGFRwt and activates the oncogenic STAT5 signaling pathway in GSCs. Besides, a circular RNA was found to encode the polymeric protein complex rtEGFR in GSCs. This complex sustains oncogenic EGFR signaling by interacting with EGFR and maintaining its membrane localization.[21] This distinct expression pattern and mechanism of EGFR variants in GSCs further emphasizes the significance of EGFR heterogeneity, as observed throughout GBM progression.[22,43,44]

Studies involving EGFRvIII$^+$ cells from primary tumors, using a specific EGFRvIII antibody, have identified the EGFRvIII$^+$ cells as the GSC population with the high degree of self-renewal and tumor-initiating ability.[41,45] Our studies suggest that EGFRx is co-expressed in the CD133 enriched GSCs and the EGFRx$^+$ cells may also represent a GSC population in vivo. However, it is important to highlight that this conclusion requires further validation, and in-depth investigation is necessary to gain a comprehensive understanding of the properties of this cell population. To address this issue, a straightforward way is to isolate and characterize the stem cell features of EGFRx$^+$ cells from primary cell cultures, as demonstrated for EGFRvIII$^+$.[41,45] However, unlike EGFRvIII, which encodes a partial ECD, EGFRx lacks the entire ECD, making it challenging to obtain a specific EGFRx antibody for isolating EGFRx$^+$ cells. Therefore, advancing our understanding of EGFRx's role in GBM progression requires the development of an efficient strategy to detect and/or isolate EGFRx$^+$ cells from GBM samples in future studies.

Amplification and overexpression of the *EGFR* gene is observed in >50% of GBM, and *EGFR* gene rearrangement is often accompanied by gene amplification, resulting in tumors expressing both EGFRwt and mutated EGFR.[46] The most common is the EGFRvIII mutant, with almost half of *EGFR*-amplified tumors harboring the mutant. In our study, the positive correlation between *EGFRx* expression and both *EGFRtotal* and *EGFRwt* expression suggest that *EGFRx* is also accompanied by *EGFR* gene amplification in GBM. However, EGFRx is less frequently detected than EGFRvIII in GBM. In a few GSC lines and even some GBMs, *EGFRx* was predominantly expressed. Our findings suggest that GBM cases with high EGFRx activity may rely on the activation of EGFRx-STAT5 signaling for pathogenesis. Despite the appealing prospect of targeting EGFR activation in GBM for therapeutic intervention, achieving therapeutic benefits has not been successful.[22] The failure may be attributed, in part, to the complexity of EGFR variants, as discussed earlier. Therefore, further studies are necessary to gain a better understanding of the activity and mechanism of these EGFR variants in GBM progression.

Supplementary Material

Supplementary material is available online at *Neuro-Oncology* (http://neuro-oncology.oxfordjournals.org/).

Keywords

EGFRx | GBM | stem cells | STAT5

Funding

This work was supported by the National Natural Science Foundation of China (Nos. 32070756 and 81903016).

Conflict of interests

The authors declare no potential conflict of interests.

Authorship

W.H. and J.L. conducted the experiments, acquired and analyzed the data. H.Z., X.Q., C.C., B.W., J.W., Y.S., X.L., Z.L., W.X., and A.H. assisted with the experiments. H.Z. and L.C. collected the GBM samples and assisted with the mouse injection. X.Y. helped with culturing of GSCs. W.W. did the bioinformatic analysis. K.S. and W.W. helped with the study design and wrote the manuscript. All authors read and approved the final manuscript.

Data Availability

The RNA-seq and CUT&Tag raw data have been deposited in Gene Expression Omnibus under the accession number GSE237135. All code used R 4.1.0 and publicly available packages cited in the paper. No custom functions were written for the analysis. Additional data used and/or analyzed will be made available upon reasonable request.

Affiliations

Department of Human Anatomy, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (W.H., X.Q., B.W., J.W., Y.S., X.L., Z.L., W.X., A.H., W.W.); Department of Integrated Traditional Chinese and Western Medicine, Tongji Hospital, Tongji Medical College,

Huazhong University of Science and Technology, Wuhan, China (J.L.); Department of Neurosurgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (H.Z., L.C., K.S.); Department of Orthopaedics, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (C.C.); Department of Histology and Embryology, School of Basic Medicine, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (X.Y.)

References

1. Stupp R, Hegi ME, Mason WP, et al; European Organisation for Research and Treatment of Cancer Brain Tumour and Radiation Oncology Groups. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 2009;10(5):459–466.
2. Stupp R, Mason WP, van den Bent MJ, et al; European Organisation for Research and Treatment of Cancer Brain Tumor and Radiotherapy Groups. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med.* 2005;352(10):987–996.
3. Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med.* 2008;359(5):492–507.
4. Lathia JD, Mack SC, Mulkearns-Hubert EE, Valentim CL, Rich JN. Cancer stem cells in glioblastoma. *Genes Dev.* 2015;29(12):1203–1217.
5. Bao S, Wu Q, McLendon RE, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature.* 2006;444(7120):756–760.
6. Dick JE. Stem cell concepts renew cancer research. *Blood.* 2008;112(13):4793–4807.
7. Jorissen RN, Walker F, Pouliot N, et al. Epidermal growth factor receptor: mechanisms of activation and signalling. *Exp Cell Res.* 2003;284(1):31–53.
8. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell.* 2006;125(6):1137–1149.
9. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol.* 2001;2(2):127–137.
10. Huang PH, Xu AM, White FM. Oncogenic EGFR signaling networks in glioma. *Sci Signal.* 2009;2(87):re6.
11. Wang Q, Hu B, Hu X, et al. Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell.* 2017;32(1):42–56.e6.
12. Neftel C, Laffy J, Filbin MG, et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell.* 2019;178(4):835–849.e21.
13. Wong AJ, Ruppert JM, Bigner SH, et al. Structural alterations of the epidermal growth factor receptor gene in human gliomas. *Proc Natl Acad Sci U S A.* 1992;89(7):2965–2969.
14. Frederick L, Eley G, Wang XY, James CD. Analysis of genomic rearrangements associated with EGRFvIII expression suggests involvement of Alu repeat elements. *Neuro Oncol.* 2000;2(3):159–163.
15. Frederick L, Wang XY, Eley G, James CD. Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. *Cancer Res.* 2000;60(5):1383–1387.
16. Nishikawa R, Ji XD, Harmon RC, et al. A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity. *Proc Natl Acad Sci U S A.* 1994;91(16):7727–7731.
17. Nagane M, Coufal F, Lin H, et al. A common mutant epidermal growth factor receptor confers enhanced tumorigenicity on human glioblastoma cells by increasing proliferation and reducing apoptosis. *Cancer Res.* 1996;56(21):5079–5086.
18. Humphrey PA, Gangarosa LM, Wong AJ, et al. Deletion-mutant epidermal growth factor receptor in human gliomas: effects of type II mutation on receptor function. *Biochem Biophys Res Commun.* 1991;178(3):1413–1420.
19. Lee JC, Vivanco I, Beroukhim R, et al. Epidermal growth factor receptor activation in glioblastoma through novel missense mutations in the extracellular domain. *PLoS Med.* 2006;3(12):e485.
20. Binder ZA, Thorne AH, Bakas S, et al. Epidermal growth factor receptor extracellular domain mutations in glioblastoma present opportunities for clinical imaging and therapeutic development. *Cancer Cell.* 2018;34(1):163–177.e7.
21. Liu Y, Li Z, Zhang M, et al. Rolling-translated EGFR variants sustain EGFR signaling and promote glioblastoma tumorigenicity. *Neuro Oncol.* 2021;23(5):743–756.
22. Eskilsson E, Rosland GV, Solecki G, et al. EGFR heterogeneity and implications for therapeutic intervention in glioblastoma. *Neuro Oncol.* 2018;20(6):743–752.
23. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods.* 2009;347(1-2):70–78.
24. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc.* 2016;11(9):1650–1667.
25. Kaya-Okur HS, Wu SJ, Codomo CA, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun.* 2019;10(1):1930.
26. Jahani-Asl A, Yin H, Soleimani VD, et al. Control of glioblastoma tumorigenesis by feed-forward cytokine signaling. *Nat Neurosci.* 2016;19(6):798–806.
27. Wlodarczyk A, Treda C, Rutkowska A, et al. Phenotypical flexibility of the EGFRvIII-positive glioblastoma cell line and the multidirectional influence of TGFbeta and EGF on these cells-EGFRvIII appears as a weak oncogene. *Int J Mol Sci* . 2022;23(20):12129.
28. Abou-Faycal C, Hatat AS, Gazzeri S, Eymin B. Splice variants of the RTK family: their role in tumour progression and response to targeted therapy. *Int J Mol Sci* . 2017;18(2):383.
29. Man J, Yu X, Huang H, et al. Hypoxic induction of vasorin regulates notch1 turnover to maintain glioma stem-like cells. *Cell Stem Cell.* 2018;22(1):104–118.e6.
30. Man J, Shoemake J, Zhou W, et al. Sema3C promotes the survival and tumorigenicity of glioma stem cells through Rac1 activation. *Cell Rep.* 2014;9(5):1812–1826.
31. Treda C, Wlodarczyk A, Pacholczyk M, et al. Increased EGFRvIII epitope accessibility after tyrosine kinase inhibitor treatment of glioblastoma cells creates more opportunities for immunotherapy. *Int J Mol Sci* . 2023;24(5):4350.
32. Quesnelle KM, Boehm AL, Grandis JR. STAT-mediated EGFR signaling in cancer. *J Cell Biochem.* 2007;102(2):311–319.
33. Yamashita H, Iwase H, Toyama T, Fujii Y. Naturally occurring dominant-negative Stat5 suppresses transcriptional activity of estrogen receptors and induces apoptosis in T47D breast cancer cells. *Oncogene.* 2003;22(11):1638–1652.
34. Bishayee A, Beguinot L, Bishayee S. Phosphorylation of tyrosine 992, 1068, and 1086 is required for conformational change of the human epidermal growth factor receptor c-terminal tail. *Mol Biol Cell.* 1999;10(3):525–536.
35. Fan QW, Cheng CK, Gustafson WC, et al. EGFR phosphorylates tumor-derived EGFRvIII driving STAT3/5 and progression in glioblastoma. *Cancer Cell.* 2013;24(4):438–449.

36. Wang Z, Zhang H, Xu S, Liu Z, Cheng Q. The adaptive transition of glioblastoma stem cells and its implications on treatments. *Signal Transduct Target Ther.* 2021;6(1):124.

37. Mack SC, Singh I, Wang X, et al. Chromatin landscapes reveal developmentally encoded transcriptional states that define human glioblastoma. *J Exp Med.* 2019;216(5):1071–1090.

38. Kim JH, Jeong K, Li J, et al. SON drives oncogenic RNA splicing in glioblastoma by regulating PTBP1/PTBP2 switching and RBFOX2 activity. *Nat Commun.* 2021;12(1):5551.

39. Song X, Wan X, Huang T, et al. SRSF3-regulated RNA alternative splicing promotes glioblastoma tumorigenicity by affecting multiple cellular processes. *Cancer Res.* 2019;79(20):5288–5301.

40. Sachamitr P, Ho JC, Ciamponi FE, et al. PRMT5 inhibition disrupts splicing and stemness in glioblastoma. *Nat Commun.* 2021;12(1):979.

41. Emlet DR, Gupta P, Holgado-Madruga M, et al. Targeting a glioblastoma cancer stem-cell population defined by EGF receptor variant III. *Cancer Res.* 2014;74(4):1238–1249.

42. Yin J, Park G, Kim TH, et al. Pigment Epithelium-Derived Factor (PEDF) expression induced by EGFRvIII promotes self-renewal and tumor progression of glioma stem cells. *PLoS Biol.* 2015;13(5):e1002152.

43. Eskilsson E, Rosland GV, Talasila KM, et al. EGFRvIII mutations can emerge as late and heterogenous events in glioblastoma development and promote angiogenesis through Src activation. *Neuro Oncol.* 2016;18(12):1644–1655.

44. Francis JM, Zhang CZ, Maire CL, et al. EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. *Cancer Discov.* 2014;4(8):956–971.

45. Liu XJ, Wu WT, Wu WH, et al. A minority subpopulation of CD133(+)/EGFRvIII(+)/EGFR(-) cells acquires stemness and contributes to gefitinib resistance. *CNS Neurosci Ther.* 2013;19(7):494–502.

46. Gan HK, Kaye AH, Luwor RB. The EGFRvIII variant in glioblastoma multiforme. *J Clin Neurosci.* 2009;16(6):748–754.