The *EGFRvIII* transcriptome in glioblastoma, a meta-omics analysis

Youri Hoogstrate, Santoesha A. Ghisai, Maurice de Wit, Iris de Heer, Kaspar Draaisma, Job van Riet, Harmen J.G. van de Werken, Vincent Bours, Jan Buter, Isabelle Vanden Bempt, Marica Eoli, Enrico Franceschi, Jean-Sebastien Frenel, Thierry Gorlia, Monique C. Hanse, Ann Hoeben, Melissa Kerkhof, Johan M. Kros, Sieger Leenstra, Giuseppe Lombardi, Slávka Lukacova, Pierre A. Robe, Juan M. Sepulveda, Walter Taal, Martin Taphoorn, René M. Vernhout, Annemiek M.E. Walenkamp, Colin Watts, Michael Weller, Filip Y.F. de Vos, Guido W. Jenster, Martin van den Bent, Pim J. French

Erasmus MC, Rotterdam, Netherlands (Y.H., S.A.G., M.d.W., I.d.H., J.v.R., H.J.G.v.W., J.M.K., S.Le, W.T., R.M.V., M.v.B., G.W.J., P.J.F.);
UMC Utrecht, Utrecht, Netherlands (K.D., P.A.R., F.Y.F.d.V.);
Université de Liège, Liège, Belgium (V.B., P.A.R.);
VU University Medical Center, Netherlands (J.B.);
University Hospitals Leuven, Belgium (I.v.B.);
Besta-IRCCS, Italy (M.E.);
Azienda USL/IRCCS Institute of Neurological Sciences, Italy (E.F.);
Institut de Cancérologie de l'Ouest, France (J.S.F.);
EORTC Headquarters, Belgium (T.G.);
Catharina Hospital, Netherlands (M.C.H.);
Maastricht UMC+, Netherlands (A.H.);
Haaglanden Medical Center, Netherlands (M.K., M.T.);
Veneto Institute of Oncology IOV-IRCCS, Italy (G.L.);
Aarhus University Hospital, Denmark (S.Lu);
Hospital Universitario 12 de Octubre, Spain (J.M.S.);
UMC Groningen, Netherlands (A.M.E.W.);
Institute of Cancer and Genomic Sciences, University of Birmingham, United Kingdom (C.W.);
University Hospital Zurich, Switzerland (M.W.);

Corresponding Author: Youri Hoogstrate, Department of Neurology, Erasmus MC, PO Box 2040, 3000CA, Rotterdam, The Netherlands (y.hoogstrate@erasmusmc.nl).

Authorship

Methodology: G.W.J., H.v.d.W., J.v.R, P.J.F., S.A.G, Y.H.; Analysis: I.d.H., K.D., M.d.W., Y.H.; Resources: A.H., A.M.E.W., C.W., E.F., F.Y.F.d.V., G.L., H.J.G.v.W, I.v.B, I.d.H., J.B., J.M.K., J.M.S., J.S.F, K.D., M.C.H., M.E.v.R., M.E., M.K., M.T., M.d.W., M.W., P.A.R., R.M.V., S.Le, S.Lu, T.G., V.B., W.T.; Drafting article: H.J.G.v.W, J.v.R., K.D., P.J.F., S.A.G.,Y.H.; Funding: M.v.B, P.J.F.

Abstract

Background. *EGFR* is among the genes most frequently altered in glioblastoma, with exons 2-7 deletions (*EGFRvIII*) being amongst its most common genomic mutations. There are conflicting reports about its prognostic role and it remains unclear whether and how it differs in signalling compared with wildtype *EGFR*.
Methods. To better understand the oncogenic role of *EGFRvIII*, we leveraged four large datasets into one large glioblastoma transcriptome dataset (n=741) alongside 81 whole-genome samples from two datasets.
Results. The *EGFRvIII/EGFR* expression ratios differ strongly between tumours and ranges from 1% to 95%. Interestingly, the slope of relative *EGFRvIII* expression is near-linear, which argues against a more positive selection pressure than *EGFR* wildtype. An absence of selection pressure is also suggested by the similar survival between *EGFRvIII* positive and negative glioblastoma patients. *EGFRvIII* levels are inversely correlated with pan-*EGFR* (all wildtype and mutant variants) expression, which indicates that *EGFRvIII* has a higher potency in downstream pathway activation. *EGFRvIII*-positive glioblastomas have a lower *CDK4* or *MDM2* amplification incidence than *EGFRvIII*-negative (p=0.007), which may point towards crosstalk between these pathways. *EGFRvIII*-expressing tumours have an upregulation of 'classical' subtype genes compared to those with *EGFR*-amplification only (p=3.873e-6). Genomic breakpoints of the *EGFRvIII* deletions have a preference towards the 3' end of the large intron-1. These preferred breakpoints preserve a cryptic exon resulting in a novel *EGFRvIII* variant and preserve an intronic enhancer.
Conclusions. These data provide deeper insights into the complex *EGFRvIII* biology and provide new insights for targeting *EGFRvIII* mutated tumours.

Keywords:
EGFRvIII, EGFR, RNA-seq, Breakpoints, glioblastoma

Importance of the Study

Glioblastoma is the most prevalent and aggressive form of malignant primary brain tumours, often characterized by EGFR mutations for which no effective treatments is available. We aimed to understand the role of its most common mutation, *EGFRvIII* (in-frame deletion of exons 2-7). By exploiting six combined datasets, we show the interplay between pan-*EGFR* and *EGFRvIII* levels, find an absence of positive selection towards *EGFRvIII* expression and demonstrate that *EGFRwt* and *EGFRvIII* largely activate similar pathways. However, significant and unique *EGFRvIII* mutation-specific associations were found with Cell Cycle (e.g. *CDK4*) and RTK/RAS/PI3K genes (e.g. *MDM2*) which provide new insights for tumour targeting. A preference in breakpoint location in intron-1 not only results in a distinct variant of *EGFRvIII* but also preserves an enhancer region, and so provides new insights into *EGFR(vIII)* gene regulation.

Key points:

- *CDK4* & *MDM2* amplifications appear less frequently in *EGFRvIII*+ compared with *EGFRvIII*- but *EGFR* amplified GBM
- Transcriptomes of *EGFR* amplified GBM differ marginally between *EGFRvIII*+ and *EGFRvIII*-
- *EGFRvIII* breakpoints preferentially retain an intronic enhancer

Introduction

Glioblastoma is the most prevalent and aggressive form of malignant primary brain tumours in adults with a short median survival time of 14.6 months[1]. Extensive research on the genetic makeup of glioblastoma has revealed recurrent genetic changes typically involving the RTK/RAS/PI3K, p53 and RB signalling pathways[2–4]. Although the diverse genetic features of glioblastoma have become increasingly better understood, no effective treatment options are currently available that specifically target the most common mutations. One of the most frequently altered genes in glioblastoma encodes the epidermal growth factor receptor (*EGFR*). *EGFR* is amplified in ~50% of all glioblastomas[4–7], typically within small circular extrachromosomal DNA copies (ecDNA)[8]. The most common mutation on top of this amplification is an in-frame deletion of exons 2-7 (*EGFRvIII*), found in ~50% of the *EGFR* amplified glioblastoma patients[9]. EGFRvIII is a constitutively, but low-level, active form of EGFR that is independent of ligand for its activation[9], likely due to the partially deleted extracellular ligand-binding receptor domain. The *EGFRvIII* variant results from a genomic deletion, not from alternative or aberrant splicing. Unfortunately, treatments aimed at targeting *EGFRvIII* have thus far not provided clinical benefit to patients[10,11].

It is assumed that *EGFRvIII* typically is a late event that arises after chromosome 7 amplification and after *EGFR* high-copy amplifications and is therefore considered subclonal[12]. However, even as subclonal mutation, it is highly prevalent in glioblastoma and contributes to and alters the biology of the tumour. *EGFRvIII* has been shown to reduce apoptosis and increase proliferation and invasiveness[9], key features of tumour progression. Protein levels of EGFRvIII vary widely across and spatially within glioblastoma tumours[13–15]. Moreover, recent observations show changes in *EGFRvIII* levels during tumour evolution after initial resection[6,16,17], including cases with complete loss of *EGFRvIII* over time. That such a common presumed driver mutation gets lost, or levels get reduced during tumour evolution is paradoxical and will complicate targeting it for clinical benefit.

In this study, we aim to unravel *EGFRvIII* specific mechanisms related to glioblastoma tumorigenesis. We examined *EGFRvIII* expression, its genomic breakpoints and co-occurrence with other genetic changes using a large combined dataset.

Methods

Sequencing data

Sequencing of the Intellance-2[18] (paired-end; 2x151bp total RNA + paired-end; 2x76bp TruSight Tumor 170 panel) and BELOB (single-end; 50bp)[19] data were described elsewhere. For G-SAM, RNA extraction was performed using the AllPrep DNA/RNA FFPE kit or the RNeasy FFPE kit (Qiagen, Venlo, The Netherlands). G-SAM samples were sequenced (150bp paired-end reads**)** on the Illumina NovaSeq at the GIGA-Genomics Core Facility University of Liège. Each of these datasets were non-poly(A)$^+$-enriched and thus also include non-polyadenylated transcripts[20]. Raw sequencing data is available (BELOB: EGAS00001004570, Intellance-2: EGAS00001005437; G-SAM: EGAS00001005436). TCGA-GBM (poly(A)$^+$ RNA and DNA mutations) and CPCT-02 and PCAWG DNA data were obtained from their public repositories.

Human specimens

Tissue and metadata from the G-SAM and Intellance-2 studies were accrued through the pan-European European Organisation for Research and Treatment of Cancer network[6,18]. Informed written consent was obtained from all patients. The study design was approved by the institutional review board of Erasmus MC (Rotterdam, the Netherlands), and conducted according to institutional and national regulations.

Sequencing data processing

For each RNA-seq sample, FASTQ files were cleaned using *fastp* (https://github.com/OpenGene/fastp), aligned to hg19 using STAR[21] and then de-duplicated with *sambamba*. For the Dr. Disco[20] pipeline, samples were first FASTQ-level de-duplicated level using *HTStream deduper* (https://github.com/ibest/HTStream). *EGFRvIII* and *EGFRwt* expression was estimated directly from BAM files using junction-reads (https://github.com/yhoogstrate/egfr-v3-determiner v0.7.4: `--spliced-reads-only`). Reads considered *EGFRvIII* spanned the splice junction of exons 1–8, and reads considered *EGFRwt* exons 1–2. Samples with <10 such reads were excluded, except for TCGA-GBM, where *EGFRwt* read counts for *EGFRvIII* negative samples were missing. Junction read counts of replicated samples were merged by summing the spliced read counts. The *EGFRvIII* percentage was defined as the average percentage from different assays were present (Intellance-2). Gene level read-counts were obtained using featureCounts and Gencode v31. *EGFRvIII* counts from TCGA-GBM were taken from elsewhere[4]. Junction-counts involving non-canonical exons A, B and C were determined using *egfr-v3-determiner* with modified exon annotations. Genomic events were taken from processed WES data or public resources (Supplementary Methods).

Expression analysis

Samples with an *EGFRwt+EGFRvIII* read count ≥10 were eligible for *EGFRvIII* status and percentage determination and for DE analysis. For DE analysis, only genes with on average ≥3 reads per sample were included. Only genes marked as "*protein_coding*" were included. DE analysis was performed using DESeq2 (Wald test)[22], in which *EGFRvIII* was excluded in *estimateSizeFactors* to avoid redundant counts. The FDR adjusted p-value reflects the q-value. For the tests with four datasets combined, the intersected protein-coding genes with on average ≥3 reads per sample, per dataset, were included. Normalised expression levels were estimated using DESeq2 followed by the VST transformation (blind=TRUE) to ensure homoscedasticity[22]. A batch correction was performed for DE and for correlation analysis to correct per-dataset differences (DESeq2 for count data; limma::removeBatchEffect[23] for VST transformed data). Volcano plots were generated with the *EnhancedVolcano* package (https://github.com/kevinblighe/EnhancedVolcano). Kaplan–Meier analysis was performed using R's *survival* package. Survival analysis on *EGFRvIII* expression was performed with a Cox Proportional Hazard survival using R's *survival* package on the normalised VST transformed expression values. Because the Depatux-M antibody binds *EGFRvIII* with high affinity[24] and the Intellance-2 trial reported a benefit from Depatux-M in *EGFRvIII* positive samples[18], Depatux-M arms were excluded from survival analysis.

Breakpoint analyses

Non-poly(A)$^{+}$-enriched RNA-seq samples include relatively large proportions of intronic reads derived from actively transcribed pre-mRNA. This allows detection of genomic breakpoints when corresponding introns are sufficiently covered[19,20]. Settings for Chimeric alignment are given in **Supplementary Methods**.

The 100-vertebrates-*phastCons* track was obtained from UCSC and smoothened by a running mean of 200bp fixed windows. H3K27ac Chip-Seq data was obtained from GSM3382305[25], GSM3670052, GSM3670055 and GSM3670058[26]. Actual genomic enhancer locations were not provided in the original manuscript[25]. Their raw CRISPRi-assay data (GSM4141363 + GSM4141364) was used to reproduce their findings according to their described methodology (**Supplementary Table 3**).

Exon-B variant experiments

To confirm the *EGFRvIII* exon-B variant, ten samples positive for the variant (RNA-seq) with remaining isolated RNA leftover from sequencing were chosen (**Supplementary Table 2**). cDNA was synthesized in a buffer of 1µl random primers, 1µl dNTP mix, 1000ng RNA and 13µl dH2O. The mixture was heated to 65°C for 5 minutes and incubated on ice for one minute. After brief centrifugation, the contents were collected, and the following was added: 4µl 5X First-Strand buffer, 1µl 0,1M DTT, 0.5µl RNaseOUT, 1µl Superscript III. The reaction was incubated at 25°C for 5 minutes, at 50°C for 45 minutes and inactivated at 70°C for 15 minutes. Partial sequences spanning the exon-B splice junction were PCR-amplified using 6 primer combinations (2x exon-B, 1x exon-8, 2x exon-9). For each reaction, the buffer consisted of: 7.9µl nuclease-free water, 3µl 5x GoTaq buffer, 0.8µl 10 mM dNTPs, 1µl 10µM forward primer, 1µl 10µM reverse primer, 1µl cDNA and 0.3µl GoTaq polymerase. Denaturation of cDNA was performed at 98°C for 30 seconds, followed by 40 cycles of 30 seconds at 98°C, 30 seconds at 60°C and 30 seconds at 72°C. The final extension was performed at 72°C for 5 minutes and brought back to 12°C. Of the ten samples, six showed bands of the expected size on agarose gel. Of these six, four were sent out for Sanger Sequencing to Macrogen Europe B.V., Amsterdam (**Supplementary Table 2)**. Three of four samples showed good per-base quality. Forward and reverse reads were assembled into consensus contigs using UGENE.

Constructs were generated to evaluate the function of EGFR variants initiating from exon B. Because exon-B lacks a translation initiation site, we generated these constructs using the first in-frame ATG in exon-2 or exon-8 (in the case of *EGFRvIII*). A total of sixteen different constructs were made: those that initiated translation in exons-2 or 8 with i) either an in-frame eGFP (located C-terminal to the transmembrane region [27]) or eGFP co-expressed via an IRES sequence; ii) with and without the L858R activating mutation (to compare the activation state of the novel isoforms with a constitutively active isoform) and; iii) without/with a canonical Kozak sequence (to ensure optimal translation of the latter). Constructs were generated by in-fusion cloning into a piggyback vector. Constructs were stably transfected in HeLa cells, imaged using an Opera Phenix (Perkin Elmer, Hamburg, Germany) high content imager and analysed using Harmony software (Perkin Elmer, Hamburg, Germany).

Results

We have collected glioblastoma data from the following cohorts: BELOB[19], Intellance-2[18], G-SAM[6], TCGA-GBM[4], CPCT-02[28] and PCAWG[29]. The compiled results are available as a study dataset: *https://zenodo.org/record/4792445*.

Molecular differences of EGFRvIII expressing tumours

To determine *EGFRwt* (spliced across exons 1–2) and *EGFRvIII* (spliced across exons 1–8) expression, we first developed *egfr-v3-determiner* (publicly available, see Methods). Out of the 839 available RNA-seq samples, we included samples with a combined *EGFR* junction read count (spliced across exons 1–2 and 1–8) of ≥10 into our combined study RNA dataset: n=**741** from 622 patients; BELOB (n=**69**/92), Intellance-2 (n=**224**/239), G-SAM (n=**285**/345) complemented with all primary TCGA-GBM samples (n=**163**). In this combined dataset, 464/741 (62.6%) samples had *EGFR* gene amplification or upregulation if copy-number data was absent. Using the transcript-specific junction-counts, we calculated the ratio *EGFRvIII* ($\frac{count\ EGFRvIII}{count\ EGFRvIII+EGFRwt}$). Of the *EGFR* amplified samples, 225/464 (48.5%) were considered *EGFRvIII* expressing ($\frac{count\ EGFRvIII}{count\ EGFRvIII+EGFRwt} \geq 1\%$), consistent with observations in literature[9,30]. These ratios revealed a high dynamic from 1% to 95%, consistent in all datasets (**Figure 1**). Lower *EGFRvIII* expression ratios were slightly over represented (1%-10%; $p=3.2e^{-9}$; Wilcoxon test on the first derivative of the ordered percentages). The total *EGFR* expression levels are on average lower for samples with higher *EGFRvIII* percentages, implying that *EGFRvIII* is more potent in *EGFR* signalling (**Figure 2,S1D**).

Several reports have indicated that *EGFRvIII* and *EGFRwt* activate different signal transduction pathways[9,31]. To assess if such differences are reflected in their transcriptomes, we performed differential gene expression (DE) analysis comparing the transcriptomes of *EGFRvIII* positive (using two cut-offs: ≥1.0% or ≥10.0%) with *EGFRvIII* negative (<1.0%) but *EGFR* amplified tumours. Tests were performed for all four datasets separately, to correlate the logarithmic fold changes (LFC) of the genes between the datasets. Markedly higher LFC correlations were found across the datasets using ≥10% *EGFRvIII* as cut-off (**Figure S2**), which suggests lower percentages (1%-10%) harbour a limited *EGFRvIII* response signal.

We therefore proceeded with the combined dataset using only ≥10% *EGFRvIII* as cut-off (n=368) and found 213 genes significantly (q-value<0.01, |LFC|>0.5) differentially expressed (**Figure 3A**) They showed enrichment in genes related to microtubule, cilium and axoneme related pathways (**Figure S3**).

The 187 significantly downregulated genes in ≥10% *EGFRvIII* included *CDK4* and *MDM2*, genes that are frequently hyper- and co-amplified in glioblastoma. Their observed differences were not a result of consistent down-regulation of *CDK4* or *MDM2* across all ≥10% *EGFRvIII* positive patients but were caused by a lower proportion of tumours with extremely high *CDK4* or *MDM2* expression levels (**Figure S4A-B**). Integration with copy-number data confirmed the negative association: *CDK4* or *MDM2* DNA amplifications appeared in significantly fewer tumours expressing *EGFRvIII* (p=0.007, Fisher's exact test, **Figure S4C**). *TP53* mutations were indeed[32] less frequently present in *EGFR* amplified tumours, both *EGFRvIII* positive and negative (**Figure 1, S1**). Similarly, *TACC3-FGFR3* fusions were indeed[33] exclusively present in *EGFR* non-amplified tumours. The overall

transcriptome differences did not show a strong separation between *EGFRvIII* positive and negative tumours, indicating the overall differences are modest (**Figure 3B**).

Glioblastomas are classified into three transcriptional subtypes: mesenchymal, proneural and classical. Classification is based on genes that are exclusively up-regulated within their subtype[34,35]. The classical subtype is characterized by *EGFR* amplifications[35]. We observed that almost all classical subtype genes tend to be upregulated in *EGFRvIII* positive tumours ($p=3.873e^{-6}$; Fisher's exact test on positive/negative LFC, **Figure 3A**) compared with *EGFRvIII* negative tumours, all harbouring *EGFR* amplifications. The classical subtype, therefore, is at least partly defined by *EGFRvIII*-specific signalling. While certain neuronal precursor and stem cell marker, sonic hedgehog pathway and notch pathway member genes are highly expressed in the classical subtype[36], these individual pathways did not differ across *EGFRvIII* positive/negative tumours (**Figure S5A-C**). According to a pathway based glioblastoma classification[37], two subtypes, proliferative/progenitor (PPR) and mitochondrial (MTC), are associated with RTK pathway amplifications such as *EGFR* and *PDGFRA*. Of these, PPR is associates positively with *EGFRvIII* (**Figure S5D-E**).

In addition to the DE analysis using a defined *EGFRvIII* expression cut-off, we interrogated the linear correlation between the expression of all genes to the *EGFRvIII* expression. This analysis was performed within the same ≥10% *EGFRvIII* positive samples. *CDK4* and *MDM2* expression levels did not linearly correlate with *EGFRvIII* expression. That there is a significant difference in *CDK4* and *MDM2* expression levels across *EGFRvIII* positive and negative tumours while their expression levels do not correlate with *EGFRvIII*, is in concordance with the difference in hyper-amplification incidence. To identify genes that correlate differently between *EGFRvIII* and *EGFR*, we performed the same test against *EGFRwt* (**Figure 3C**). We then calculated per gene to what extent the correlation with *EGFRwt* and *EGFRvIII* differs, and tested which differences were beyond what may be expected by chance (**Supplementary Methods**). This revealed 6 additional genes that significantly differ in their correlation to *EGFRwt* in contrast to *EGFRvIII* (*NSG1*, *GALNT15*, *RFWD3*, *NCAPD3*, *ARHGEF26* and *PHF19*; $q < 0.01$). *RFWD3* was positively correlated with *EGFRvIII* (coef=0.33) while negatively correlated with *EGFRwt* (coef=-0.20). Similar to using a defined *EGFRvIII* expression cut-off, we found that the classical subtype genes correlate positively stronger with *EGFRvIII* compared with *EGFRwt* ($p=1.1e^{-9}$; two-sided t-test on Z-score difference).

The difference in correlation between *EGFRvIII* and *EGFRwt* and the difference in gene expression by *EGFRvIII* presence, showed correlation (Spearman coef=0.4, **Figure S6**). For classical subtype genes this correlation was stronger (Spearman coef=0.7), indicating consistency in the outcome of the tests. In particular, genes that showed strong concordant results were *PHF19*, *NSG1* and Sprouty/Spred family members *SPRED2*, *SPRY4* and *SPRY2* (**Figure S6B**). Furthermore, *PTPRZ1,* occasionally found in glioma as donor partner in fusions such as *PTPRZ1-ETV1* and *PTPRZ1-MET*[38], positively associates with *EGFRvIII*.

EGFRvIII prognostic value

There have been conflicting data on the association of *EGFRvIII* with prognosis[9]. We interrogated the patient survival between *EGFRvIII* positive and negative patients in the BELOB, G-SAM and TCGA-GBM and Intellance-2 (control arm) datasets. Within patients with *EGFR* amplified tumours, there was no significant difference in overall survival between patients with *EGFRvIII* positive and negative tumours (n=327) in each dataset or combined

(**Figure 4, S7**). There was no significant association between relative *EGFRvIII* expression levels and patient survival (HR: BELOB=1.1, G-SAM=0.96, Intellance-2=1.2). In summary, we found no evidence for an association of *EGFRvIII* with survival in patients with *EGFR* amplified tumours.

EGFRvIII breakpoints preferentially retain intronic enhancer

With a transcription rate of 1-6 kb/min[39], transcription of the ~120 kb *EGFR* intron-1 can take up to two hours. The closer the breakpoint of the causal *EGFRvIII* deletion is to exon-1, the shorter its intron. Given the large size of intron-1, breakpoints at the beginning of the intron (early breakpoints) may provide an energetic and temporal benefit over breakpoints at the end of the intron (late breakpoints). We screened ≥1% *EGFRvIII* positive samples for their genomic *EGFRvIII* breakpoints based on the presence of pre-mRNA[19,20]. We found 44 breakpoints within our datasets (**Supplementary Table 1; Figure 5**). One sample harboured two unique *EGFRvIII* breakpoints. We complemented these breakpoints with those identified from CPCT-02 (8/41 patients)[28] and PCAWG (11/40 patients)[29] whole-genome sequencing datasets. In several samples, we observed multiple, unique *EGFRvIII* breakpoints (**Figure S8A**) that could not have evolved from a tumour-specific ancestor *EGFRvIII* variant. In these cases, *EGFRvIII* thus has independently reoccurred within the same tumour.

Interestingly, the genomic breakpoints found in intron-1 show a difference in breakpoint density, where the region close to exon-1 contains 3.63 times fewer breakpoints per base than the region close to exon 2 ($p=4.9e^{-13}$; Fisher exact test; decision-boundary: chr7:55.182.397). Genomic breakpoints between exons 7-8 were more uniformly distributed (**Figure S8B**). The breakpoint preference in intron-1 may suggest preserving functional regions that confer a selective advantage to the tumour. Upon closer inspection, *EGFR* intron-1 contains three non-canonical exons[40] of which their expression is only rarely observed. We refer to these as exons A, B and C. The *EGFRvIII* breakpoint preference region is located 3' of exon-B (**Figure 5**) and thus preserves this exon at the genomic level. All datasets examined revealed junction-reads that initiated in exon-B and were spliced to exon-2 (*EGFRwt*) or exon-8 (*EGFRvIII*) (**Figure 6**). However, the fraction of transcripts containing exon-B was low compared to those initiating in exon-1 (≤1.05%; **Figure S8C**), indicating exon-B expression is driven by a weak promoter. Transcripts spliced from exons A or C to exon-2 were extremely rare.

In samples with breakpoints retaining exon-B, a novel exon-B-exon-8 *EGFR(vIII/B)* variant is created (**Figure 6**). This variant was confirmed with RT-PCR in six out of ten tested tumour samples (**Supplementary Table 2**). We verified the presence of the exon-B→exon-8→exon-9 sequence in three samples (GenBank: MZ484953, MZ484954 and MZ484955). *EGFR* transcripts that initiate in exon-B lack part of the extracellular domain on protein level as the translation initiation sites are located in exon-2 or exon-8. To test the potential functional role of exon-B variants, we created constructs of EGFR starting in exon-B and spliced to either exon-2 or exon-8.

Even after optimizing the Kozak sequence surrounding the translation initiation site, we failed to see the expression of 'exon-B' variants in any of the 16 constructs generated. This absent expression was not due to a potential lethality of exon-B constructs as (1) RT-PCR did show expression of the *EGFR* transgene and (2) biscistronic constructs (in which eGFP was independently translated from EGFR constructs as they were separated by an IRES sequence) did express eGFP. These data argue for an inferior protein translation of exon-B

transcripts. Given the inadequate translation into protein combined with the low level of transcripts incorporating exon-B, we deemed it unlikely these constructs significantly impact the tumour biology.

We explored the possibility that late breakpoints retain regulatory sequences further. H3K27ac ChIP-seq data from recent studies on *EGFR* enhancers[25,26] were plotted onto the *EGFR* locus (**Figure 5, S9**). An enhancer, previously referred to as "E3"[25], located just 5' to the *EGFRvIII* breakpoint preference region and is thus more often preserved. This region is conserved across 100 vertebrates. Previous experiments using CRSIPRi demonstrated its functional relevance in cell fitness. Unfortunately, too few samples with detected breakpoints and combined RNA-seq and DNA-seq data were available to determine whether late breakpoints have a higher fractional *EGFRvIII* expression (**Figure S10**).

Discussion

EGFR is commonly amplified, mutated and activated in glioblastoma, resulting in increased cell invasion and proliferation[41]. *EGFRvIII* is a specific tumour marker often present in glioblastoma, that has been intensively investigated[9,18]. Here, we report on this genomic mutation using a large glioblastoma *EGFRvIII* omics dataset. To maximize statistical power, analysis was performed across a combined cohort of four RNA datasets and two independent whole-genome sequencing datasets. Previous data on the prognostic value of *EGFRvIII* was conflicting, with some suggesting *EGFRvIII* is a negative[42,43] or a positive[44] prognostic marker, where other studies also suggested it did not affect survival[45,46]. Here, we demonstrate that within patients with *EGFR* amplified glioblastoma, we observed no difference in survival between *EGFRvIII* positive and negative tumours. Because *EGFRvIII* is known to be spatially heterogeneously distributed[13,14], *EGFRvIII* positive tumours can therefore, through sampling, be marked *EGFRvIII*-negative by omics analysis. Tumour sampling is therefore a limitation potentially influencing this survival analysis.

The expression levels of *EGFRvIII* and *EGFRwt* were anti-correlated and the total *EGFR* levels were generally lower when higher levels of *EGFRvIII* were present. This is in agreement with the hypothesis that *EGFRvIII* lowers the tumours' dependency on high *EGFR*-amplification levels[18].

Transcriptomes of *EGFRvIII* positive and negative tumours showed only minor differences (**Figure 3B**). A possibly related factor of this limited difference may be the ability of *EGFRvIII* to alter expression in *EGFRvIII*-negative tumour cells[15]. Within *EGFR* amplified tumours, those with ≥10% *EGFRvIII* were found to have significantly lower expression of *CDK4* and *MDM2* due to a lower incidence of respective amplifications. This inverse correlation may point towards crosstalk or redundancy between these pathways. Of the genes correlated positively to *EGFRvIII* expression, *RFWD3* can form a complex with *MDM2*, known for regulating p53[47]. Furthermore, Sprouty/Spred family genes were consistently associated with *EGFRvIII* presence and subsequent expression and are known for their inhibiting role in Ras/Raf/ERK[48] and involvement in *EGF*/*EGFR* signaling. The presented results are not supporting the standpoint that *EGFRvIII* is causing large distinct changes in downstream gene expression compared with *EGFRwt* amplifications.

Overall, our molecular analysis demonstrated that glioblastomas expressing *EGFRvIII* show a distinct but limited difference in their transcriptome compared with *EGFRwt*. The clearest observed signal is an increased correlation with classical glioblastoma subtype genes, which may indicate that the constitutively active *EGFRvIII* is, in the context of *EGFR* amplified glioblastoma, stronger in downstream *EGFR* signalling than (amplified) *EGFRwt*. This is in line with the lower total *EGFR* levels for tumours having higher *EGFRvIII* levels.

We found a broad range of *EGFRvIII*/total *EGFR* expression levels (1%-95%). Such range is puzzling because, if *EGFRvIII* is only a variant that is stronger in activating downstream EGFR signalling, it is possible that *EGFRvIII* would simply outcompete the *EGFRwt* ecDNA copies. This would likely take place relatively quickly since ecDNA amplifications are notorious for increasing tumour heterogeneity[8]. However, presence of extrachromosomal *EGFRwt* copies lasts in virtually all analysed *EGFRvIII* positive tumours. An explanation could be that *EGFRvIII* depends on the presence of *EGFRwt*[49], for instance, to form dimers to complete *EGFRvIII* phosphorylation[50] or in an inter-cellular context, for instance by *EGFRvIII* dependent secretion of cytokines[15]. Such dependencies would likely come with a preferred *EGFRvIII*/*EGFRwt* ratio. Alternatively, the linear slope is indicative for an absence of selection pressure to retain *EGFRvIII* over *EGFRwt*. This absence can explain the highly-heterogeneous spatial and temporal expression pattern of the mutant. It may also explain the near identical survival between *EGFRvIII* positive and negative glioblastoma patients. However, if there is no selection pressure to retain *EGFRvIII*, it remains puzzling why this particular mutant is found at such a high frequency. *EGFR* signalling in glioblastomas is highly complex as the tumour can adopt various methods to enhance its pathway activation. Multiple mutations can co-exist in the same tumour, sometimes subclonal and with reported longitudinal differences, with a unique, different ligand dependency.

An earlier study proposed defining samples with a read count of at least 1% or 10% *EGFRvIII* compared with total *EGFR* as *EGFRvIII* positive[4]. We recommend similarly rather than using the presence of any *EGFRvIII* read, as mapping artefacts and index hopping/switching derived reads are common in multiplexed RNA-seq and because higher *EGFRvIII* percentages showed a stronger response signal.

Determination of the subclonal breakpoints in pre-mRNA data was more complicated than in datasets where breakpoints were clonal[20]. Breakpoints were found predominantly in samples with high fractions of *EGFRvIII*. The median *EGFRvIII* percentage in samples with detected breakpoints was 55%, whereas 29% in samples without.

Intriguingly, we find a minority of *EGFR* transcripts starting with a cryptic exon preferentially preserved in *EGFRvIII* expressing tumours. The first translation initiation site is located in exon-8, but the total exon-B read count is very low and, combined with a weak Kozak sequence, we did not consider this variant to be the main reason for a breakpoint preference.

Recently, the promotor and functional enhancers specifically retained in extrachromosomal *EGFR* fragments in glioblastoma and neuroblastoma cells have been interrogated[25]. These enhancers, including ‘E3’, were discovered using 4C-seq, H3K27ac ChIP-seq and a CRISPRi knock-down proliferation dropout assay. The E3 enhancer also showed H3K27ac in an independent dataset[26]. The preferential retention of intragenic enhancer E3 in *EGFRvIII* is in line with these observations. As the E3 enhancer is also conserved across vertebrates, it likely results in higher *EGFR* transcription rates. Unfortunately, both absolute

and relative *EGFRvIII* levels differ essentially between samples, which combined with a high level of *EGFRwt* heterogeneity makes it difficult to confirm this hypothesis.

In summary, using the largest combined *EGFRvIII* omics dataset to date, we find that the expression profiles of *EGFRvIII* positive tumours differ only marginally from *EGFRvIII* negative tumours. The results suggest that *EGFRvIII* mainly performs a similar role as *EGFRwt* but with a stronger affinity to activate *EGFR* downstream pathways, possibly linked to persistent activity independent of ligand(s). Furthermore, genomic breakpoints in intron-1 retain an enhancer that likely increases the expression of *EGFRvIII* transcripts. In this retrospective setting, no prognostic difference was found between *EGFRvIII* positive patients compared with those harbouring *EGFRwt* amplifications. However, associations between *EGFRvIII* and genes such as *CDK4*, *MDM2* and *PTPRZ1* suggest that the relation between *EGFR* and *EGFRvIII* is not fully understood and further research is needed, ideally to find therapies targeting both isoforms.

Funding

This study was supported by Télévie, Brussels, Belgium, AbbVie, Inc., a grant from the “Westlandse ride”, the Brain Tumour Charity (grant-number ET_2019_/2_10470) and Stichting Stophersentumoren.nl 2013.

Acknowledgements

We thank the GIGA facility of University of Liège for RNA-sequencing. This publication and the underlying study have been made possible on the basis of data that Hartwig Medical Foundation and the Center of Personalised Cancer Treatment (CPCT) have made available. The authors thank the European Organization for Research and Treatment of Cancer for permission to use the data from EORTC studies EORTC_1410 (Intellance-2) and EORTC_1542 (G-SAM) for this research. We thank Martin E. van Royen for contributing.

Conflicts of Interest

None declared.

References

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *N Engl J Med*. 2005;352(10):987-996.
2. McLendon R, Friedman A, Bigner D, et al. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*. 2008;455(7216):1061-1068.
3. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science*. 2008;321(5897):1807-1812.
4. Brennan CW, Verhaak RGW, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell*. 2014;157(3):753.
5. French PJ, Eoli M, Sepulveda JM, et al. Defining EGFR amplification status for clinical trial inclusion. *Neuro Oncol*. 2019;21(10):1263-1272.
6. Draaisma K, Chatzipli A, Taphoorn M, et al. Molecular evolution of IDH wild-type glioblastomas treated with standard of care affects survival and design of precision medicine trials: A report from the EORTC 1542 study. *J Clin Oncol*. 2020;38(1):81-99.
7. Lassman AB, Aldape KD, Ansell PJ, et al. Epidermal growth factor receptor (EGFR) amplification rates observed in screening patients for randomized trials in glioblastoma. *J Neurooncol*. 2019;144(1):205-210.
8. Verhaak RGW, Bafna V, Mischel PS. Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. *Nat Rev Cancer*. 2019;19(5):283-288.
9. Gan HK, Cvrljevic AN, Johns TG. The epidermal growth factor receptor variant III (EGFRvIII): Where wild things are altered. *FEBS J*. 2013;280(21):5350-5370.
10. Lassman A, Pugh S, Wang T, et al. ACTR-21. A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED PHASE 3 TRIAL OF DEPATUXIZUMAB MAFODOTIN (ABT-414) IN EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR) AMPLIFIED (AMP) NEWLY DIAGNOSED GLIOBLASTOMA (nGBM). *Neuro Oncol*. 2019;21(Supplement_6):vi17-vi17. doi:10.1093/neuonc/noz175.064
11. Weller M, Butowski N, Tran DD, et al. 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(10):1373-1385.
12. 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.
13. Nathanson DA, Gini B, Mottahedeh J, et al. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science*. 2014;343(6166):72-76.
14. Del Vecchio CA, Giacomini CP, Vogel H, et al. EGFRvIII gene rearrangement is an early event in glioblastoma tumorigenesis and expression defines a hierarchy modulated by epigenetic mechanisms. *Oncogene*. 2013;32(21):2670-2681.
15. Zanca C, Villa GR, Benitez JA, et al. Glioblastoma cellular cross-talk converges on NF-κB to attenuate EGFR inhibitor sensitivity. *Genes Dev*. 2017;31(12):1212-1227.
16. Van Den Bent MJ, Gao Y, Kerkhof M, et al. Changes in the EGFR amplification and EGFRvIII expression between paired primary and recurrent glioblastomas. *Neuro Oncol*. 2015;17(7):935-941.

17. Wang J, Cazzato E, Ladewig E, et al. Clonal evolution of glioblastoma under therapy. *Nat Genet*. 2016;48(7):768-776.

18. Hoogstrate Y, Vallentgoed W, Kros JM, et al. EGFR mutations are associated with response to depatux-m in combination with temozolomide and result in a receptor that is hypersensitive to ligand. *Neuro-Oncology Adv*. 2020;2(1):1-15.

19. Erdem-Eraslan L, Van Den Bent MJ, Hoogstrate Y, et al. Identification of patients with recurrent glioblastoma who may benefit from combined bevacizumab and CCNU Therapy: A Report from the BELOB Trial. *Cancer Res*. 2016;76(3):525-534.

20. Hoogstrate Y, Komor MA, Riet J Van, et al. Detection of fusion transcripts and their genomic breakpoints from RNA sequencing data. *bioRxiv*. 2021;(2021.05.17.441778):1-27. doi:10.1101/2021.05.17.441778

21. Dobin A, Davis CA, Schlesinger F, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.

22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12).

23. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.

24. Reilly EB, Phillips AC, Buchanan FG, et al. Characterization of ABT-806, a humanized tumor-specific anti-EGFR monoclonal antibody. *Mol Cancer Ther*. 2015;14(5):1411-1451.

25. Morton AR, Dogan-Artun N, Faber ZJ, et al. Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell*. 2019;179(6):1330-1341.e13.

26. Jameson NM, Ma J, Benitez J, et al. Intron 1-mediated regulation of EGFR expression in EGFR-dependent malignancies is mediated by AP-1 and BET proteins. *Mol Cancer Res*. 2019;17(11):2208-2220.

27. Erdem-Eraslan L, Gao Y, Kloosterhof NK, et al. Mutation specific functions of EGFR result in a mutation-specific downstream pathway activation. *Eur J Cancer*. 2015;51(7):893-903.

28. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature*. 2019;575(7781):210-216.

29. Campbell PJ, Getz G, Korbel JO, et al. Pan-cancer analysis of whole genomes. *Nature*. 2020;578(7793):82-93.

30. Lassman AB, Van Den Bent MJ, Gan HK, et al. Safety and efficacy of depatuxizumab mafodotin + temozolomide in patients with EGFR -amplified, recurrent glioblastoma: Results from an international phase I multicenter trial. *Neuro Oncol*. 2019;21(1):106-114.

31. Tanaka K, Babic I, Nathanson D, et al. Oncogenic EGFR signaling activates an mTORC2-NF-κB pathway that promotes chemotherapy resistance. *Cancer Discov*. 2011;1(6):524-538.

32. Watanabe K, Tachibana O, Sato K, Yonekawa Y, Kleihues P, Ohgaki H. Overexpression of the EGF receptor and p53 mutations are mutually exclusive in the evolution of primary and secondary glioblastomas. *Brain Pathol*. 1996;6(3):217-223.

33. Di Stefano AL, Picca A, Saragoussi E, et al. Clinical, molecular, and radiomic profile of gliomas with FGFR3-TACC3 fusions. *Neuro Oncol*. 2020;22(11):1614-1624.

34. 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.

35. Orzan F, Pagani F, Cominelli M, et al. A simplified integrated molecular and immunohistochemistry-based algorithm allows high accuracy prediction of glioblastoma transcriptional subtypes. *Lab Investig*. 2020;100(10):1330-1344.

36. Verhaak RGW, Hoadley KA, Purdom E, et al. Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010;17(1):98-110.

37. Garofano L, Migliozzi S, Oh YT, et al. Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. *Nat Cancer*. 2021;2(2):141-156.

38. Matjašič A, Zupan A, Boštjančič E, Pižem J, Popović M, Kolenc D. A novel PTPRZ1-ETV1 fusion in gliomas. *Brain Pathol*. 2020;30(2):226-234.

39. Jonkers I, Lis JT. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 2015;16(3):167-177.

40. Ota T, Suzuki Y, Nishikawa T, et al. Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet*. 2004;36(1):40-45.

41. Talasila KM, Soentgerath A, Euskirchen P, et al. EGFR wild-type amplification and activation promote invasion and development of glioblastoma independent of angiogenesis. *Acta Neuropathol*. 2013;125(5):683-698.

42. Feldkamp MM, Lala P, Lau N, Roncari L, Guha A. Expression of activated epidermal growth factor receptors, Ras-guanosine triphosphate, and mitogen-activated protein kinase in human glioblastoma multiforme specimens. *Neurosurgery*. 1999;45(6):1442-1453.

43. Shinojima N, Tada K, Shiraishi S, et al. Prognostic Value of Epidermal Growth Factor Receptor in Patients with Glioblastoma Multiforme. *Cancer Res*. 2003;63(20):6962-6970.

44. Montano N, Cenci T, Martini M, et al. Expression of EGFRvIII in glioblastoma: Prognostic significance revisited. *Neoplasia*. 2011;13(12):1113-1121.

45. Heimberger AB, Hlatky R, Suki D, et al. Prognostic effect of epidermal growth factor receptor and EGFRvIII in glioblastoma multiforme patients. *Clin Cancer Res*. 2005;11(4):1462-1466.

46. Van Den Bent MJ, Brandes AA, Rampling R, et al. Randomized phase II trial of erlotinib versus temozolomide or carmustine in recurrent glioblastoma: EORTC brain tumor group study 26034. *J Clin Oncol*. 2009;27(8):1268-1274.

47. Fu X, Yucer N, Liu S, et al. RFWD3-Mdm2 ubiquitin ligase complex positively regulates p53 stability in response to DNA damage. *Proc Natl Acad Sci U S A*. 2010;107(10):4579-4584.

48. Kawazoe T, Taniguchi K. The Sprouty/Spred family as tumor suppressors: Coming of age. *Cancer Sci*. 2019;110(5):1525-1535.

49. Zadeh G, Bhat KPL, Aldape K. EGFR and EGFRvIII in Glioblastoma: Partners in Crime. *Cancer Cell*. 2013;24(4):403-404.

50. Li L, Chakraborty S, Yang CR, et al. An EGFR wild type-EGFRvIII-HB-EGF feed-forward loop regulates the activation of EGFRvIII. *Oncogene*. 2014;33(33):4253-4264.

Figure captions

Figure 1 Range of *EGFRvIII* percentages relative to total *EGFR*. Results are split per dataset (top) and combined (bottom). Grey vertical lines (Intellance-2) indicate levels determined by both full and panel-based RNA-seq where the actual percentages reflect their mean. Mutation statuses are indicated underneath. N/A-values are indicated in black.

Figure 2 EGFRwt/EGFRvIII correlations. (**A**) *EGFRwt* and *EGFRvIII* correlation and (**B**) total *EGFR* and percentage of correlation, per dataset. **A**) the correlations between *EGFRwt* and *EGFRvIII* are negative. **B**) Y-axis represents a surrogate for the total *EGFR* level (VST transformed sum of *EGFRwt* + *EGFRvIII* junction reads, because the full gene *EGFR* read count is negatively affected by exons missing in *EGFRvIII)*. Correlations are negative, indicating that tumours with higher proportions of *EGFRvIII* have lower levels of both variants combined.

Figure 3 (**A**) DE analysis between EGFR amplified samples with (≥10%) and without *EGFRvIII* (<1%), with batch correction for the four datasets (Intellance-2, G-SAM, BELOB and TCGA-GBM). 213/15.617 protein coding genes were differentially expressed, including *DLX1, DLX2, TSPAN31*, *TMPRSS7*, *PPBP* and *DPT*. Classical subtype genes are marked black. Overall LFCs were more often negative while the majority of the classical subtype genes had a positive LFC. (**B**) First two components of a supervised principal component analysis (213 DE genes). (**C**) Z-scores of Pearson correlation tests between genes and the relative *EGFRvIII* (x-axis) and *EGFRwt* (y-axis) levels, in samples with ≥10% *EGFRvIII*. Values near 0 represent no correlation, negative values a negative correlation and positive values represent a positive correlation. Classical subtype genes are marked black. Genes with a significant difference (t-test; q-value < 0.01) are marked purple. Genes showing a trend (q-value < 0.1) are marked blue.

Figure 4: Kaplan-Meier survival plots of patients with *EGFR* amplification with/without *EGFRvIII*. Patients included were from the BELOB trial, Intellance-2 TMZ/control arm, primary G-SAM tumours and primary TCGA-GBM tumours. Difference in patient survival between *EGFR* amplified glioblastoma patients with/without *EGFRvIII* (≥1% and ≥10%) was not significant and neither in each dataset separately (**Figure S7**).

Figure 5: Overview of genomic *EGFR* locus (exons 1-11) and *EGFRvIII* breakpoints. From bottom to top: chr7, transcript annotations, late and early breakpoint regions, conservation (purple), H3K27ac intensity in GSC23 cells[25] (green) and the actual breakpoints (blue and mustard). Genomic *EGFRvIII* breakpoints are indicated with mustard (RNA detected) and blue (DNA detected) bars on top.

Figure 6 Exon-B expression. Spliced read counts for exon-B (exon-B→exon-2: bars up & exon-B→exon-8: bars down) in tumours with RNA detected genomic *EGFRvIII* breakpoint. Tumours with a 'late' intron-1 breakpoint (≥chr7:55.182.397) are marked with a square and 'early' with a cross. Regular (**A**) and high (**B**) depth datasets were split. *EGFRvIII* exon-B variant reads (exon-B→exon-8) are only present in tumours with a late *EGFRvIII* breakpoint, which retain exon-B.

Figure 1

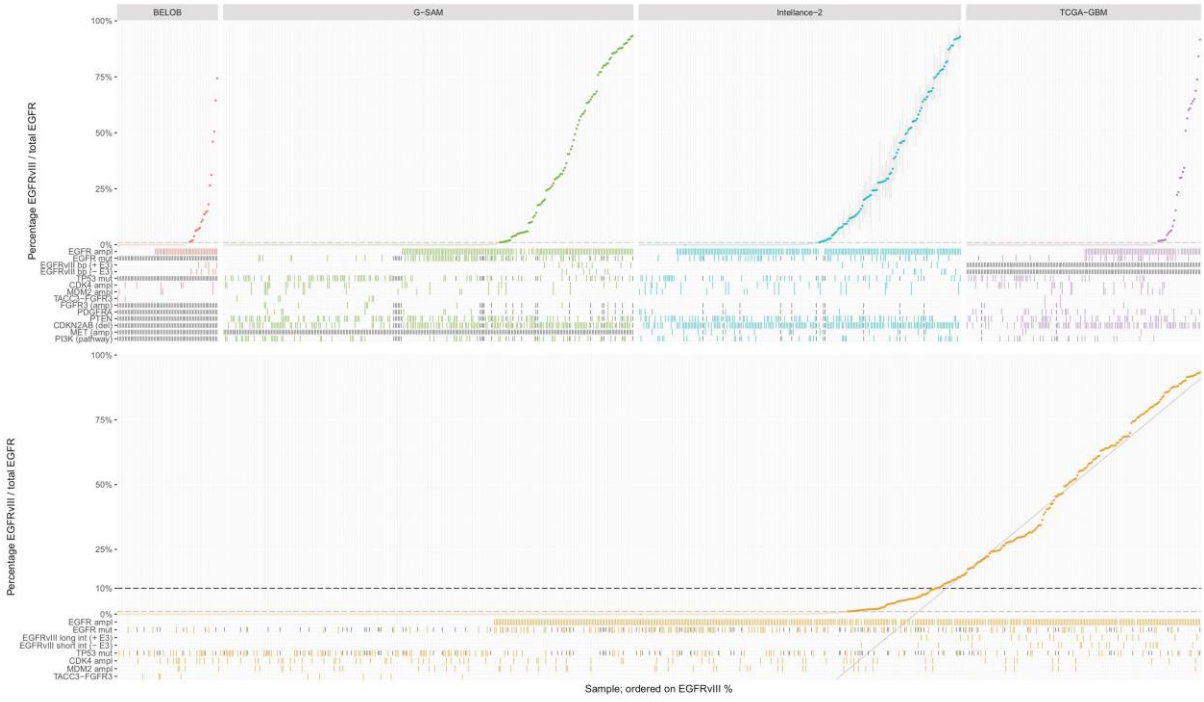

Figure 2

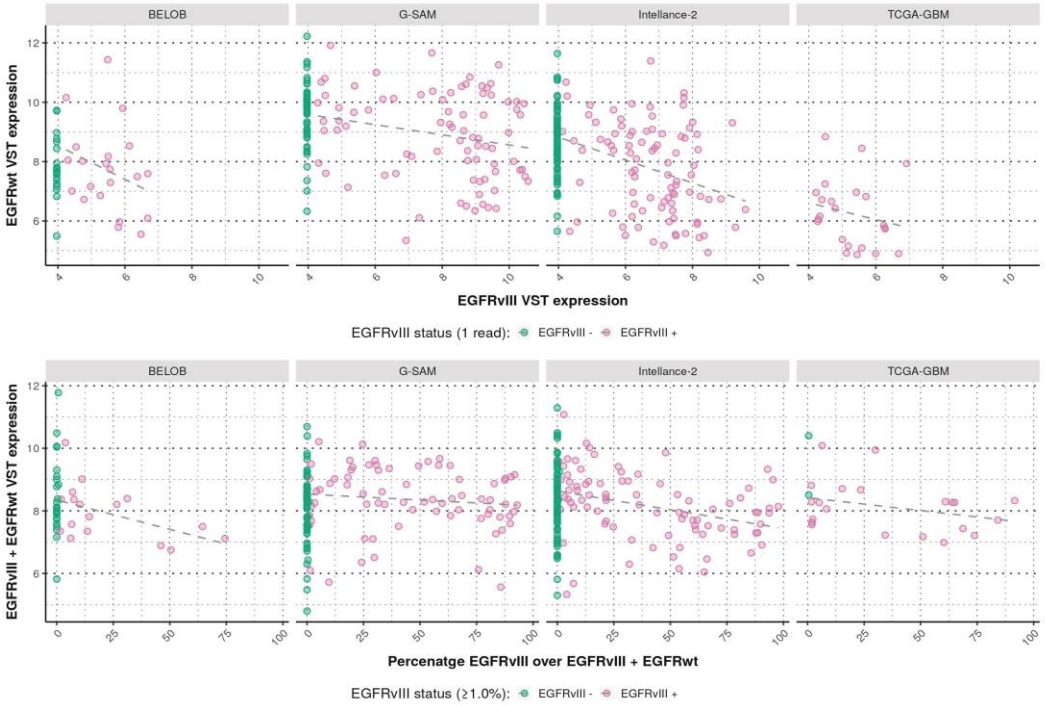

Figure 3

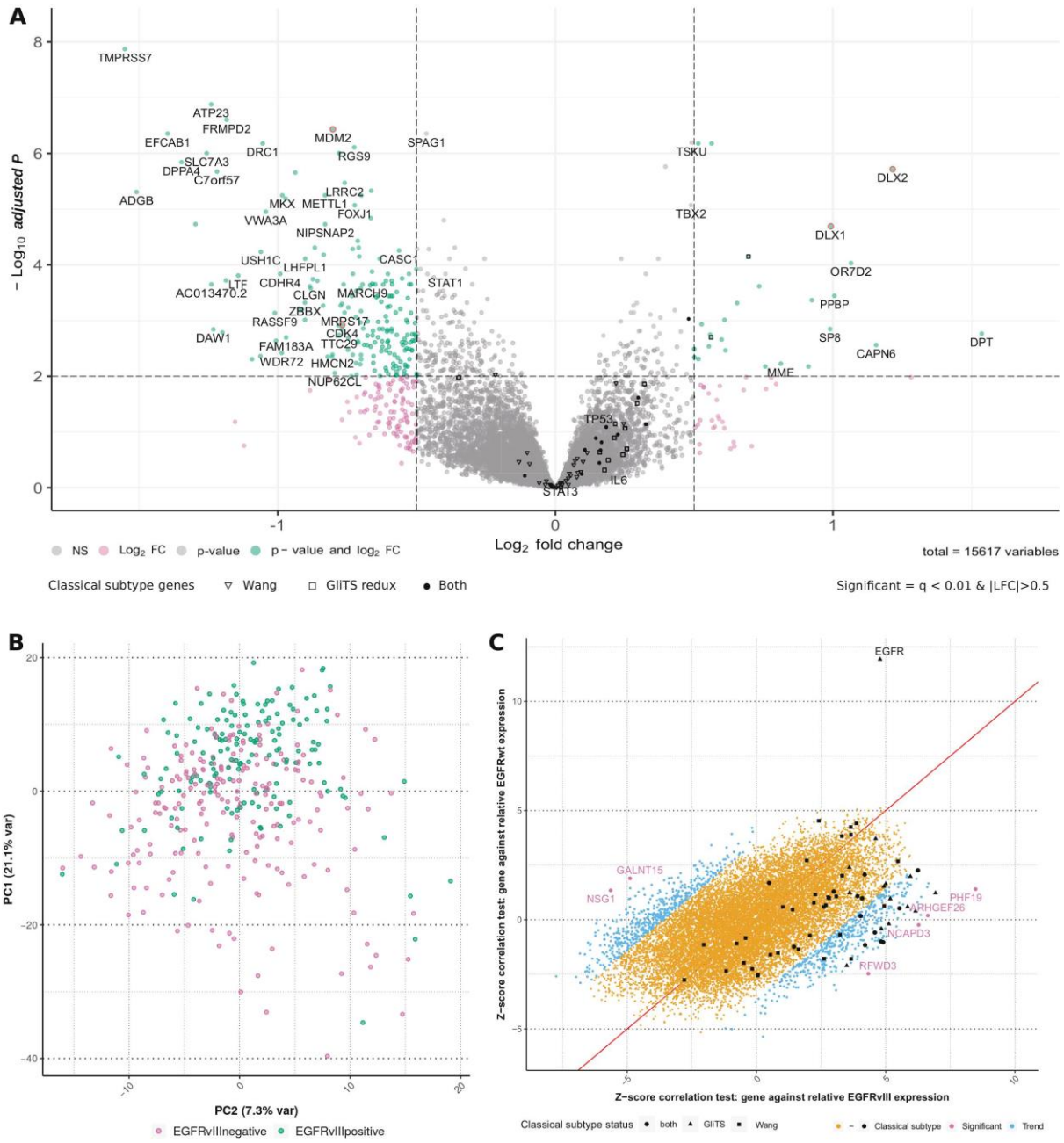

Figure 4

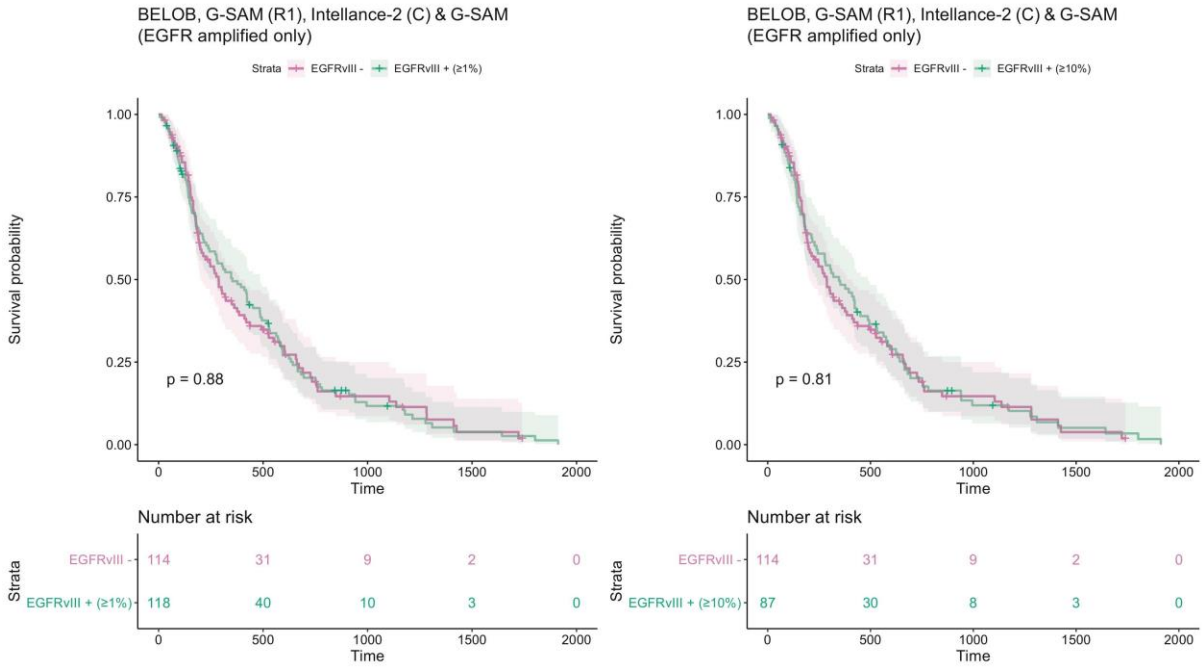

Figure 5

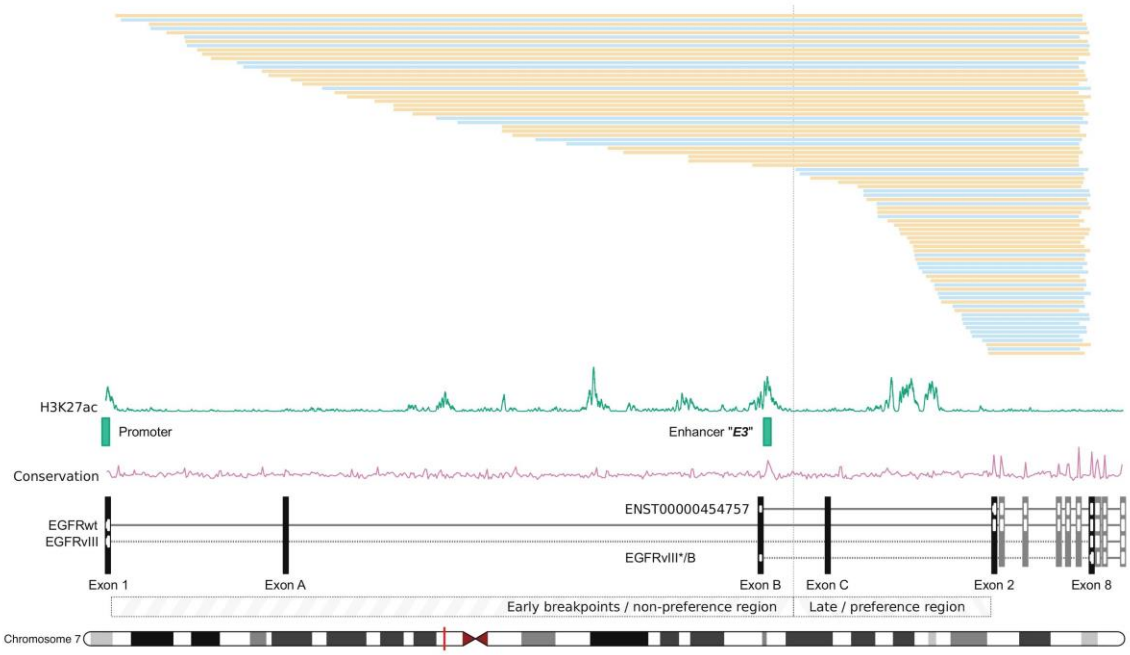

Figure 6