nature cell biology

ARTICLES

https://doi.org/10.1038/s41556-021-00639-4

Check for updates

Circular RNA-encoded oncogenic E-cadherin variant promotes glioblastoma tumorigenicity through activation of EGFR–STAT3 signalling

Xinya Gao[1,9], Xin Xia[1,9], Fanying Li[1,9], Maolei Zhang[1,9], Huangkai Zhou[1,2], Xujia Wu[1], Jian Zhong[1], Zheng Zhao[3], Kun Zhao[1], Dawei Liu[4], Feizhe Xiao[5], Qiang Xu[6], Tao Jiang[3], Bo Li[7], Shi-Yuan Cheng[8]✉ and Nu Zhang[1]✉

Activated EGFR signalling drives tumorigenicity in 50% of glioblastoma (GBM). However, EGFR-targeting therapy has proven ineffective in treating patients with GBM, indicating that there is redundant EGFR activation. Circular RNAs are covalently closed RNA transcripts that are involved in various physiological and pathological processes. Herein, we report an additional activation mechanism of EGFR signalling in GBM by an undescribed secretory E-cadherin protein variant (C-E-Cad) encoded by a circular E-cadherin (*circ-E-Cad*) RNA through multiple-round open reading frame translation. C-E-Cad is overexpressed in GBM and promotes glioma stem cell tumorigenicity. C-E-Cad activates EGFR independent of EGF through association with the EGFR CR2 domain using a unique 14-amino-acid carboxy terminus, thereby maintaining glioma stem cell tumorigenicity. Notably, inhibition of C-E-Cad markedly enhances the antitumour activity of therapeutic anti-EGFR strategies in GBM. Our results uncover a critical role of C-E-Cad in stimulating EGFR signalling and provide a promising approach for treating EGFR-driven GBM.

Glioblastoma (GBM) is a common and the most malignant primary tumour in the brain and has a poor prognosis[1,2]. In addition to the difficulty of blood–brain barrier penetration, intratumoural heterogeneity with high plasticity limits the effects of current therapies for GBM[3,4]. EGFR amplification or mutation occurs in ~50% of primary GBM and is an established oncogenic driver for GBM tumorigenesis[1,5]. However, efforts in targeting EGFR in combination with standard care or other therapies have proven ineffective for GBM treatment[6–8]. Although heterogeneity of the EGFR signalling network may render therapy unsuccessful[5], other unidentified molecular modulators co-existing in EGFR-expressing tumours could contribute to GBM malignancy and therapy resistance.

Circular RNAs (circRNAs) are covalently closed RNA transcripts that are generally expressed at lower levels than their associated linear mRNAs[9–11]. However, certain circRNAs are highly enriched during neurogenesis[12], which implies that circRNAs and their linear counterparts may have distinct functions. circRNAs have been demonstrated as translatable RNAs[13]. Open reading frames (ORFs) in translatable circRNAs may not have an in-frame stop codon, which results in multiple rounds of translation and the production of certain proteins[14].

In this study, we sought to identify coding circRNAs in GBM and whether multiple rounds of circRNA translation generate proteins that could serve as potential targets for GBM.

Results

Circular E-cadherin RNA is selectively expressed in GBM. Since several circRNAs have been found to be enriched in tumours[15,16], we determined differentially expressed circRNAs (DEcRs) in GBM by RNA sequencing (RNA-seq) analyses of GBM specimens and paired normal brain (NB) tissues from 12 patients[17] (Fig. 1a, upper left). Among a total of identified 139,090 circRNAs, 21,535 have been annotated in circBase[18] (Extended Data Fig. 1a). The majority of the circRNAs were less than 1,500-nucleotides long (Extended Data Fig. 1b). We identified nominally significant DEcRs between GBM and NB (Extended Data Fig. 1c (left) and 1d, and Supplementary Table 1). Of the identified 2,289 DEcRs, 1,305 were downregulated while 984 were upregulated when comparing circRNAs in GBM to that in NB (Extended Data Fig. 1c, right).

To discover translatable circRNAs, we performed ribosome profiling of ten paired GBM and NB samples[19] (Fig. 1a, upper right). We focused on the head-to-tail junction reads that were specific for translating circRNAs. We excluded mismatches and limited the minimum read-junction overlap to 9 nucleotides on either side. A candidate circRNA was called only when the unique junction reads were found in more than three samples and more than ten total junction reads were seen (Extended Data Fig. 1c and Supplementary Table 2). A total of 1,879 coding circRNAs were called, of which 1,109 were annotated in circBase and 4 circRNAs were recently described[20] (Fig. 1a, lower left, and Extended Data

[1]Department of Neurosurgery, The First Affiliated Hospital of Sun Yat-sen University, Guangdong Provincial Key Laboratory of Brain Function and Disease, Guangdong Translational Medicine Innovation Platform, Guangzhou, China. [2]Gene Denovo Biotechnology (Guangzhou), Guangzhou, China. [3]Beijing Neurosurgical Institute, Capital Medical University, Beijing, China. [4]Department of Pathology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China. [5]Department of Scientific Research Section, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China. [6]GenomiCare Biotechnology (Shanghai), Shanghai, China. [7]Department of Biochemistry and Molecular Biology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, China. [8]The Ken and Ruth Davee Department of Neurology, Lou and Jean Malnati Brain Tumor Institute, The Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [9]These authors contributed equally: Xinya Gao, Xin Xia, Fanying Li, Maolei Zhang. ✉e-mail: shiyuan.cheng@northwestern.edu; zhangnu2@mail.sysu.edu.cn

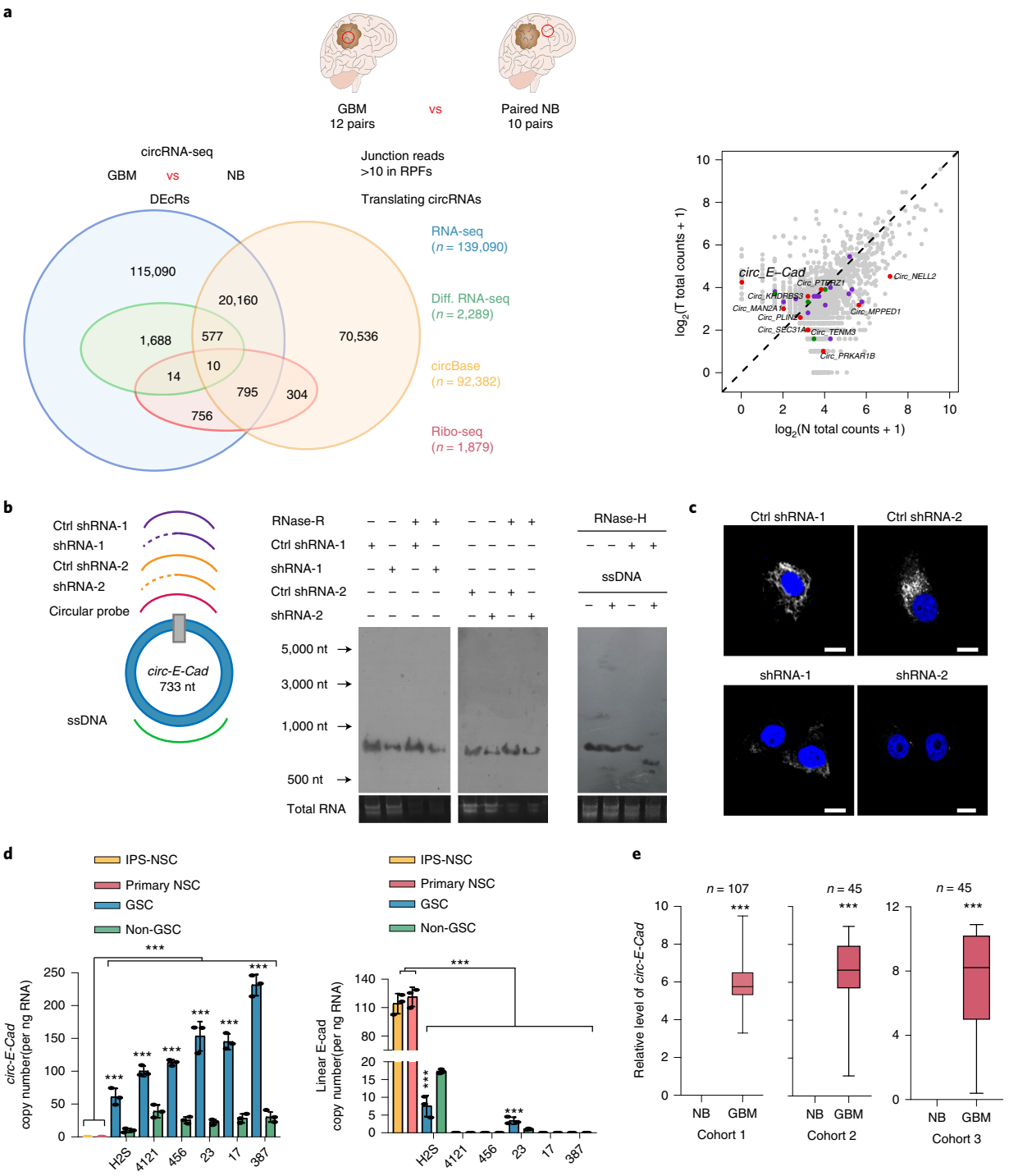


Fig. 1 | *circ-E-Cad* RNA is a potential coding circRNA and is overexpressed in GBM and GSCs. a, Upper: strategies used for circRNA-seq and ribosome profiling (Ribo-seq). A total of 12 pairs of GBM and NB samples were subjected to circRNA-seq, while 10 pairs of GBM and NB samples were subjected to ribosome profiling. Lower left: a Venn diagram of coding circRNAs and DEcRs intersecting in GBM or NB. Lower right: DEcRs annotated in circBase are labelled red, DEcRs are labelled purple, and DEcRs discovered by ribosome profiling in human heart[20] and our data are labelled in green; RPFs, ribosome protected frames. **b**, Left: northern blot of *circ-E-Cad* RNA in GSCH2S using junction probes with or without RNase-R treatment. Two junction shRNAs were used to confirm specificity. Ctrl, control. Right: different shifting rates of *circ-E-Cad* RNA and its linear form were determined using a ssDNA matched to *circ-E-Cad* RNA that was treated with RNase-H ($n=3$ independent experiments). **c**, Fluorescence in situ hybridization for *circ-E-Cad* RNA cellular localization in GSCH2S (H2S) cells ($n=3$ independent experiments). Scale bar, 20 μm. **d**, Real-time qPCR of circular and linear E-cadherin (E-cad) RNA expression in iPSC-derived NSCs (IPS-NSC), primary human NSCs, GSCs and their paired non-GSCs. $n=3$ independent experiments, data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P<0.001$. The *P* values for each group are listed in Supplementary Table 13. **e**, *circ-E-Cad* RNA expression in GBM and their paired NB samples. Cohort 1: 107 randomly selected GBM and their NB samples from the Department of Neurosurgery, The First Affiliated Hospital of Sun Yat-sen University. Cohort 2: 45 randomly selected GBMs and their paired NB from the Department of Neurosurgery, The Sun Yat-sen University Cancer Centre. Cohort 3: 45 randomly selected GBMs and their paired NB from the Department of Neurosurgery, The First Affiliated Hospital of Nanjing Medical University. Data are presented as boxes containing the first and third quartiles. The whiskers indicate the maxima and minima. Wilcoxon test, $^{***}P<0.001$. Source data are provided.

Fig. 1e, left). When these coding circRNAs were cross-referenced to the DEcRs (Fig. 1a, lower left), 24 circRNAs were differentially translated between GBM and NB (Fig. 1a, lower right, and Extended Data Fig. 1e, right).

We validated the ten candidates annotated in circBase in paired GBM and NB. Among them, circular E-cadherin (*circ-E-Cad*) RNA was most differentially expressed (Extended Data Fig. 2a, upper). To exclude biases induced by nontumoural cells, we performed quantitative PCR (qPCR) in a panel of patient-derived glioma stem cells (GSCs), including GSC456, GSC4121, GSC3691, GSCH2S, GSC387 (ref. [21]), GSC17 and GSC23 (ref. [22]), as well as neural stem cells (NSCs) and normal astrocytes (NHAs). The differential expression levels of *circ-E-Cad* RNA ranked at the top of these candidates (Extended Data Fig. 2a, lower).

Circularization of exons 7–10 forms the 733-nucleotide *circ-E-Cad* RNA (Extended Data Fig. 2b, upper). After the junction was confirmed (Extended Data Fig. 2b, lower), we found that *circ-E-Cad* RNA was resistant to RNase-R[23], and two short hairpin RNAs (shRNAs) targeting the circular junction successfully reduced *circ-E-Cad* RNA levels (Fig. 1b, left and middle). When a complementary single-stranded DNA (ssDNA) was applied, *circ-E-Cad* RNA displayed a different shift rate compared with its linear form during RNase-H digestion[24] (Fig. 1b, right). The *circ-E-Cad RNA* was localized in the cytoplasm (Fig. 1c), was upregulated in multiple CD133$^+$ GSCs versus CD133$^-$ non-GSCs and was undetectable in NSCs[25] (Fig. 1d). In three independent cohorts of clinical samples, *circ-E-Cad RNA* levels were higher in GBM compared to NB (Fig. 1e).

***circ-E-Cad* RNA encodes a 254-amino-acid protein.** circRNAs are reported to be translatable, and circRNA-encoded functional proteins have been described during GBM tumorigenesis[17,26,27]. *circ-E-Cad* RNA junction reads in ribosome profiling were detected in five out of ten GBM samples, whereas no junction reads were found in ten NB samples (Fig. 2a). We identified a potential internal ribosomal entry site (IRES) and validated its activity (Extended Data Fig. 2c). A cross-junction and multiple-round ORF driven by this IRES potentially encoded a 254-amino acid protein. We named this product as a circRNA-encoded E-cadherin (C-E-Cad). Due to the lack of a stop codon in the first-round read, C-E-Cad has a unique 14-aa tail at its carboxy terminus formed by a natural frameshift in the second-round translation (Fig. 2b). We produced and characterized a monoclonal antibody against the tail 14-aa sequences (Extended Data Fig. 2d). This antibody detected a specific protein at 28 kDa in GSCH2S cells overexpressing *circ-E-Cad* RNA, as well as endogenous C-E-Cad in GSC387 cells (Fig. 2c). The 14-aa residues in C-E-Cad were further authenticated by mass spectrometry (MS) in these GSCs (Extended Data Fig. 2e).

These 14-aa residues displayed no homology to any known proteins. C-E-Cad was highly expressed in CD133$^+$ GSCs compared to CD133$^-$ non-GSCs and in GBM compared to NB (Fig. 2d and Extended Data Fig. 2f). In 107 primary GBM and 85 NB tissues, C-E-Cad was detected in 90 out of 107 GBM (84.1%, Fig. 2e), whereas C-E-Cad was undetectable in all NB. Based on the quantification results, we divided the GBM into a C-E-Cad-high group and a C-E-Cad-low group. C-E-Cad expression inversely correlated with overall survival of patients with GBM (Fig. 2f, upper, and Extended Data Fig. 2g, left and middle), whereas there was no correlation of E-cadherin expression to GBM prognosis (Fig. 2f, lower, and Extended Data Fig. 2g, right). We examined whether C-E-Cad was co-expressed with SOX-2, an established marker for cell stemness[28]. Expression levels of C-E-Cad and *circ-E-Cad* RNA were positively correlated with that of SOX-2 in GBM (Fig. 2g).

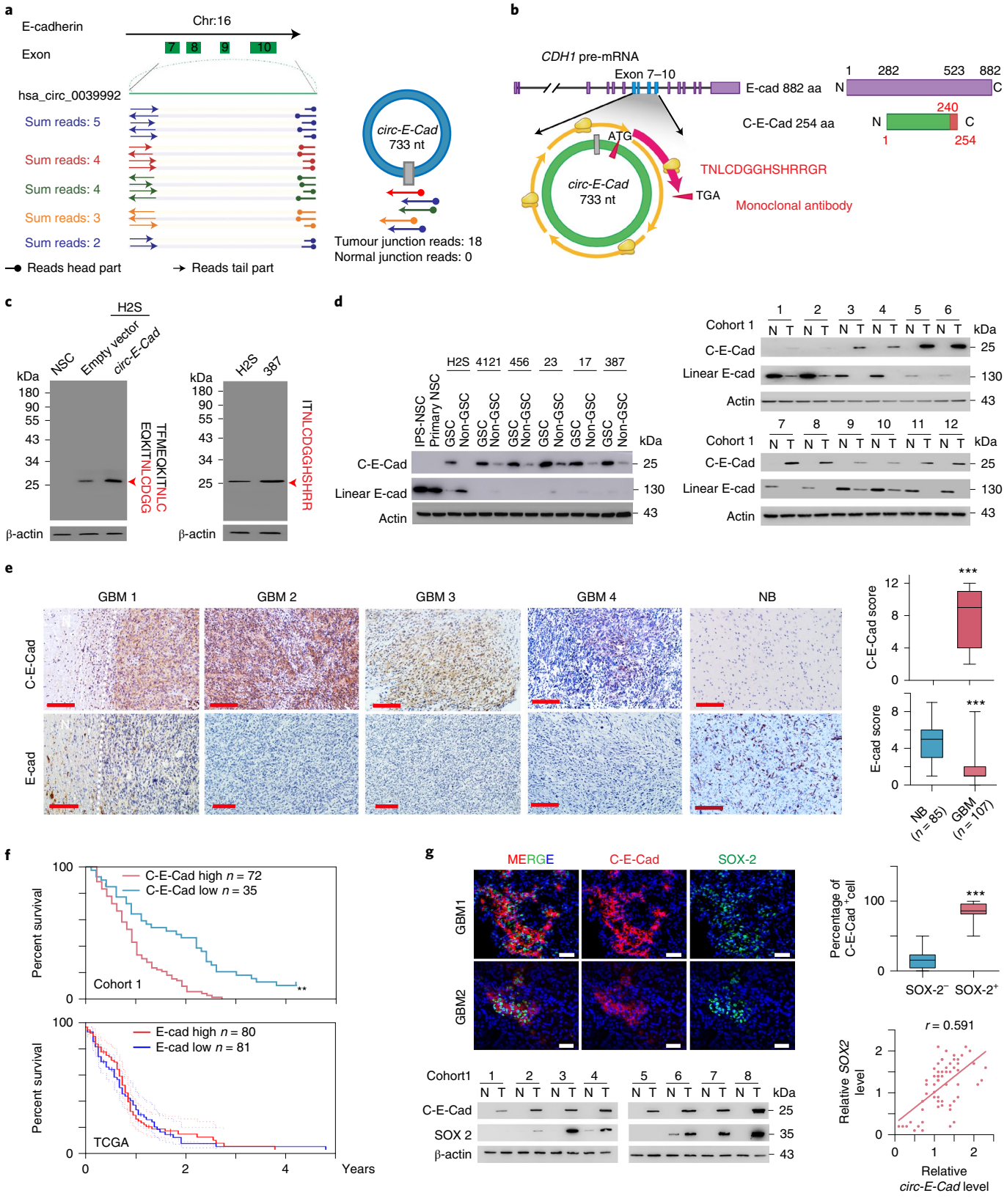
C-E-Cad regulates the biological properties and tumorigenicity of GSCs. We stably expressed *circ-E-Cad* RNA (or C-E-Cad ORF) in GSCH2S cells and stably knocked down circ-E-Cad RNA in GSC387 and GSC4121 cells. Knockdown (KD) of circ-E-cad reduced the expression levels of stemness markers and increased the levels of differentiation markers (Fig. 3a and Extended Data Fig. 3a). Conversely, overexpression of *circ-E-Cad* RNA or C-E-Cad ORF enhanced the stemness properties of the cells (Fig. 3a and Extended Data Fig. 3a). Functionally, sphere-forming frequency (Fig. 3b and Extended Data Fig. 3b, left), proliferation (Fig. 3c and Extended Data Fig. 3b, right), invasion (Fig. 3d and Extended Data Fig. 3c), anti-apoptosis and senescence resistance (Extended Data Fig. 3d,e) were reduced with KD of C-E-Cad in GSCs, whereas exogenous expression of C-E-Cad promoted the above phenotypes. In vivo, KD of C-E-Cad inhibited, while C-E-Cad expression enhanced, GSC intracranial tumour growth, which in turn altered the overall survival of animals bearing intracranial GSC xenografts (Fig. 3e,f and Extended Data Fig. 3f). Re-expression of a shRNA-resistant C-E-Cad (Extended Data Fig. 3g) restored the biological properties of cells in vitro and in vivo (Extended Data Fig. 3h–l). Additionally, expression of a mutated *circ-E-Cad* RNA (Extended Data Fig. 3g) failed to enhance these malignant phenotypes in vitro or in vivo (Extended Data Fig. 3h–l).

We identified symmetrical ALU sequences, which are critical for circRNA formation[29,30], on both sides of exons 7–10 of the gene encoding E-cadherin (*CDH1*). Knockout (KO) of downstream ALU using CRISPR–Cas9 techniques impaired the formation of *circ-E-Cad* RNA and decreased C-E-Cad expression in GSC387 and GSC4121 cells (Extended Data Fig. 4a,b). ALU-KO GSCs showed diminished sphere-forming, which was similar to that observed with the stable KD of *circ-E-Cad* RNA. The impaired stemness

Fig. 2 | *circ-E-Cad* RNA encodes a protein through multiple rounds of translation. a, Visualization of RFPs in the *circ-E-Cad* RNA junction. Five out of 10 GBMs had a total of 18 junction reads, whereas no junction reads were found in 10 paired NB samples. **b**, Illustration of the C-E-Cad protein encoded by *circ-E-Cad* RNA. The C terminus of C-E-Cad is translated by the second-round read (showing in red). A monoclonal mouse antibody was generated against the indicated C-terminal sequences. **c**, The antibody against C-E-Cad was characterized in GSCH2S cells that stably overexpressed empty vector or *circ-E-Cad* RNA (left) and in GSC387 (387) cells (right). $n=3$ independent experiments. **d**, Immunoblot (IB) of C-E-Cad and E-cadherin expression in NSCs, GSCs and their paired non-GSCs sorted by CD133 (left) and in 12 randomly selected GBM samples (T) (cohort 1) and their paired NB (N) samples (right). $n=3$ independent experiments. **e**, Immunohistochemistry (IHC) images of C-E-Cad and E-cadherin expression in GBM and NB tissues (left) and semiquantitative scoring of C-E-Cad and E-cadherin expression in 85 NB and 107 GBM samples. Scale bar, 250 μm (cohort 1, right). Data are presented as boxes containing the first and third quartiles. The whiskers indicate the maxima and minima. Wilcoxon test, $^{***}P<0.001$. **f**, Upper: Kaplan–Meier survival analysis of patients with GBM (cohort 1, $n=107$ biologically independent samples) with C-E-Cad levels. Lower: two-sided, log-rank analysis of E-cadherin expression in GBM of The Cancer Genome Atlas (TCGA) dataset. **g**, Upper left: immunofluorescence (IF) images of C-E-Cad and SOX-2 expression in GBM. Scale bar, 100 μm ($n=3$ independent experiments). Upper right: comparison of the percentages of C-E-Cad$^+$ cells among SOX-2$^+$ versus SOX-2$^-$ cells in 50 randomly selected microscopy fields of each tumour image (cohort 1, $n=50$ biologically independent samples). Data are presented as boxes containing the first and third quartiles. The whiskers indicate the maxima and minima. Wilcoxon test, $^{***}P<0.001$. Lower left: IB of C-E-Cad and SOX-2 expression in eight randomly selected GBM samples and their paired NB (cohort 1, $n=3$ independent experiments). Lower right: Pearson's correlation between *circ-E-Cad* RNA and *SOX2* mRNA expression in the 60 GBM sample cohort (cohort 1, $n=60$ biologically independent samples). Source data are provided.

properties of ALU-KO could be rescued by the re-expression of C-E-Cad (Extended Data Fig. 4c,d).

C-E-Cad activates STAT3, PI3K–AKT and MAPK–ERK signalling in GSCs. To elucidate the mechanism of C-E-Cad in GBM, we analysed RNA-seq data from GSC4121 cells with stable KD of *circ-E-Cad* RNA and found that 9 out of the 39 altered pathways were related to STAT3. Additionally, PI3K–AKT and MAPK signalling ranked at the top of these pathways (Extended Data Fig. 4e and Supplementary Table 3). To determine whether C-E-Cad



participates in STAT3 activation, we analysed RNA-seq data from GSC387 and GSC4121 cells in which STAT3 was stably depleted. Gene sets directly regulated by STAT3 in GSC387 and GSC4121 cells were obtained by cross-referencing the STAT3-regulated genes with the STAT3-binding genes[31,32]. A positive correlation between genes directly regulated by STAT3 and C-E-Cad expression was evident (Fig. 4a). Consistent with the knowledge that phosphorylation and nuclear translocation of STAT3 indicate STAT3 activation[33], C-E-Cad regulated p-STAT3 levels and nuclear localization (Fig. 4b and Extended Data Fig. 4f,g). C-E-Cad also stimulated p-AKT and p-ERK1/2, which are critical for the survival and proliferation of GSCs[34–36] (Fig. 4b and Extended Data Fig. 4f).

To determine whether C-E-Cad regulates the stemness properties of GSCs through STAT3, we re-expressed a constitutively activating (CA) STAT3 (T705E) in ALU-KO GSC387 and GSC4121 cells. CA-STAT3 rescued the impaired sphere-forming ability of ALU-KO GSC387 and GSC4121 cells (Extended Data Fig. 4h). CA-STAT3 also induced the expression of mesenchymal markers and reduced the expression of epithelial markers in ALU-KO GSC387 and GSC4121. Additionally, a STAT3 inhibitor, WP1066, inhibited the C-E-Cad-induced the stemness properties of the cells (Extended Data Fig. 4i).

Soluble-cleaved E-cadherin can promote cancer progression[37,38], and E-cadherin can regulate p-STAT3 in embryonic stem cells[39]. We generated *CDH1* KOs in human NSCs and in GSC17 and GSC23 cells using the CRISPR–Cas9 system and separately re-expressed C-E-Cad or E-cadherin. KO of *CDH1* depleted both *circ-E-Cad* RNA and *CDH1* mRNA (Extended Data Fig. 4j). Levels of p-STAT3 were markedly reduced in GSC17 and GSC23 *CDH1* KO cells, but were restored by C-E-Cad expression (Fig. 4c). Moreover, p-STAT3 was moderately reduced, whereas C-E-Cad expression elevated p-STAT3 in *CDH1* KO NSCs (Extended Data Fig. 4k). In all the *CDH1* KO cells, expression of E-cadherin had minimal effects on p-STAT3 abundance (Fig. 4c and Extended Data Fig. 4k). GSC17 and GSC23 *CDH1* KO cells showed impairments in stemness properties and brain tumorigenicity. These attenuated phenotypes were rescued by expression of C-E-Cad but not E-cadherin. *CDH1* KO did not render any appreciable effects on the biological behaviours of NSCs, while expression of C-E-Cad enhanced the stemness properties of these NSCs (Fig. 4d–g and Extended Data Fig. 4l–n).

C-E-Cad is a secretory protein and activates EGFR. C-E-Cad was soluble and secreted out of cells (Fig. 5a,b). We then purified His-tagged C-E-Cad (rC-E-Cad) and found that rC-E-Cad, but not a rC-E-Cad mutant with deleted (Δ)14 aa, stimulates p-STAT3 (Extended Data Fig. 5a,b), promoted sphere-forming, cell proliferation, invasion, apoptosis resistance and senescence resistance of *CDH1* KO cells. These stimulatory effects were diminished when a specific anti-C-E-Cad antibody was included (Fig. 5c–f and Extended Data Fig. 5b–d).

In GBM, receptor tyrosine kinases (RTKs) are frequently amplified[40–42]. We determined which RTKs were stimulated by C-E-Cad using MS and identified EGFR as a potential C-E-Cad binding partner (Extended Data Fig. 5e and Supplementary Table 3). We next stably knocked down *EGFR*, *PDGFRA*, *MET* or *IL6R* in GSC387 and GSC4121 cells (Extended Data Fig. 5f) followed by treatments with rC-E-Cad[43,44]. Notably, *circ-E-Cad* RNA KD reduced interleukin-6 (IL-6) secretion (Extended Data Fig. 5g). When compared to the controls, rC-E-Cad less effectively activated downstream signals in GSCs with separate KD of *PDGFRA*, *MET* or *IL6R*, but almost failed to activate p-STAT3, p-AKT and p-ERK1/2 in GSCs with *EGFR* KD, which suggests that EGFR-mediated C-E-Cad activation is necessary for these signalling events (Fig. 5g and Extended Data Fig. 5h, left). The less effectiveness of downstream signal activation by rC-E-Cad in GSCs with separate KD of *PDGFRA*, *MET* or *IL6R* may be due to a positive activation loop of EGFR signalling, such as EGFR–AKT–IL-6–IL-6R.

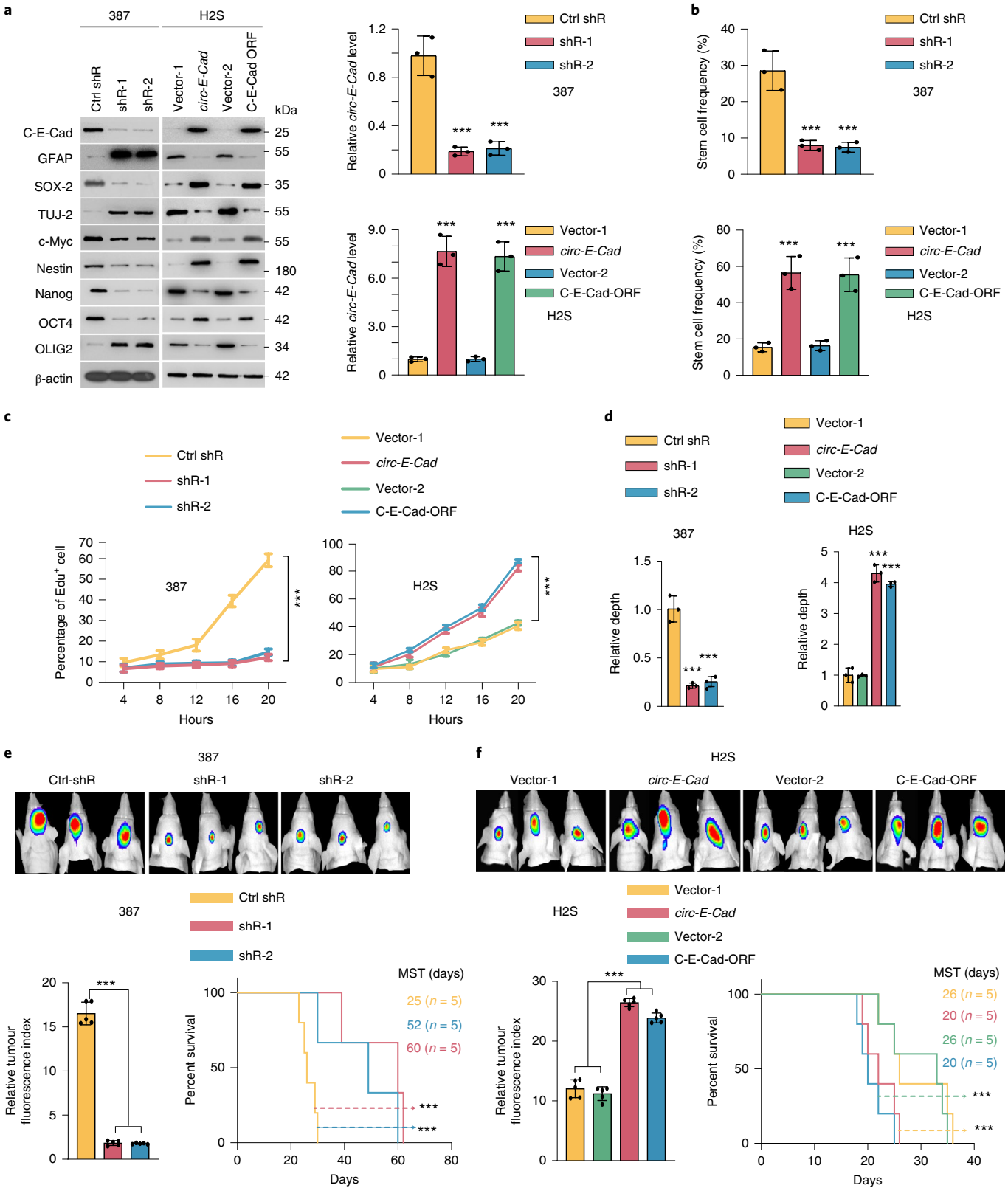
The RTK activity of EGFR is determined by the induced phosphorylation of several tyrosine (Y) residues in the C terminus[45]. We expressed EGFR wild-type (WT) or a Y1068A mutant in *EGFR* KD GSC387 and GSC4121 cells and then treated them with rC-E-Cad. EGFR depletion markedly reduced the levels of p-STAT3, p-AKT and p-ERK1/2 that were induced by C-E-Cad. WT EGFR, but not the Y1068A mutant, rescued the diminished levels of p-STAT3, p-AKT and p-ERK1/2 in the modified GSCs (Fig. 5h and Extended Data Fig. 5h, right). To determine whether C-E-Cad had an independent role in EGFR activation, we expressed C-E-Cad in CD133$^-$ non-GSCs that are deficient in the expression of EGFR and C-E-Cad. C-E-Cad was unable to induce p-STAT3 or stimulate the stemness properties of CD133$^-$ non-GSCs. When C-E-Cad was co-expressed, appreciable levels of p-STAT3 and sphere-forming were induced in CD133$^-$ non-GSCs (Extended Data Fig. 5i,j).

The C terminus of C-E-Cad and the CR2 domain of EGFR mediate their direct interaction. We next determined whether C-E-Cad is associated with EGFR. We found that both recombinant proteins interact with each other and that endogenous C-E-Cad associates with EGFR in GSCs (Fig. 6a). Additionally, C-E-Cad and EGFR were co-localized in cells (Fig. 6b). Then, we assessed whether the unique 14-aa residues at the C terminus of C-E-Cad mediate this association. C-E-Cad-Δ14aa was unable to interact and co-localize with EGFR (Fig. 6b,c, left, and Extended Data Fig. 5k). Compared to WT C-E-Cad, C-E-Cad-Δ14aa also failed to activate p-STAT3 (Fig. 6c, right). Conversely, mutants that contain the CR2 domain of EGFR, but not other mutants[46], were able to interact with the recombinant C-E-Cad, which suggests that C-E-Cad binds to the CR2 domain (Fig. 6d).

Fig. 3 | C-E-Cad maintains self-renewal and tumorigenicity of GSCs in vitro and in vivo. a, Left: IB of stemness and differentiation marker expression in GSC387 cells with stable *circ-E-Cad* RNA KD (referred as to Ctrl shR, shR-1 and shR-2, hereafter) and in GSCH2S with *circ-E-Cad* RNA or linearized C-E-Cad overexpression (direct overexpression of C-E-Cad-ORF) (referred as to Vector-1, *circ-E-Cad*, Vector-2, C-E-Cad ORF, hereafter). β-actin was used as a loading control. Right: qPCR of the above modified GSCs using junction primers specific for *circ-E-Cad* RNA. $n=3$ independent experiments, two-sided *t*-test, $^{***}P<0.001$. **b**, The stem cell frequency of LDA analysis of GSC387 and GCSH2S cells with the indicated modifications. $n=3$ independent experiments, two-sided *t*-test, $^{***}P<0.001$. **c**, Percentage of EdU$^+$ GSC387 and GCSH2S cells with the indicated modifications in randomly selected microscopy fields at different time points. $n=3$ independent experiments, two-sided *t*-test, $^{***}P<0.001$. **d**, Brain slice invasion. Relative invasion depths of GSC387 and GCSH2S cells with the indicated modifications. $n=3$ independent experiments, two-sided *t*-test, $^{***}P<0.001$. **e**, Upper: representative BLI images of in vivo tumorigenicity assays using GSC387 cells with the indicated modifications. A total of 2,000 of each of the indicated cell types were intracranially injected into each mouse, five mice per group were used. Lower left: the tumour volume was determined by calculating the BLI fluorescence index. $n=5$ animals, two-sided *t*-test, $^{***}P<0.001$. Lower right: two-sided, log-rank analysis of mice intracranially implanted with GSC387 cells with indicated modifications, MST, middle survival time; $^{***}P<0.001$. **f**, Upper: representative BLI images of in vivo tumorigenicity assays using GSCH2S cells with the indicated modifications. $n=5$ animals, two-sided *t*-test, $^{***}P<0.001$. Lower left: the tumour volume was determined by calculating the BLI fluorescence index. Lower right: two-sided, log-rank analysis of mice intracranially implanted with GSCH2S cells with the indicated modifications. MST, middle survival time; $^{***}P<0.001$. In **b**–**f**, data are presented as the mean ± s.d. Source data are provided.

We used the ClusPro server[47] to perform molecular docking simulations of C-E-Cad to the CR2 domain. The Ramachandran plot indicated that the three-dimensional (3D) structure of the model was reasonable (Fig. 6e). Docking simulation revealed that residues in the unique C terminus of C-E-Cad were involved in binding to residues in the CR2 domain through salt bridge and hydrogen bond interactions (Fig. 6e and Extended Data Fig. 5l).

We performed surface plasmon resonance (SPR)[48] analysis using a recombinant CR2 domain, WT C-E-Cad and C-E-Cad-Δ14aa. Recombinant C-E-Cad effectively bound to the recombinant CR2



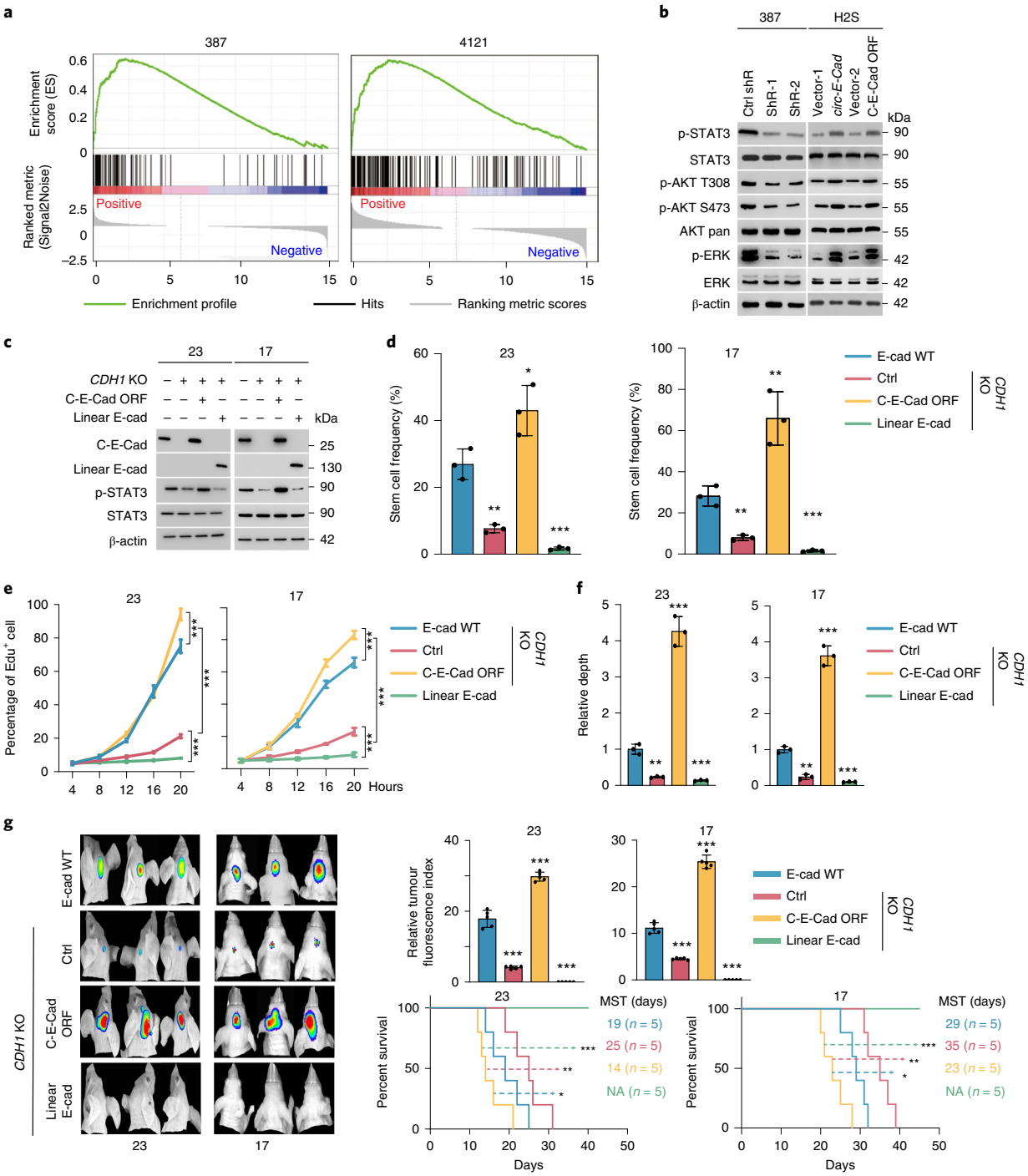


Fig. 4 | C-E-Cad promotes the activation of STAT3, PI3K–AKT and MAPK signalling in GSCs. a, Gene set enrichment analysis of the STAT3 directly regulated gene set in GSC387 and GSC4121 (4121) cells with stable circular *CDH1* RNA KD (versus controls) conducted using RNA-seq data (nominal *P* value of 0; false-discovery rate *q* value of 0). **b**, IB of p-STAT3, p-AKT and p-ERK expression levels in GSC387 and GCSH2S with the indicated modifications. $n = 3$ independent experiments. **c**, IB of the p-STAT3, E-cadherin and C-E-Cad expression levels in GSC23 (23)-WT, GSC23-*CDH1* KO, GSC17 (17)-WT and GSC17-*CDH1* KO cells. *circ-E-Cad* RNA or E-cadherin was re-expressed in the indicated cells. $n = 3$ independent experiments. **d**, The stem cell frequency of LDA of GSC23-WT, GSC23-*CDH1* KO, GSC17-WT and GSC17-*CDH1* KO cells. *circ-E-Cad* RNA or E-cadherin was re-expressed in the above cells as indicated (referred to as WT, KO + Ctrl, KO + C-E-Cad ORF, KO + Linear E-cad). $n = 3$ independent experiments, data are presented as the mean ± s.d., two-sided *t*-test. Left, $^{*}P = 0.018$, $^{**}P = 0.001$, $^{***}P < 0.001$. Right, $^{**}P = 0.001$ (KO + Ctrl), $^{**}P = 0.005$ (KO + C-E-Cad ORF), $^{***}P < 0.001$. **e**, Percentages of EdU^{+} cells in five randomly selected microscope fields at different time points in GSC23 and GSC17 with the indicated modifications. $n = 3$ independent experiments. Data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P < 0.001$. **f**, Brain slice invasion. The relative invasion depths of GSC23 and GSC17 cells with the indicated modifications. $n = 3$ independent experiments. Data are presented as the mean ± s.d., two-sided *t*-test, $^{**}P = 0.002$, $^{***}P < 0.001$. **g**, Left: in vivo tumorigenicity assay of GSC17 (WT, KO + Ctrl, KO + C-E-Cad ORF, KO + Linear E-cad). Right upper: tumour volumes as determined by the in vivo tumorigenicity assay and of mice bearing indicated GSC tumour xenografts ($n = 5$ animals). Data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P < 0.001$. Right lower: two-sided, log-rank analysis of mice ($n = 5$ animals). Left: $^{*}P = 0.018$, $^{**}P = 0.009$, $^{***}P < 0.001$. Right: $^{*}P = 0.017$, $^{**}P = 0.009$, $^{***}P < 0.001$. NA, not applicable. Source data are provided.

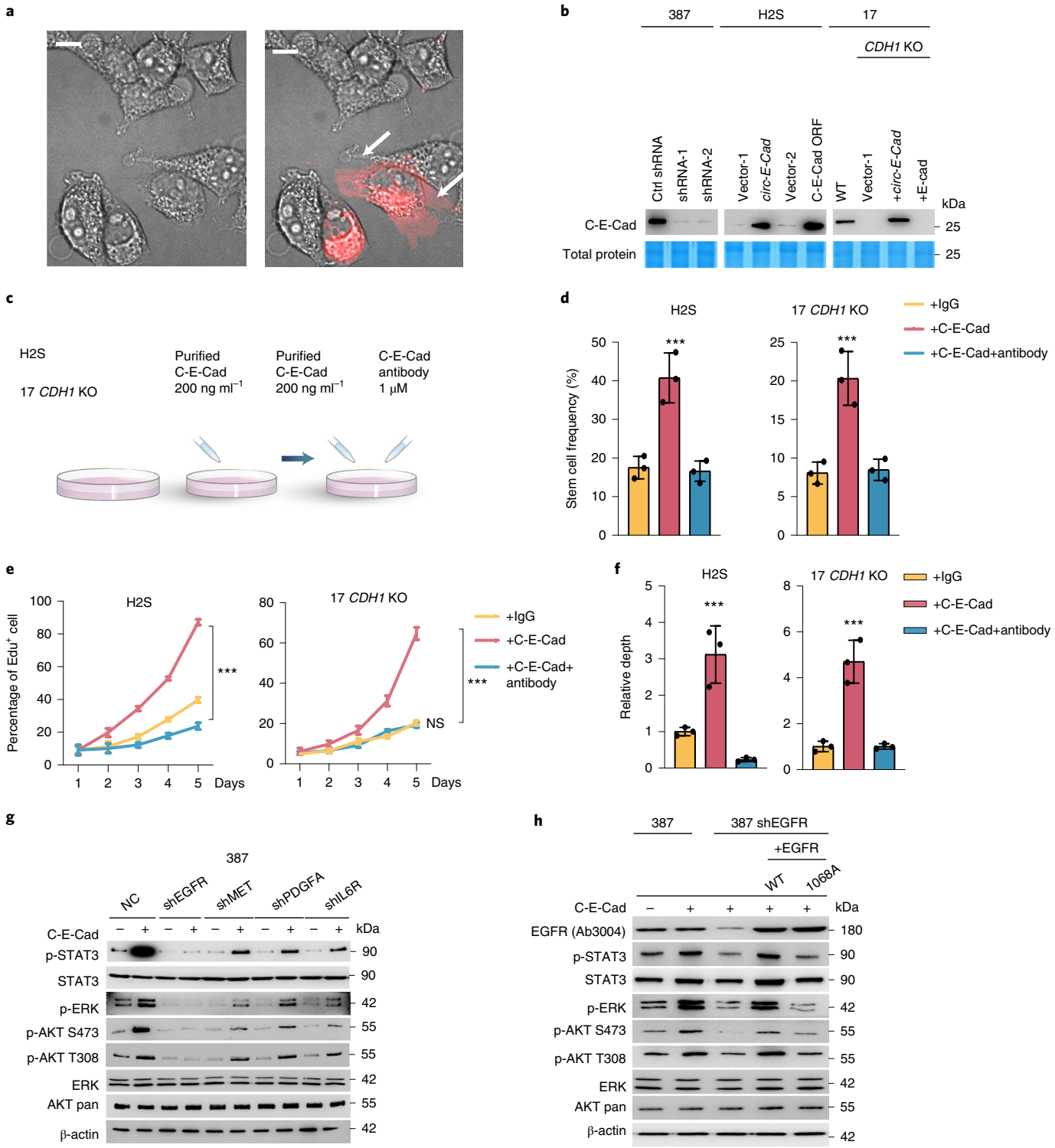


Fig. 5 | C-E-Cad is a secretory protein that activates STAT3 through EGFR. a, Live cell images of HEK293T cells transfected with C-E-Cad–RFP. The arrows indicate secretory C-E-Cad. $n=3$ independent experiments. Scale bar, 20 μm. **b**, IB of concentrated supernatant from GSC387, GSCH2S and GSC17 cells with the indicated modifications. Coomassie blue staining of total proteins was used as a loading control. $n=3$ independent experiments. **c**, Illustration of the purified C-E-Cad stimulation and antibody blocking strategy. **d**, The stem cell frequency of LDA of GSCH2S and GSC17 *CDH1* KO cells treated with purified C-E-Cad (200 ng ml^{-1}) or in combination with a neutralization antibody (1 μM, referred as to +IgG, +C-E-Cad, +C-E-Cad+antibody). $n=3$ independent experiments, data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P<0.001$. **e**, EdU assay analysis of GSCH2S and GSC17 *CDH1* KO cells with the indicated treatments. $n=3$ independent experiments. Data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P<0.001$. **f**, Brain slice invasion assay of GSCH2S and GSC17 *CDH1* KO cells with the indicated treatments. $n=3$ independent experiments. Data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P<0.001$. **g**, IB of p-STAT3, p-ERK and p-AKT expression in GSC387 cells with stable *EGFR*, *MET*, *PDGFRA* or *IL6R* knockdown separately and treated with or without recombinant C-E-Cad (200 ng ml^{-1}). $n=3$ independent experiments. NC, negative control shRNA. **h**, IB of EGFR, p-STAT3, p-AKT and p-ERK expression in GSC387 cells with stable *EGFR* KD plus re-expression of EGFR-WT or stable *EGFR* KD plus re-expression of EGFR-1068A. $n=3$ independent experiments. Source data are provided.

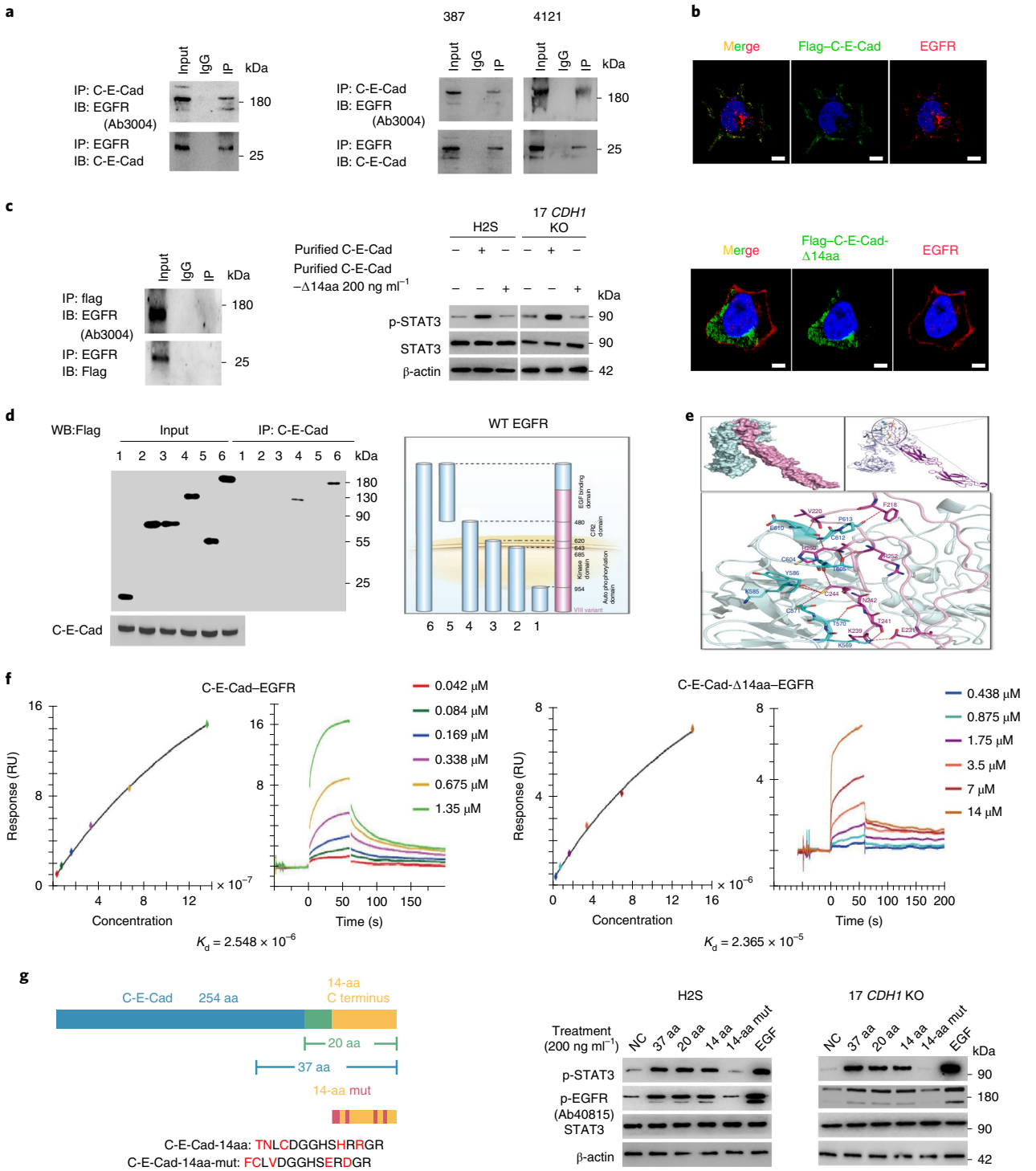


Fig. 6 | The unique C terminus of C-E-Cad and the CR2 domain of EGFR mediate the mutual interaction. a, Immunoprecipitation (IP) and IB. Left: recombinant C-E-Cad associates with EGFR in vitro. Right: C-E-Cad associates with EGFR in GSC387 and GSC4121 cells. $n=3$ independent experiments. **b**, IF images of the colocalization of EGFR, C-E-Cad and C-E-Cad-Δ14aa in HEK293T cells transfected with Flag–C-E-Cad and Flag–C-E-Cad-Δ14aa. Scale bar, 20 μm. $n=3$ independent experiments. **c**, Left: IP and IB of recombinant C-E-Cad-Δ14aa associating with EGFR in vitro. Right: IB of p-STAT3 and STAT3 expression in GSCH2S and GSC17 *CDH1* KO cells treated with purified C-E-Cad and C-E-Cad-Δ14aa (200 ng ml^{-1}). $n=3$ independent experiments. **d**, IP and IB of recombinant C-E-Cad associating with different His-tagged EGFR truncations in vitro. An illustration of the truncated EGFR domains is shown on the right. $n=3$ independent experiments. **e**, Molecular docking analysis of the C-E-Cad and EGFR CR2 domains. Cyan, EGFR CR2 domain; purple, C-E-Cad. **f**, SPR (BIAcore) analysis of the interaction between C-E-Cad (left) and C-E-Cad-Δ14aa (right) with the EGFR CR2 domain. The EGFR CR2 domain peptide was immobilized on a SPR sensor chip. The raw response (RU) curves (coloured curves) from a representative experiment were fitted to a one-site-specific kinetic model (black curves) to derive on and off rates and then calculated K_d values for interactions with C-E-Cad or C-E-Cad-Δ14aa. **g**, IB of p-STAT3, p-EGFR and STAT3 expression in GSCH2S and GSC17 *CDH1* KO cells treated with different synthetic polypeptides (200 ng ml^{-1}), as illustrated on the left. $n=3$ independent experiments. Source data are provided.

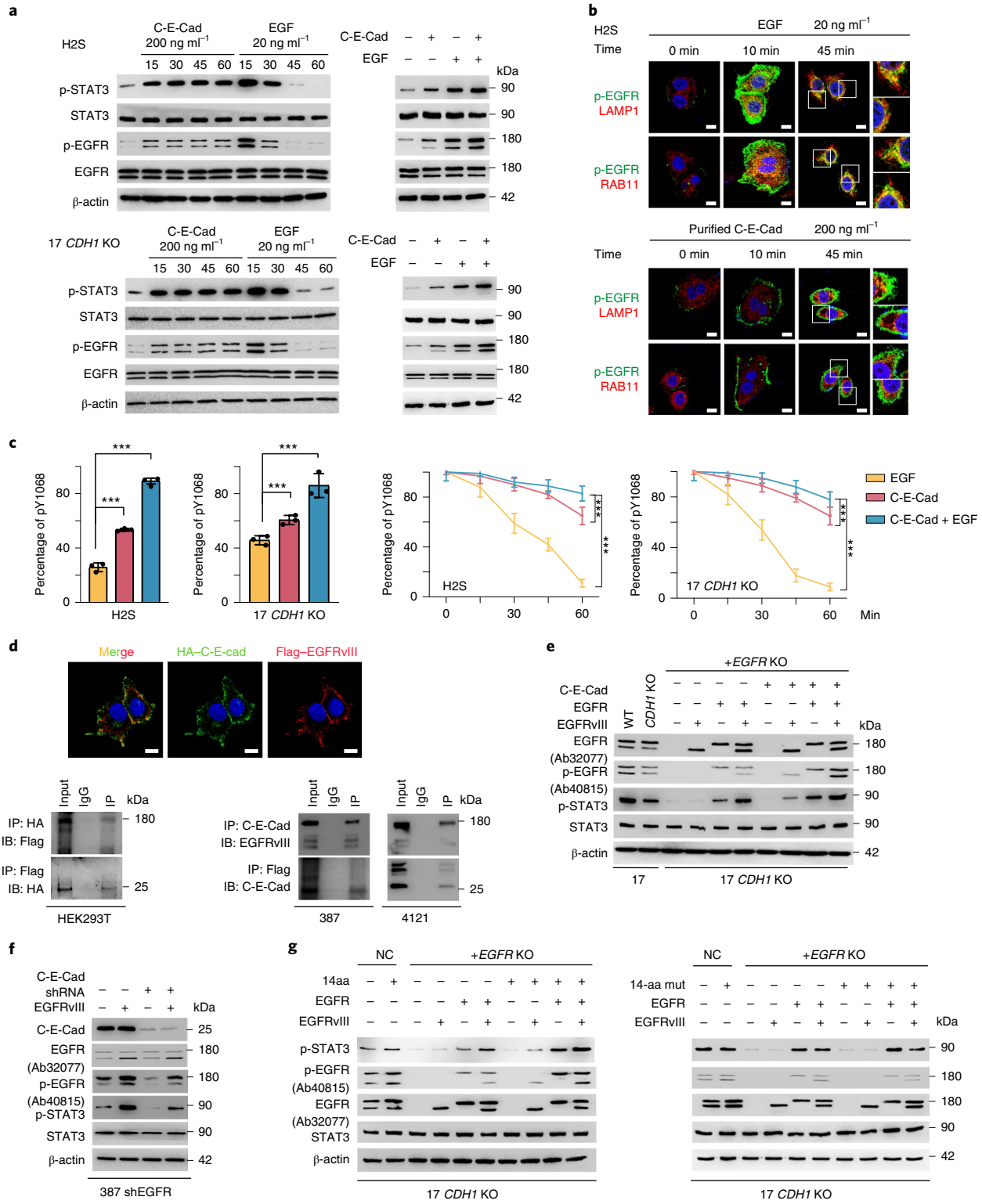


Fig. 7 | C-E-Cad independently activates EGFR and EGFRvIII. a, IB of p-STAT3, p-EGFR, STAT3 and EGFR expression in GSCH2S (upper) and GSC17 *CDH1* KO (lower) cells. **b**, IF of p-Y1068-EGFR and LAMP1 or RAB11 co-localization detected in GSCH2S cells treated with EGF or purified C-E-Cad at the indicated time points. Scale bar, 20 μm. **c**, p-Y1068-EGFR activation rate (left) and duration (right) in GSCH2S and GSC17 *CDH1* KO cells. Data are presented as the mean ± s.d., two-sided *t*-test, $***P < 0.001$. **d**, Upper: IF of the colocalization of Flag-tagged C-E-Cad and HA-tagged EGFRVIII in HEK293T cells. Scale bar, 20 μm. Lower: IP and IB of recombinant C-E-Cad associating with EGFRvIII in vitro (left) and C-E-Cad associating with EGFRvIII in GSC387 and GSC4121 cells (right). **e**, IB of p-STAT3, p-EGFR, STAT3 and EGFR expression in WT and *CDH1* KO cells together with *EGFR* KO and *EGFR* WT or EGFRvIII re-expressed in GSC17 cells. Recombinant C-E-Cad treatment (200 ng ml^{-1}) was added as indicated. **f**, IB of p-STAT3, p-EGFR, STAT3 and EGFR expression in EGFR stably KD GSC387 cells with C-E-Cad stable KD and re-expression of EGFRvIII. **g**, IB of p-STAT3, p-EGFR, STAT3 and EGFR expression in WT and *CDH1* KO together with EGFR WT and *EGFR* KO or EGFRvIII re-expressed GSC17 cells. The recombinant and synthetic C-E-Cad C-terminal 14-aa peptide or mutated peptide (200 ng ml^{-1}) was added as indicated. In **a**–**g**, $n = 3$ independent experiments. Source data are provided.

immobilized SPR chip, with a K_d of 2.548×10^{-6} M (Fig. 6f, left). C-E-Cad-Δ14aa displayed a markedly reduced ability (approximately tenfold reduction, K_d of 2.368×10^{-5} M) to interact with the recombinant CR2 (Fig. 6f, right). To further determine whether the C terminus of E-C-Cad is sufficient to stimulate EGFR signalling, we treated GSCH2S and GSC17 *CDH1* KO cells with several C-terminal peptide variants, which encompassed the sequences involved in the EGFR interaction (Fig. 6g, left). In both GSCs, the 14-aa, 20-aa and 37-aa peptides, but not the 14-aa peptide mutant, induced p-EGFR and p-STAT3 (Fig. 6g, right). Furthermore, re-expression of C-E-Cad, but not C-E-Cad Δ14aa, rescued tumorigenicity in vivo (Extended Data Fig. 5m–o).

C-E-Cad independently activates EGFR. Unlike C-E-Cad, EGF binds to the L1 and L3 domains of EGFR[49]. We treated GSCH2S and GSC17 *CDH1* KO cells with EGF, recombinant C-E-Cad or both. C-E-Cad was sufficient to induce p-EGFR and p-STAT3 in both GSCs, while combined treatment of C-E-Cad and EGF displayed a synergistic effect (Fig. 7a, left). Importantly, C-E-Cad was able to induce a prolonged activation of p-EGFR compared to that by EGF (Fig. 7a, right).

Internalization followed by recycle and degradation of the ligand-bound EGFR negatively regulates the intensity and duration of EGFR activation[50,51]. Since C-E-Cad did not affect EGFR expression, we next investigated whether C-E-Cad affected EGFR internalization. C-E-Cad-associated EGFR stayed on the cell membrane for 45 min after stimulation, which suggests that C-E-Cad retains EGFR on the cell surface for a prolonged period of time (Fig. 7b and Extended Data Fig. 6a). rC-E-Cad was able to stimulate EGFR activity more than that caused by EGF, while the combined treatment further elevated EGFR activity (Fig. 7c).

EGFRvIII is frequently amplified and co-expressed with EGFR in GBM[5]. We found that EGFRvIII was co-localized with C-E-Cad in GSCs (Fig. 7d). Since EGFRvIII was detected in various GSCs (Extended Data Fig. 6b), we knocked out *EGFR* in GSC17 *CDH1* KO cells and NSCs (Extended Data Fig. 6c–f) and re-expressed EGFR or EGFRvIII or both in these cells. KO of *EGFR* and *CDH1* markedly diminished p-STAT3. When these cells were cultured without EGF and C-E-Cad, Y1068 of EGFRvIII was not phosphorylated, and p-STAT3 levels were not elevated. C-E-Cad was able to stimulate p-EGFR and p-STAT3 in cells that only expressed EGFR or EGFRvIII. C-E-Cad displayed the strongest activation of p-EGFR and p-STAT3 in cells that co-expressed EGFR and EGFRvIII (Fig. 7e; Extended Data Fig. 6g). In EGFRvIII-expressing GSCs with stable *EGFR* KD, similar results were observed (Fig. 7f and Extended Data Fig. 6h). The 14-aa peptide stimulated *CDH1* + *EGFR* KO GSC17 cells and NSCs, in which EGFR or EGFRvIII or both were re-expressed. EGFR, EGFRvIII and the 14-aa peptide together exhibited the strongest activation of p-STAT3 (Fig. 7g and Extended Data Fig. 6i).

Targeting C-E-Cad enhances anti-EGFR therapy for inhibiting GSC tumorigenicity. In GBM, most of the C-E-Cad$^+$ cells harboured activated STAT3 signalling (Fig. 8a and Extended Data Fig. 7a). In patients positive for p-EGFR, those positive for C-E-Cad had shorter overall survival compared with those who were C-E-Cad$^-$ (Fig. 8b,c).

EGFR-targeting treatments for GBM have been disappointing[7,52]. We found that monotreatment with nimotuzumab failed to inhibit C-E-Cad-induced p-STAT3 in GSCs. An anti-C-E-Cad antibody displayed a moderate impact on the reduction of p-STAT3 in GSCs (Extended Data Fig. 7b). A combination of nimotuzumab and an anti-C-E-Cad antibody exhibited a synergistic effect on p-STAT3 inhibition and cell growth (Extended Data Fig. 7c), and this inhibition was also accomplished by robust KD of *EGFR* (Extended Data Fig. 7d,e). Lapatinib shows appreciable effects in EGFR mutated cancer cells[53]. Treatment with lapatinib or an anti-C-E-Cad partially inhibited p-STAT3 and cell growth, whereas a combination treatment with both achieved maximal inhibitory effects (Extended Data Fig. 7f,g). A previously reported oncogenic EGFR mutant, G598V, with higher intrinsic activity[54] rescued the inhibitory effects of anti-C-E-Cad; this result reinforces the notion that anti-C-E-Cad mainly blocks EGFR signalling (Extended Data Fig. 7h–j).

Next, we treated immunodeficient mice bearing GSC brain tumour xenografts with nimotuzumab and an anti-C-E-Cad antibody separately or in combination. When compared to the controls or nimotuzumab-treated animals, the anti-C-E-Cad antibody alone displayed an appreciable effect. Combination treatment with the anti-C-E-Cad antibody and nimotuzumab markedly suppressed the growth of GSC brain tumour xenografts and increased survival of the animals (Fig. 8d,f and Extended Data Fig. 8a,b). When these mice were treated with lapatinib and the anti-C-E-Cad antibody, separately or in combination, similar results were achieved (Extended Data Fig. 8c,d).

Discussion

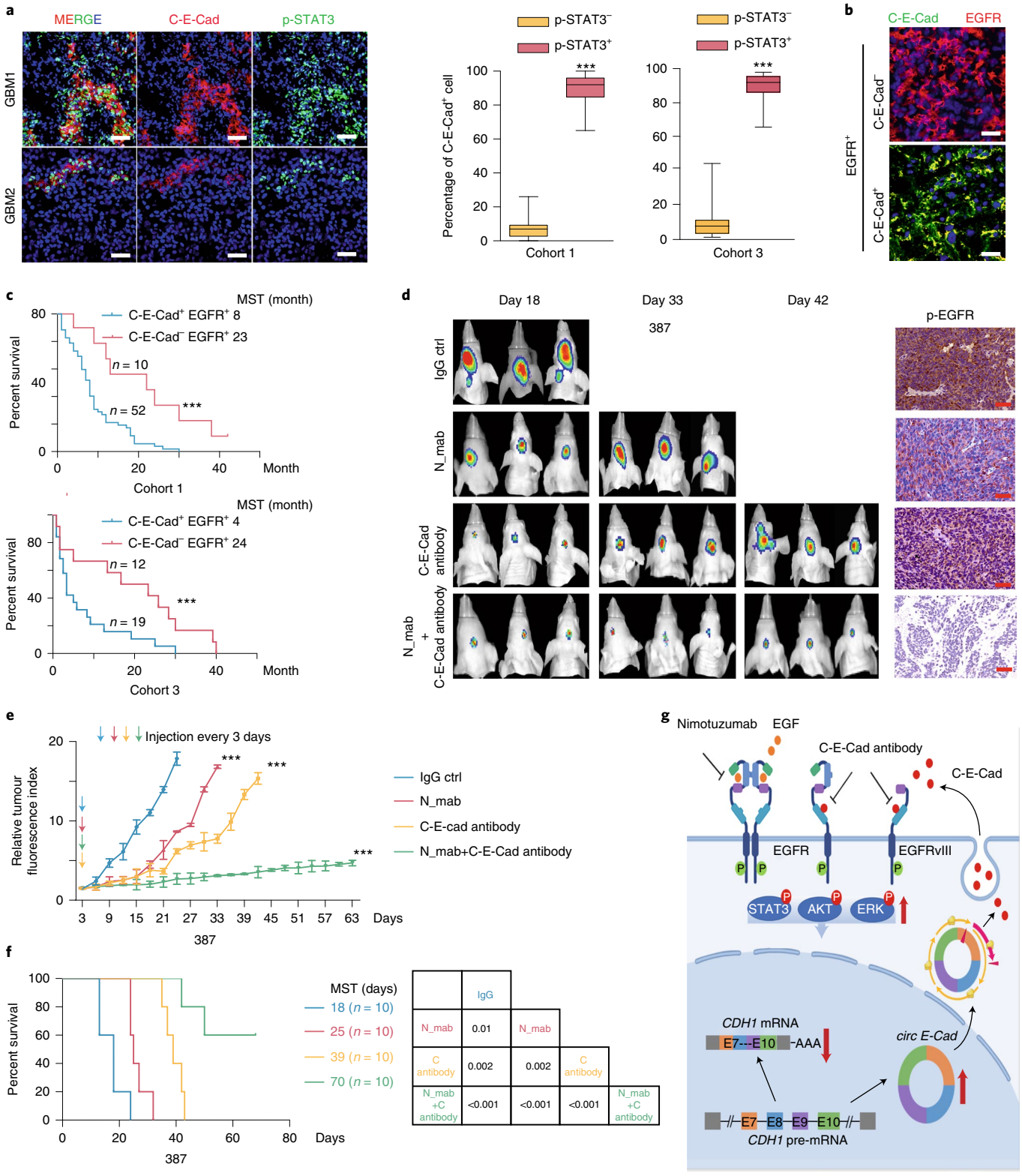
circRNAs have been described as generally downregulated transcripts in cancer[15]. However, a subset of circRNAs is also found upregulated in tumours, which indicates that circRNAs may be functionally diverse and that their clinical implications remain to be investigated[15]. In GBM, 468 circRNAs have been found to be overexpressed, and the levels of 21 circRNAs are inversely correlated with that of their cognate linear mRNAs[55]. Our evidence suggests that *circ-E-Cad* RNA is an oncogenic circRNA and an independent prognostic factor in GBM. Moreover, the upstream mechanisms governing the equilibrium between E-cadherin mRNA and *circ-E-Cad* RNA in GBM warrants further investigation.

We described that *circ-E-Cad* RNA encodes a unique E-cadherin variant. This translation pattern has been described in viruses[14] but not in mammalian cells. The C terminus provides an exceptional

Fig. 8 | Clinical implications and applications of C-E-Cad. a, Left: IF of p-STAT3 and C-E-Cad in GBM specimens from cohort 1 (GBM1) and cohort 3 (GBM2). Scale bar, 100 μm. Right: comparison of the percentages of C-E-Cad$^+$ cells among p-STAT3$^+$ cells in 50 randomly selected microscopy fields of each tumour image. $n = 50$ independent experiments. Data are presented as boxes containing the first and third quartiles. The whiskers indicate the maxima and minima. Wilcoxon test, $^{***}P < 0.001$. **b**, IF of GBM specimens classified via EGFR and C-E-Cad expression. $n = 3$ independent experiments. Scale bar, 100 μm. **c**, Two-sided, log-rank analysis of EGFR-expressing GBM stratified by both EGFR and C-E-Cad expression (upper, cohort 1, $n = 62$ biologically independent samples; lower, cohort 3, $n = 31$ biologically independent samples). In cohort 1, the median patient survival of EGFR$^+$/C-E-Cad$^+$ GBM was 8 months, while that of EGFR$^+$/C-E-Cad$^-$ GBM was 23 months. In cohort 3, the median patient survival of EGFR$^+$/C-E-Cad$^+$ GBM was 4 months, while that of EGFR$^+$/C-E-Cad$^-$ GBM was 24 months, $^{***}P < 0.001$. **d**, BLI images of mice bearing GSC tumour xenografts ($n = 10$ animals) and IHC of p-EGFR ($n = 3$ animals) brain sections with tumour xenografts treated with different therapeutic strategies. Scale bar, 250 μm. **e**, The C-E-Cad antibody therapeutic strategy in an animal model. For anti-EGFR therapy, nimotuzumab (N_mab; 1 μg μl^{-1}, 3 μl every 3 days) was used. The C-E-Cad antibody (1 μg μl^{-1}, 3 μl every 3 days) was used alone or in combination with anti-EGFR therapy. The total intracranial injection volume was 3 μl at each time. $n = 10$ animals. Data are presented as the mean ± s.d., two-sided *t*-test, $^{***}P < 0.001$. **f**, Two-sided, log-rank analysis of mice treated with the indicated therapeutic strategies. $n = 10$ animals. The *P* values are shown in the table. **g**, Summary figure of the mechanism of *circ-E-Cad*. *circ-E-Cad* RNA (generated from exon 7–9 of *CDH1*) encodes C-E-Cad. C-E-Cad interacts with the CR2 domain of EGFR and activates downstream STAT3, AKT and ERK signalling, thus sustaining tumour growth and survival. Source data are provided.

target for effective anticancer therapy. In a previous study[16], circPOK and POKEMON were shown to have antithetical functions. However, antithetical proteins generated from circular/linear RNA of the same gene have not yet been described. We showed that alternative translation of the same gene results in distinct protein variants with antithetical functions and that a dysregulated equilibrium between circular and linear variants profoundly affects tumorigenicity.

Finally, we identified C-E-Cad as a new ligand that activates the oncogenic EGFR signalling pathway by directly binding to and persistently activating EGFR signalling. Previous studies reported that a *PTEN* mutation mediated resistance to EGFR kinase inhibitors[56,57]. A combination of an anti-C-E-Cad antibody and lapatinib repressed PTEN mutant GSC tumorigenicity in vivo, which highlights that anti-C-E-Cad treatments could enhance the efficiency of EGFR-targeting therapy. Although an optimized delivery system



and small-molecule inhibitors are attractive approaches to pursue, our findings hold particular promise for antibody development for innovative clinical translation due to the exclusive 14-aa sequence and expression pattern of C-E-Cad. We hope to develop clinical protocols of combinations of anti-C-E-Cad antibody and EGFR for treating recurrent GBMs with EGFR overexpression.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-021-00639-4.

Received: 4 May 2020; Accepted: 24 January 2021;
Published online: 04 March 2021

References

1. Reifenberger, G., Wirsching, H. G., Knobbe-Thomsen, C. B. & Weller, M. Advances in the molecular genetics of gliomas—implications for classification and therapy. *Nat. Rev. Clin. Oncol.* **14**, 434–452 (2017).
2. Aldape, K. et al. Challenges to curing primary brain tumours. *Nat. Rev. Clin. Oncol.* **16**, 509–520 (2019).
3. Brennan, C. W. et al. The somatic genomic landscape of glioblastoma. *Cell* **155**, 462–477 (2013).
4. Neftel, C. et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* **178**, 835–849.e21 (2019).
5. Furnari, F. B., Cloughesy, T. F., Cavenee, W. K. & Mischel, P. S. Heterogeneity of epidermal growth factor receptor signalling networks in glioblastoma. *Nat. Rev. Cancer* **15**, 302–310 (2015).
6. Thorne, A. H., Zanca, C. & Furnari, F. Epidermal growth factor receptor targeting and challenges in glioblastoma. *Neuro Oncol.* **18**, 914–918 (2016).
7. Westphal, M., Maire, C. L. & Lamszus, K. EGFR as a target for glioblastoma treatment: an unfulfilled promise. *CNS Drugs* **31**, 723–735 (2017).
8. Prados, M. D. et al. Phase II study of erlotinib plus temozolomide during and after radiation therapy in patients with newly diagnosed glioblastoma multiforme or gliosarcoma. *J. Clin. Oncol.* **27**, 579–584 (2009).
9. Guo, J. U., Agarwal, V., Guo, H. & Bartel, D. P. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* **15**, 409 (2014).
10. Salzman, J., Chen, R. E., Olsen, M. N., Wang, P. L. & Brown, P. O. Cell-type specific features of circular RNA expression. *PLoS Genet.* **9**, e1003777 (2013).
11. Memczak, S. et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**, 333–338 (2013).
12. You, X. et al. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. *Nat. Neurosci.* **18**, 603–610 (2015).
13. Kristensen, L. S. et al. The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* **20**, 675–691 (2019).
14. AbouHaidar, M. G., Venkataraman, S., Golshani, A., Liu, B. & Ahmad, T. Novel coding, translation, and gene expression of a replicating covalently closed circular RNA of 220 nt. *Proc. Natl Acad. Sci. USA* **111**, 14542–14547 (2014).
15. Vo, J. N. et al. The landscape of circular RNA in cancer. *Cell* **176**, 869–881. e13 (2019).
16. Guarnerio, J. et al. Intragenic antagonistic roles of protein and circRNA in tumorigenesis. *Cell Res.* **29**, 628–640 (2019).
17. Zhang, M. et al. A peptide encoded by circular form of *LINC-PINT* suppresses oncogenic transcriptional elongation in glioblastoma. *Nat. Commun.* **9**, 4475 (2018).
18. Glazar, P., Papavasileiou, P. & Rajewsky, N. circBase: a database for circular RNAs. *RNA* **20**, 1666–1670 (2014).
19. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* **7**, 1534–1550 (2012).
20. van Heesch, S. et al. The translational landscape of the human heart. *Cell* **178**, 242–260.e29 (2019).
21. Alvarado, A. G. et al. Coordination of self-renewal in glioblastoma by integration of adhesion and microRNA signaling. *Neuro Oncol.* **18**, 656–666 (2016).
22. Bao, S. et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756–760 (2006).
23. Suzuki, H. et al. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* **34**, e63 (2006).
24. Schneider, T., Schreiner, S., Preusser, C., Bindereif, A. & Rossbach, O. Northern blot analysis of circular RNAs. *Methods Mol. Biol.* **1724**, 119–133 (2018).
25. Wang, R. et al. Glioblastoma stem-like cells give rise to tumour endothelium. *Nature* **468**, 829–833 (2010).
26. Zhang, M. et al. A novel protein encoded by the circular form of the *SHPRH* gene suppresses glioma tumorigenesis. *Oncogene* **37**, 1805–1814 (2018).
27. Pamudurti, N. R. et al. Translation of circRNAs. *Mol. Cell* **66**, 9–21.e7 (2017).
28. Bulstrode, H. et al. Elevated FOXG1 and SOX2 in glioblastoma enforces neural stem cell identity through transcriptional control of cell cycle and epigenetic regulators. *Genes Dev.* **31**, 757–773 (2017).
29. Chuang, T. J. et al. Integrative transcriptome sequencing reveals extensive alternative *trans*-splicing and *cis*-backsplicing in human cells. *Nucleic Acids Res.* **46**, 3671–3691 (2018).
30. Zhang, X. O. et al. Complementary sequence-mediated exon circularization. *Cell* **159**, 134–147 (2014).
31. Zhang, J. X. et al. Unique genome-wide map of *TCF4* and *STAT3* targets using ChIP-seq reveals their association with new molecular subtypes of glioblastoma. *Neuro Oncol.* **15**, 279–289 (2013).
32. Mogil, J. S. Stats: multiple experiments test biomedical conclusions. *Nature* **569**, 192 (2019).
33. Levy, D. E. & Lee, C. K. What does Stat3 do? *J. Clin. Invest.* **109**, 1143–1148 (2002).
34. Bleau, A. M. et al. PTEN/PI3K/Akt pathway regulates the side population phenotype and ABCG2 activity in glioma tumor stem-like cells. *Cell Stem Cell* **4**, 226–235 (2009).
35. Wang, X. et al. MYC-regulated mevalonate metabolism maintains brain tumor-initiating cells. *Cancer Res.* **77**, 4947–4960 (2017).
36. Day, B. W. et al. The dystroglycan receptor maintains glioma stem cells in the vascular niche. *Acta Neuropathol.* **138**, 1033–1052 (2019).
37. Qin, M. et al. Hsa_circ_0001649: a circular RNA and potential novel biomarker for hepatocellular carcinoma. *Cancer Biomark.* **16**, 161–169 (2016).
38. Hu, Q. P., Kuang, J. Y., Yang, Q. K., Bian, X. W. & Yu, S. C. Beyond a tumor suppressor: soluble E-cadherin promotes the progression of cancer. *Int. J. Cancer* **138**, 2804–2812 (2016).
39. Soncin, F. et al. Abrogation of E-cadherin-mediated cell–cell contact in mouse embryonic stem cells results in reversible LIF-independent self-renewal. *Stem Cells* **27**, 2069–2080 (2009).
40. Zanca, C. et al. Glioblastoma cellular cross-talk converges on NF-κB to attenuate EGFR inhibitor sensitivity. *Genes Dev.* **31**, 1212–1227 (2017).
41. Vitucci, M. et al. Cooperativity between MAPK and PI3K signaling activation is required for glioblastoma pathogenesis. *Neuro Oncol.* **15**, 1317–1329 (2013).
42. Akhavan, D. et al. De-repression of PDGFRβ transcription promotes acquired resistance to EGFR tyrosine kinase inhibitors in glioblastoma patients. *Cancer Discov.* **3**, 534–547 (2013).
43. Jones, S. A. & Jenkins, B. J. Recent insights into targeting the IL-6 cytokine family in inflammatory diseases and cancer. *Nat. Rev. Immunol.* **18**, 773–789 (2018).
44. Aldape, K., Zadeh, G., Mansouri, S., Reifenberger, G. & von Deimling, A. Glioblastoma: pathology, molecular mechanisms and markers. *Acta Neuropathol.* **129**, 829–848 (2015).
45. Downward, J., Parker, P. & Waterfield, M. D. Autophosphorylation sites on the epidermal growth factor receptor. *Nature* **311**, 483–485 (1984).
46. Ferguson, K. M. Structure-based view of epidermal growth factor receptor regulation. *Annu. Rev. Biophys.* **37**, 353–373 (2008).
47. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Res.* **32**, W96–W99 (2004).
48. Lee, T. H., Hirst, D. J., Kulkarni, K., Del Borgo, M. P. & Aguilar, M. I. Exploring molecular–biomembrane interactions with surface plasmon resonance and dual polarization interferometry technology: expanding the spotlight onto biomembrane structure. *Chem. Rev.* **118**, 5392–5487 (2018).
49. Garrett, T. P. et al. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* **110**, 763–773 (2002).
50. Wells, A. et al. Ligand-induced transformation by a noninternalizing epidermal growth factor receptor. *Science* **247**, 962–964 (1990).
51. Caldieri, G. et al. Reticulon 3-dependent ER–PM contact sites control EGFR nonclathrin endocytosis. *Science* **356**, 617–624 (2017).
52. Mellinghoff, I. K. et al. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N. Engl. J. Med.* **353**, 2012–2024 (2005).
53. Vivanco, I. et al. Differential sensitivity of glioma- versus lung cancer-specific *EGFR* mutations to EGFR kinase inhibitors. *Cancer Discov.* **2**, 458–471 (2012).

54. Lee, J. C. et al. Epidermal growth factor receptor activation in glioblastoma through novel missense mutations in the extracellular domain. *PLoS Med.* **3**, e485 (2006).
55. Song, X. et al. Circular RNA profile in gliomas revealed by identification tool UROBORUS. *Nucleic Acids Res.* **44**, e87 (2016).
56. Mellinghoff, I. K., Cloughesy, T. F. & Mischel, P. S. PTEN-mediated resistance to epidermal growth factor receptor kinase inhibitors. *Clin. Cancer Res.* **13**, 378–381 (2007).
57. Fenton, T. R. et al. Resistance to EGF receptor inhibitors in glioblastoma mediated by phosphorylation of the PTEN tumor suppressor at tyrosine 240. *Proc. Natl Acad. Sci. USA* **109**, 14164–14169 (2012).

Methods

Experimental model and subject details. *Mice and animal housing.* Athymic (Ncr nu/nu) female mice at 6–8 weeks of age were purchased from Nanjing University Farms. Five mice were grouped in each cage. All experiments using animals were conducted using approved protocols by the Institutional Animal Care and Use Committee at Sun Yat-sen University in accordance with US NIH and institutional guidelines. The study was compliant with all relevant ethical regulations regarding animal research.

Xenograft studies. Mice were randomly assigned to experimental groups for all the experiments. For the animal survival analysis, mice were intracranially injected with 2,000 GSCs prelabelled with Fluc in 5 µl PBS and maintained until pathological symptoms from tumour burden developed or 70 days after injection. Tumour volume was estimated by calculating the relative Fluc index.

For the intracranial tumour xenograft model treated with nimotuzumab or the anti-C-E-Cad antibody (using a microsyringe guided by stereotactic system), mice intracranially receiving GSCs were randomized into four groups: (1) IgG control (3 µl, 3 µg); (2) anti-C-E-Cad antibody (3 µl, 3 µg); (3) anti-EGFR antibody (nimotuzumab, 3 µl, 3 µg); and (4) nimotuzumab and anti-C-E-Cad antibody (3 µl, 3 µg per antibody) co-administered at every injection point. For the combination of apatinib and the C-E-Cad antibody, the same dose as nimotuzumab was used. All mice were monitored every 3 days for the development of neurological symptoms due to tumour burden. The growth of intracranial GSC tumour xenografts was monitored using bioluminescence imaging (BLI) with an IVIS machine (PhotoSound PAFT/256). Mice were maintained until the development of neurological symptoms. For in vivo BLI assessments, mice were injected with luciferin, and the fluorescence intensity was determined. The BLI images were taken using a Xenogeny imaging system (PhotoSound PAFT/256).

Human GBM and paired adjacent specimen. Human GBM tumour tissues were obtained after confirmation by board-certified neuropathologists from the following surgical suites: the Department of Neurosurgery, The First Affiliated Hospital of Sun Yat-sen University; the Department of Neurosurgery, Sun Yat-sen University Cancer Centre; and the Department of Neurosurgery, First Affiliated Hospital of Nanjing Medical University. Informed and signed content were obtained from each patient, and the study was approved by the Ethics Institutional Review Boards of The First Affiliated Hospital of Sun Yat-sen University, the Sun Yat-sen University Cancer Centre, the First Affiliated Hospital of Nanjing Medical University and the Beijing Neurosurgical Institute, Capital University. The samples were de-identified before being subjected to the described experiments in this study. GBM samples enrolled from the CGGA (Chinese Glioma Genome Atlas) were from the Beijing Neurosurgical Institute, Capital University. The study was compliant with all relevant ethical regulations regarding research involving human participants.

Cell lines and cell culture. The GSC456, GSC4121, GSC387, GSC3691 and GSCH2S cells were gifts from the UCSD. The GSC17 and GSC23 cell lines were generated using standard procedures[22]. GSCs were cultured as glioma tumour spheres in DMEM/F12 medium supplemented with B27 supplement (Life Technologies), bFGF and EGF (20 ng ml^{-1} each). Induced pluripotent stem cell (iPSC)-derived NSCs were a gift from P. Xiang, Sun Yat-sen University. Human NSCs (primary) were obtained from Gibco (TM A15654) and cultured with StemPro NSC SFM (A10509-01) supplemented with 2 mM GlutaMAX-I supplement (35050), 6 U ml^{-1} heparin (Sigma, H3149) and 200 µM ascorbic acid (Sigma, A8960). Human NHAs were obtained from Lonza/ThermoFisher. HEK293T cells (from The American Type Culture Collection) were cultured in DMEM with 10% fetal bovine serum.

Techniques. *RNA-seq.* Total RNA was extracted using a Trizol reagent kit (Invitrogen). After digestion, purification and reverse transcription, the segments were subjected to Illumina sequencing adapters. Candidate circRNAs were selected if the unique junction reads were more than 2. Reads per million mapped reads (RPKM) were applied to quantify circRNAs. circRNAs with a fold change of >2 and $P < 0.01$ were identified as DEcRs. For linear RNA expression analysis, data were mapped to the reference genome using TopHat2 (v.2.1.1), then transcript abundance was quantified using the software RSEM (v.1.2.19).

Gene set enrichment analysis. For the correlation between C-E-Cad and STAT3, two treatments (a negative control and shC-E-Cad) were conducted in GSC387 and GSC4121 cells in parallel. The intersection of STAT3-binding genes and STAT3 positively regulated genes were obtained from GSC387 and GSC4121 cells, respectively. These genes were considered as STAT3 directly regulated gene sets in GSC4121 or GSC387 cells. Gene set enrichment analysis was conducted using the parameters from signal2Noise to analyse the enrichment of STAT3 directly regulated gene sets in comparison to shC-E-Cad treatment versus the negative control.

Ribosomal profiling sequencing. The ribosomal profiling technique was carried out as previously described[19]. After obtaining ribosome footprints, ribosomal profiling libraries were constructed using a NEBNext Multiple Small RNA Library Prep set for Illumina (E7300S, E7300L). Briefly, adapters were added to both ends of ribosome frames (RFs), followed by reverse transcription and PCR amplification. The 140–160-bp sized PCR products were enriched to generate a complementary DNA library and sequenced using Illumina HiSeqTM X10.

Metagene analysis. Low-quality reads were filtered using FASTP[19]. The short reads alignment tool Bowtie2 was used for mapping reads to a ribosome RNA and transfer RNA database. The reads mapped to ribosome RNA and transfer RNA were removed. The retained reads from each sample were mapped to the reference genome using Bowtie2, with no mismatches allowed. To visualize the RFs surrounding the start and stop codons of metagenes, reads were counted at each position of each gene, then these counts were summed across all genes. Metagene plots were generated using R (https://www.r-project.org/), taking the read counts and the CDS (coding sequence) boundaries in the transcript coordinates as input.

Identification of junction ribosome sequencing reads on circRNA. circRNA data were download from the circBase database (http://www.circbase.org/). The ribosome sequencing reads from junction sites of circRNA were searched following the methods described[58,59]. The number of junction reads that indicated a translation signal was calculated for each circRNA.

Plasmid construction and transfection. The *C-E-Cad* ORF expression plasmid was generated by cloning the linearized full-length ORF of *C-E-Cad*. The *circ-E-Cad* expression plasmid was generated by cloning the sequence of exons 7–10 of *CDH1* using pCDH-CMV-MCS-EF1–GFP + Puro (SBI pCD513B-1). An additional circulation promoter sequence and an AG/GT splicing sequence were added 83-bp upstream and 53-bp downstream. The C-E-Cad–RFP expression plasmid was generated by cloning the ORF fused to red fluorescent protein (RFP). Other plasmids were generated according to the key resource table (Suplementary Table 11). Plasmids were transfected using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's protocol.

Lentiviral production and establishment of stable cell lines. Lentiviral vectors expressing *circ-E-Cad*, *C-E-Cad* ORF, sh*C-E-Cad*, *C-E-Cad*–RFP, *STAT3*-705E, *EGFR*-WT, *EGFRvIII*, *EGFR*-1068A or *EGFR* truncated mutations were co-transfected with the packaging vectors psPAX2 (Addgene) and pMD2G (Addgene) into HEK293T cells for lentivirus production using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. To establish stable cell lines, GSCs were transduced by using the above lentiviruses with polybrene (8 mg ml^{-1}, Sigma). After incubating for 72 h, cells were selected with 2 mg ml^{-1} puromycin for 3 days.

CRISPR–Cas9-mediated gene KO. The target sequences of guide RNA (gRNA) were designed using the online tool at http://crispr.mit.edu/. To produce CRISPR lentivirus, HEK293T cells seeded in 100-mm plates were transfected with 10 µg lentiCRISPRv2 gRNA or lentiCRISPRv2 control[59] (Addgene, plasmid 52961) plasmids, 5 µg psPAX$_2$ and 2.5 µg PMD$_2$G plasmids using Lipofectamine 2000 according to the manufacturer's instructions. After incubation for 48–72 h, the supernatants containing lentivirus were collected and used to infect GSCs for 4–6 h in the incubator. After 5 days of culture. KO stable cell lines were collected and characterized.

Limiting dilution assay. In vitro limiting dilution assays (LDAs) were performed according to a previously described protocol[60]. Glioma sphere-forming frequency was calculated using extreme LDA software (http://bioinf.wehi.edu.au/software/elda/).

Immunohistochemistry. As previously described[61], all paraffin-embedded GBM tumour sections (8–10 µm thick) were deparaffinized and blocked. Primary antibodies were diluted in bovine serum albumin and applied overnight at 4 °C in a wet chamber. After incubation in secondary antibodies, diaminobenzidine reagent was added to these tumour sections, which were then counterstained with haematoxylin to visualize nuclei.

We quantitatively scored the tumour tissue sections according to the percentage of positive cells and staining intensity. The scores were multiplied to give a scoring range of 0–12, whereby 0–6 was considered low expression and 7–12 was considered overexpression.

RNA fluorescence in situ hybridization. Cells were incubated at 37 °C in a solution containing 50% formamide, 2× SSC, 0.25 mg ml^{-1} *Escherichia coli* transfer RNA, 0.25 mg ml^{-1} salmon sperm DNA (Life Technologies), 2.5 mg ml^{-1} BSA (Roche) and 125 nM fluorescently labelled junction probe (Generay). After 12 h, the cells were washed and mounted in ProLong Gold (Life Technologies) and incubated overnight at room temperature. Confocal microscopy imaging (Olympus FV100) was then performed.

Northern blotting. Approximately 20 µg of total RNA was extracted and separated by 1.2% agarose gel. After transfer to a membrane and fixed, specific probes were applied at 37 °C and washed with 0.1% SDS at room temperature. Data were analysed using Quantity One or Image Lab software (Bio-Rad).

Reverse transcription and real-time PCR. Total RNA samples were extracted and reverse-transcribed. The resulting cDNA was then subjected to real-time PCR analysis with SYBR Select master mix (Thermo Fisher Scientific) in a StepOne Plus real time PCR system (Applied Biosystems). Results were normalized to the β-actin mRNA in each sample.

Dual-luciferase reporter system. The renilla luciferase (Rluc) and the firefly luciferase (Luc) sequences were amplified from a psicheck2 vector (Promega). The Rluc sequence was placed in front and the Luc sequence was placed in the back. The full-length sequences of Rluc and Luc were obtained by overlapping PCR, and the flank sequences were connected to pCDNA 3.1(+) vector using two restriction enzyme sites (NheI and XhoI). The potential IRES sequences of *circ-E-Cad* were amplified and inserted in the middle of Rluc and Luc using two restriction enzyme sites (KpnI and EcoRI) introduced by primers. The IRES of hepatitis C virus was applied as a positive control.

Immunoblotting. Equal loading of cell lysate and tissue lysate were added to each well of a 12% SDS–PAGE system. After electrophoresis, membrane transfer and blocking, membranes were incubated with indicated primary antibodies and secondary horseradish-peroxidase-tagged antibodies, and the signals were visualized by enhanced chemiluminescence.

Immunofluorescence. Cultured cells and GSCs were fixed with 4% formaldehyde (Fisher) for 10 min and then blocked with 5% BSA with 0.1% Triton X-100 in PBS for 30 min at room temperature. Immunostaining was performed using the appropriate primary and secondary antibodies. Nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI). Images were taken using an Olympus FV1000 microscope.

Signalling of human EGFR. To analyse ligand-induced EGFR activation, cells were washed with ice-cold binding buffer (10 mM HEPES, 150 mM NaCl, 1% BSA, pH 8.0) and stimulated with C-E-Cad and the control in this buffer for 10 min on ice. Cells were quickly lysed in binding buffer containing 1% NP-40, 1 mM phenylmethylsulfonyl fluoride, 1 mg ml^{-1} aprotinin, 1 mg ml^{-1} leupeptin, 5 mM sodium orthovanadate with phosphatase inhibitor cocktail (Thermo Fisher Scientific), and lysis supernatants were subjected to immunoblotting with 3 mg ml^{-1} anti-EGFR ab-10 (Lab Vision) and 1:500 diluted anti-phospho-tyrosine p-Y20 (Santa Cruz Biotechnology), with detection using a LI-COR Odyssey Fc instrument.

EGFR activation. Quantification of phosphorylated and total EGFR was performed using LI-COR Image Studio software. The ratio of these signal intensities (calculated as phosphorylated EGFR divided by total EGFR) at each ligand concentration was determined, and the background value from the unstimulated sample subtracted. Data are plotted as log[ligand] versus response using GraphPad Prism, from which the maximum response for each experiment was determined. Results for each concentration were then normalized by the maximum response for the relevant experiment, and values of mean response ± s.d. are plotted.

Liquid chromatography–MS/MS analysis. Total protein was collected and separated by 12% SDS gel, and the band at about 25 kDa was excised and subjected to digestion. The resulting peptide was analysed using a QExactive mass spectrometer coupled to a nano-LC (AdvanceLC). The acquired spectra were analysed using the SEQUEST HT algorithm.

Edu assay. For the 5-ethynyl-2′-deoxyuridine (Edu) assays, cells were co-cultured with an Edu-labelling reagent following the manfacturer's protocol. Five fields of view were taken for each cell line. Images were taken using an Olympus FV1000 microscope.

CCK-8 assay. Two hundred of the indicated cells were seeded into each well of a 96-well plate. The viability of cells was determined using a CCK-8 reagent (Dojindo) every 24 h by measuring the absorbance at 450 nm (BioTek) of different cell lines following the manufacturer's instructions. Treatment conditions are described in each figure legend.

Live-cell imaging. HEK293T cells were transfected with RFP-tagged C-E-Cad. After transfection for 72 h, live cell images (red fluorescence and bright-field) were taken using an Olympus FV100 microscope.

Immunoprecipitation. HEK293T cells were transfected with different plasmids for 72 h. Cells were lysed in ice-cold lysis buffer (0.3% CHAPS, 10 mM β-glycerol phosphate, 10 mM pyrophosphate, 40 mM HEPES (pH 7.4), 2.5 mM $MgCl_2$ and EDTA-free protease inhibitor). The soluble fractions from cell lysates were immunoprecipitated using primary antibodies. Then, the mix was rotated overnight at 4 °C. Beads were added and incubated for 2 h at room temperate. Immunoprecipitates were washed three times with PBST and subjected to immunoblotting analyses.

Brain slice invasion assay. As previously described[62,63], the indicated number of cells labelled with green fluorescent protein (GFP) were seeded into non-adherent dishes in medium containing 20% methylcellulose (Sigma, M7140). After 8 h of incubation, cells assembled as neurospheres or glioma tumour spheres. We collected the medium and separated the spheres and planted the spheres onto a fresh mouse brain slice. After incubating for 12 h, the depth of neurosphere invasion into the brain slice was measured. Five spheres were applied for each cell line. Images were taken using a confocal microscope (Olympus FV100).

SPR analysis. A BIAcore S200 instrument (GE Healthcare) was used to detect binding interactions using a direct-binding assay format. SPR equilibrium binding data, consisting of Req values from several concentration series, were analysed by fitting a simple 1:1 binding to yield R_{max} and K_d values using BIAcore S200 Evaluation software.

Molecular docking. The 3D structure of the EGFR protein was downloaded from RCSB Protein Data Bank (PDB ID: 4UV7). Protein–protein docking in ClusPro server4-8 (https://cluspro.org) was used for molecular-docking simulations of C-E-Cad and for predicting the binding affinity to EGFR. Molecular graphics were generated using PyMOL.

Statistics and reproducibility. Statistical analysis was carried out using Microsoft Excel 2013 and GraphPad Prism v.5.00 for Windows. Experimental data are represented as the average ± s.d. of a minimum of three biological replicates. Unless otherwise indicated, Student's two-tailed unpaired *t*-test was used to determine statistical significance of in vitro experiments. Gehan–Breslow–Wilcoxon test or log-rank test was used to determine the statistical differences of the survival data. All statistical tests were two-sided, and a *P* value of less than 0.05 was considered statistically significant. For each experiment, data are representative of three replications, with similar results obtained.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The RNA-seq data that support the findings of this study have been deposited in the NCBI with the identifiers PRJNA525736, SRA714646 and PRJNA525736. The human GBM data were derived from the TCGA Research Network: http://cancergenome.nih.gov/. The dataset derived from this resource that supports the findings of this study is available at http://gepia.cancer-pku.cn/detail.php?gene. The CGGA data were derived from the 301 dataset available at http://www.cgga.org.cn/download.jsp. All other data supporting the findings of this study are available from the corresponding authors upon reasonable request. Source data are provided with this paper.

References

58. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
59. Salzman, J., Gawad, C., Wang, P. L., Lacayo, N. & Brown, P. O. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* **7**, e30733 (2012).
60. Cohen, D. R. & Townsend, C. A. A dual role for a polyketide synthase in dynemicin enediyne and anthraquinone biosynthesis. *Nat. Chem.* **10**, 231–236 (2018).
61. van Galen, P. et al. Reduced lymphoid lineage priming promotes human hematopoietic stem cell expansion. *Cell Stem Cell* **14**, 94–106 (2014).
62. Yang, L. et al. Metadherin/astrocyte elevated gene-1 positively regulates the stability and function of forkhead box M1 during tumorigenesis. *Neuro Oncol.* **19**, 352–363 (2017).
63. Ohnishi, T., Matsumura, H., Izumoto, S., Hiraga, S. & Hayakawa, T. A novel model of glioma cell invasion using organotypic brain slice culture. *Cancer Res.* **58**, 2935–2940 (1998).

Acknowledgements

This study was approved by the IRB of the First Affiliated Hospital of Sun Yat-sen University. We gratefully thank all patients who participated in these studies. We also sincerely thank J. Rich for providing the GSCs and P. Xiang for providing the iPSC-derived human NSCs. This work was supported in part by the National Natural Science Outstanding Youth Foundation of China (81822033 to N.Z.), the National Key Research and Development Program of China (2016YFA0503000 to N.Z.), the National Natural Science Foundation of China (81572477 and 81772683 to N.Z.), the Guangdong Innovative and Entrepreneurial Research Team Program (2016ZT06S638 to B.L.), and the Lou and Jean Malnati Brain Tumor Institute at Northwestern Medicine (S.-Y.C.). We thank D. Tiek for reading, editing and commenting on this manuscript. We also thank S. Xiao and T. Huang for helping with the graphical abstract.

Author contributions

Conceptualization: N.Z.. Data production, analysis and investigation: X.G., X.X., F.L., M.Z., K.Z., H.Z., X.W., J.Z., D.L., Q.X., F.X., N.Z. and S.-Y.C. Resources: Q.X., B.L., Z.Z. and T.J. Writing, reviewing and editing: N.Z., X.G. and S.-Y.C. Supervision: N.Z. Funding acquisition: N.Z. Comments and revision: S.-Y.C.

Competing interests

The authors declare no competing interests.

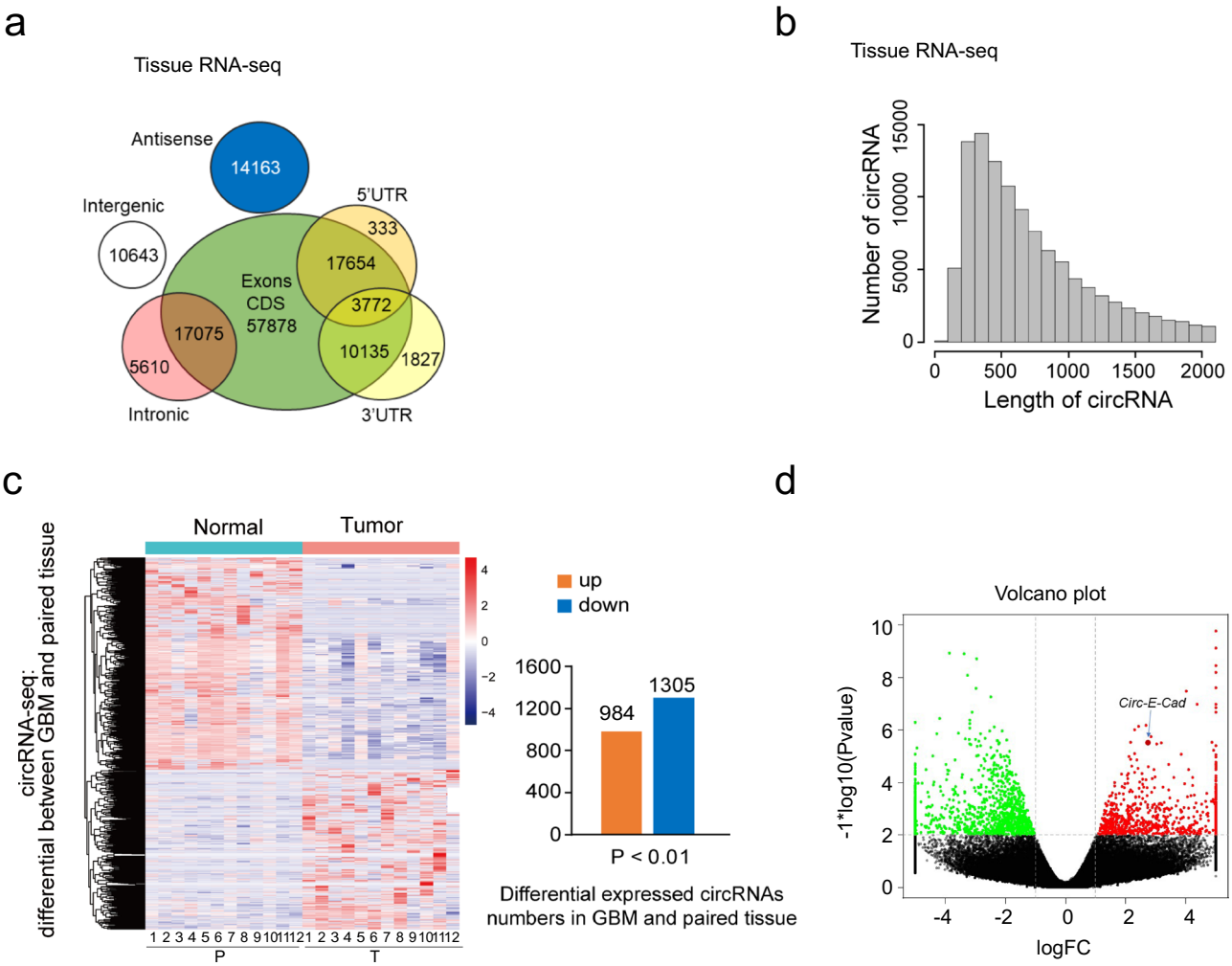
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41556-021-00639-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41556-021-00639-4.

Correspondence and requests for materials should be addressed to S.-Y.C. or N.Z.

Peer review information *Nature Cell Biology* thanks Justin Lathia, Igor Vivanco and Markus Landthaler for their contribution to the peer review of this work. Peer reviewer reports are available.

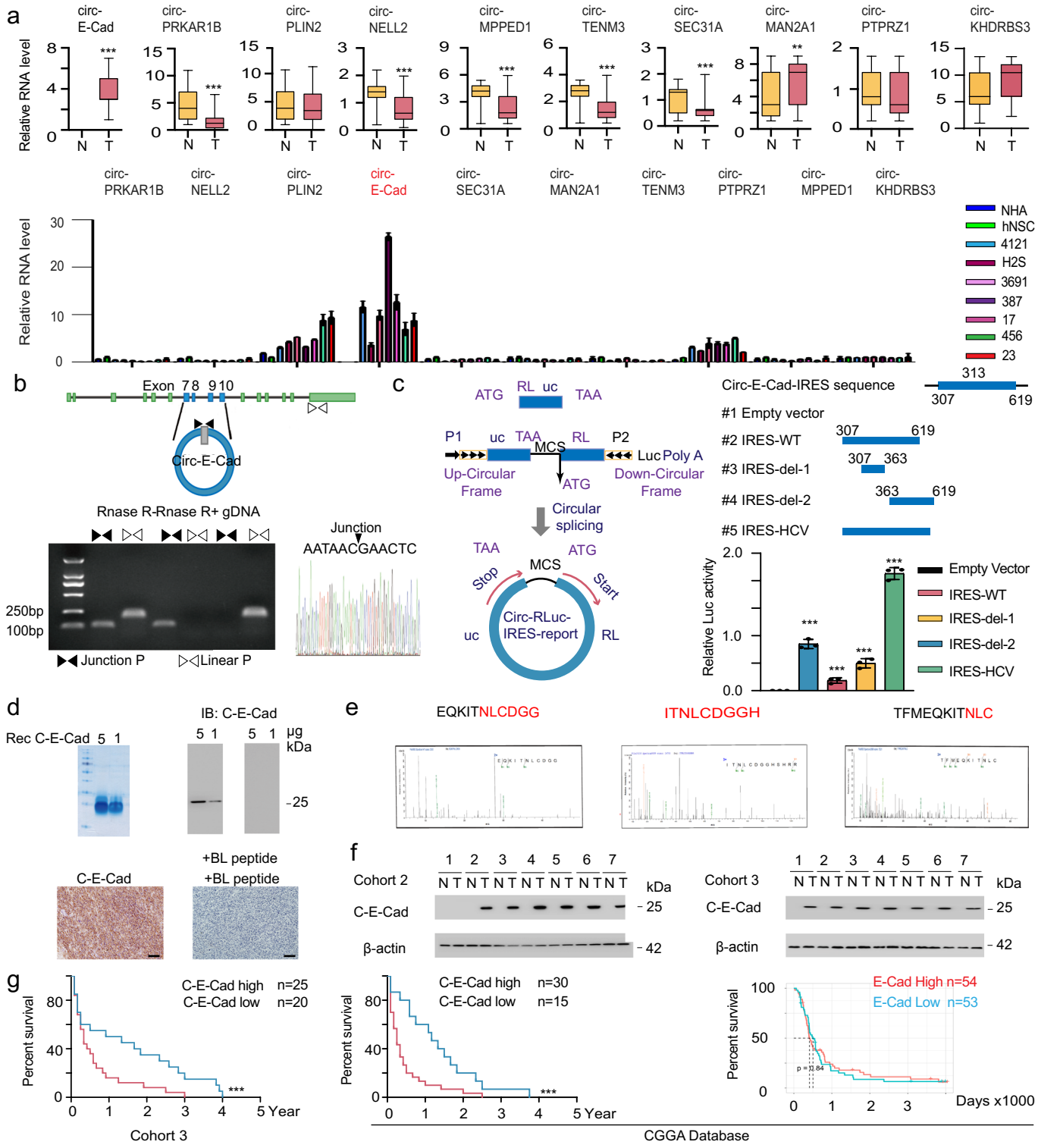


e

SourceGene	circBase_ID	Ribo Count	
		P	T
circ_PSEN1	hsa_circ_0003848	15	14
circ_RELL1	hsa_circ_0001400	2	12
circ_GUCY1A2	hsa_circ_0008602	1	0
circ_XPO	hsa_circ_0001016	10	2

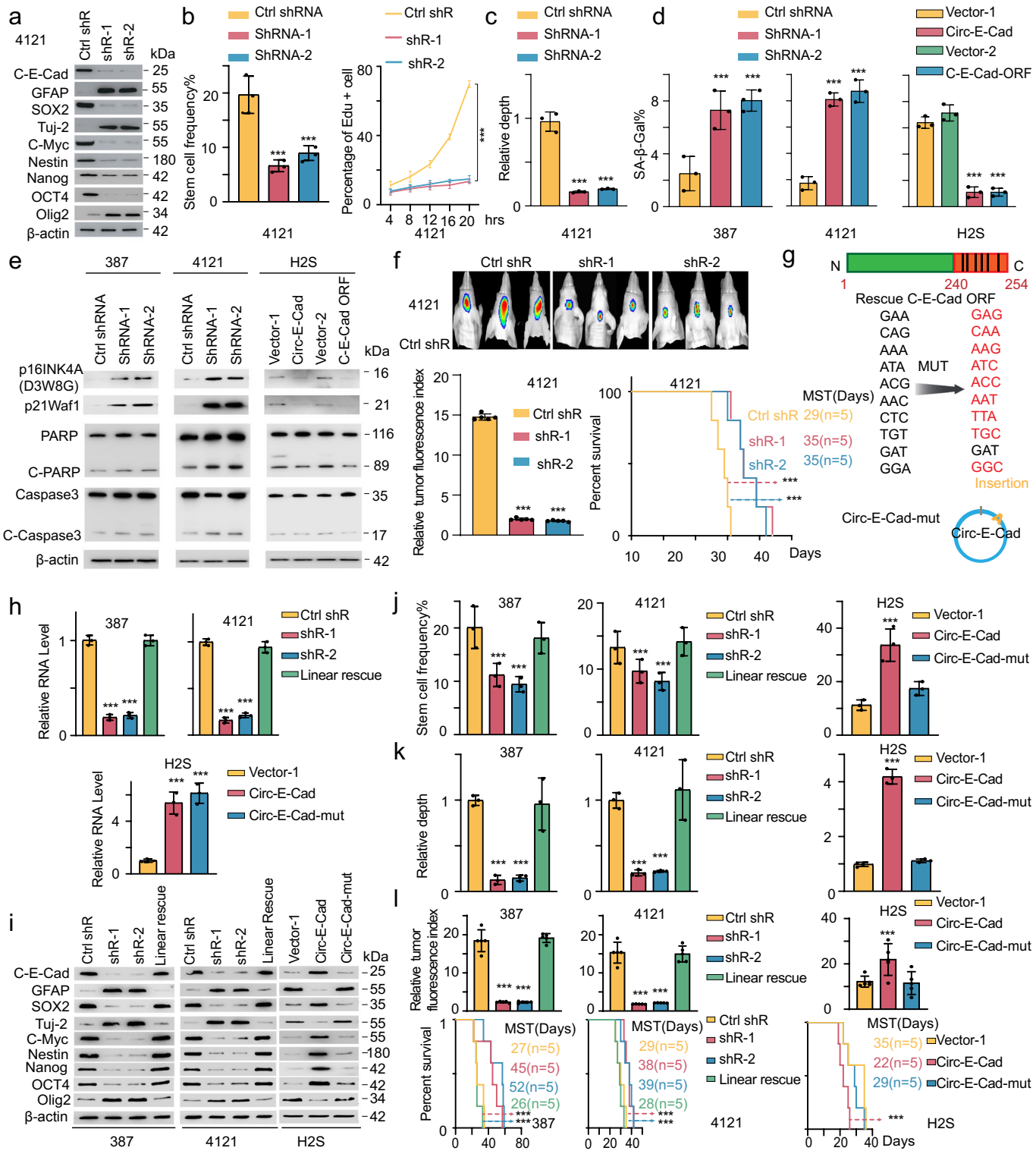
SourceGene	circBase_ID	Ribo Count	
		P	T
circ_E-Cad	hsa_circ_0039992	0	18
circ_PRKAR1B	hsa_circ_0079039	14	1
circ_PLIN2	hsa_circ_0086455	6	5
circ_NELL2	hsa_circ_0098585	137	22
circ_MPPED1	hsa_circ_0116713	48	22
circ_TENM3	hsa_circ_0125881	8	3
circ_SEC31A	hsa_circ_0127087	8	3
circ_MAN2A1	hsa_circ_0127540	3	7
circ_PTPRZ1	hsa_circ_0133169	13	14
circ_KHDRBS3	hsa_circ_0135839	8	11

Extended Data Fig. 1 | Profiling of circular RNAs in GBM and NB; in GSCs and normal cells. (Related to Fig. 1). a, Venn diagram of circRNAs derived from different genomic regions. Data was from GBM samples and paired NB. **b**, RNA-seq read of abundance distributions of identified circRNAs. Data originated from GBM samples and paired adjacent normal tissues. X-axis: back-spliced read numbers of circRNAs. Y-axis: abundance of circRNAs classified by different read numbers. The majority of circRNAs identified were confirmed by more than 10 reads. **c**, Differentially expressed circRNAs (DEcRs). Data originated from GBM samples and paired adjacent normal tissues. DEcRs with $p < 0.01$ and fold change > 2 were considered significant. **d**, Volcano diagram of DEcRs from GBM samples and paired adjacent normal tissues. Circ-E-cad RNA was highly expressed in GBM samples ($p < 0.01$). **e**, Left: 4 circRNAs with coding potential in both brain tissues and hearts. Right: 10 differentially expressed coding circRNAs with circBase annotation and their RPFs in GBM and normal brain tissues.



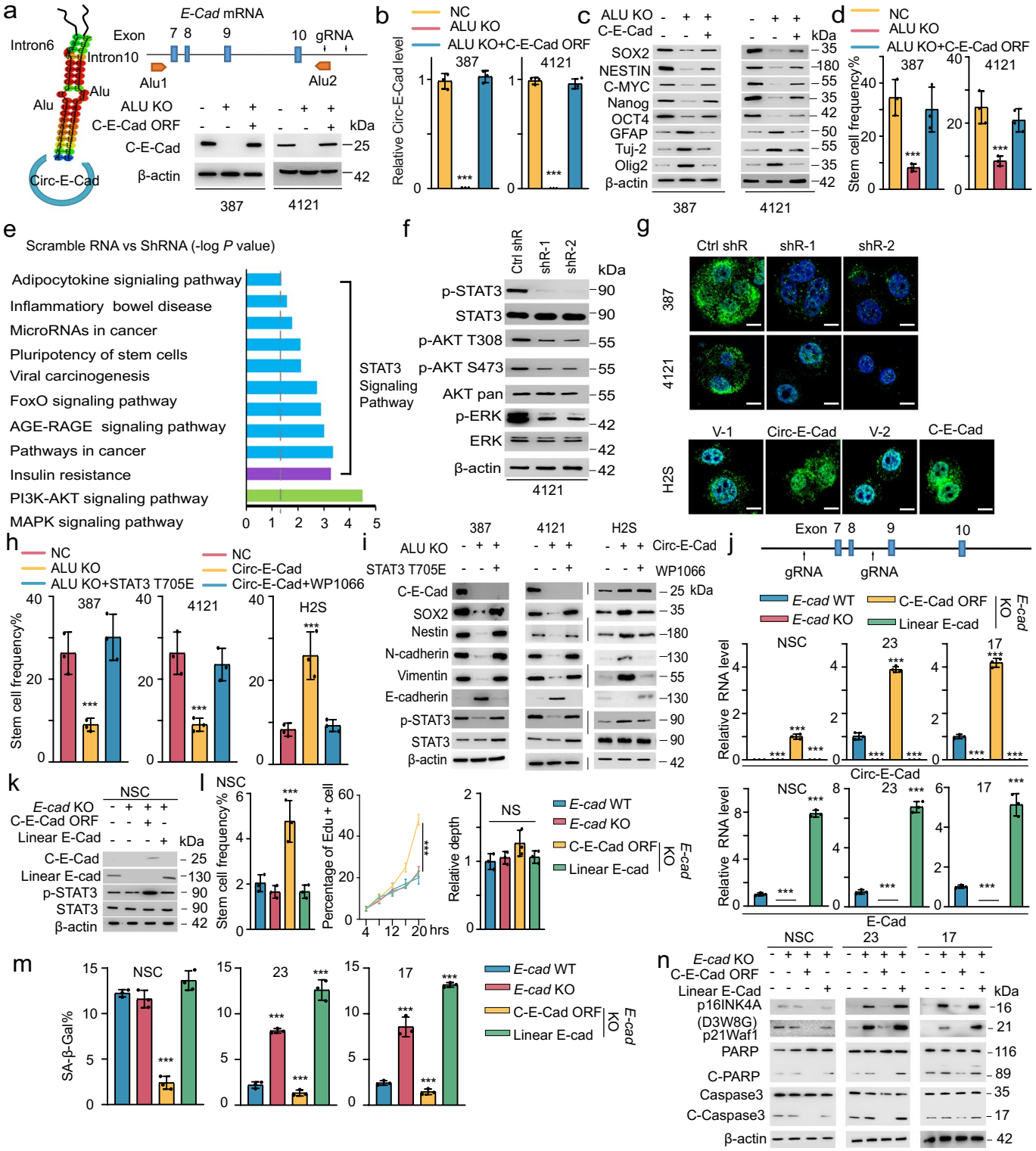
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Expression of circ-E-cad RNA in GBM; identification of IRES in circ-E-cad RNA; characterization of the anti-C-E-cad antibody (Related to Fig. 2). **a**, qPCR of the relative expression levels of 10 differential protein-coding circRNAs GBM and normal tissues, n = 44 biologically independent samples, data were presented as boxes containing the first and third quartiles. The whiskers indicate the maxima and minima. Wilcoxon test, **p = 0.005, ***p < 0.001. In GSCs and NSC/NHA. Circ-E-cad RNA had lowest expression in normal brain tissues among all candidate circRNAs, n = 3 independent experiments, data were presented as the mean ± SD, two-sided t test, the p value was detailed in Source data. **b**, Upper, exons 7-10 of E-cadherin formed circ-E-cad RNA. Lower left, RT-PCR of circular and linear E-cadherin RNA in GSC H2S with or without RNase R treatment.Lower right, Sanger sequencing. n = 3 independent experiments. **c**, Left, sketches of the strategy for circular RNA based IRES verification. Right, WT or different truncated IRES predicated in circ-E-cad RNA in circ-Rluc-IRES report vector as indicated. n = 3 independent experiments. data were presented as the mean ± SD, two-sided t test, ***p < 0.001. **d**, Upper, left, a Coomassie blue-stained gel; right, anti-circ-E-Cad antibody validation by IB. BL peptide, blocking peptide with the 14a.a. peptide sequence. Lower, IHC. anti-circ-E-Cad antibody validation in a clinical GBM tumor sample. n = 3 independent experiments. scale bar, 250 μm **e**, MS/Mass-spectra identified C-E-cad unique 14a.a. peptide sequences in GSCH2S and 387. **f**, IB of C-E-Cad in 14 randomly selected GBM samples (cohort 2 and cohort 3) and their paired NB. n = 3 independent experiments. **g**, Two-sided, Log-rank analysis of GBM patients (n = 45 biologically independent samples) from cohort 3 correlated with C-E-Cad levels (left); CGGA database with C-E-Cad levels (middle) and E-Cad levels (right).***p < 0.001.



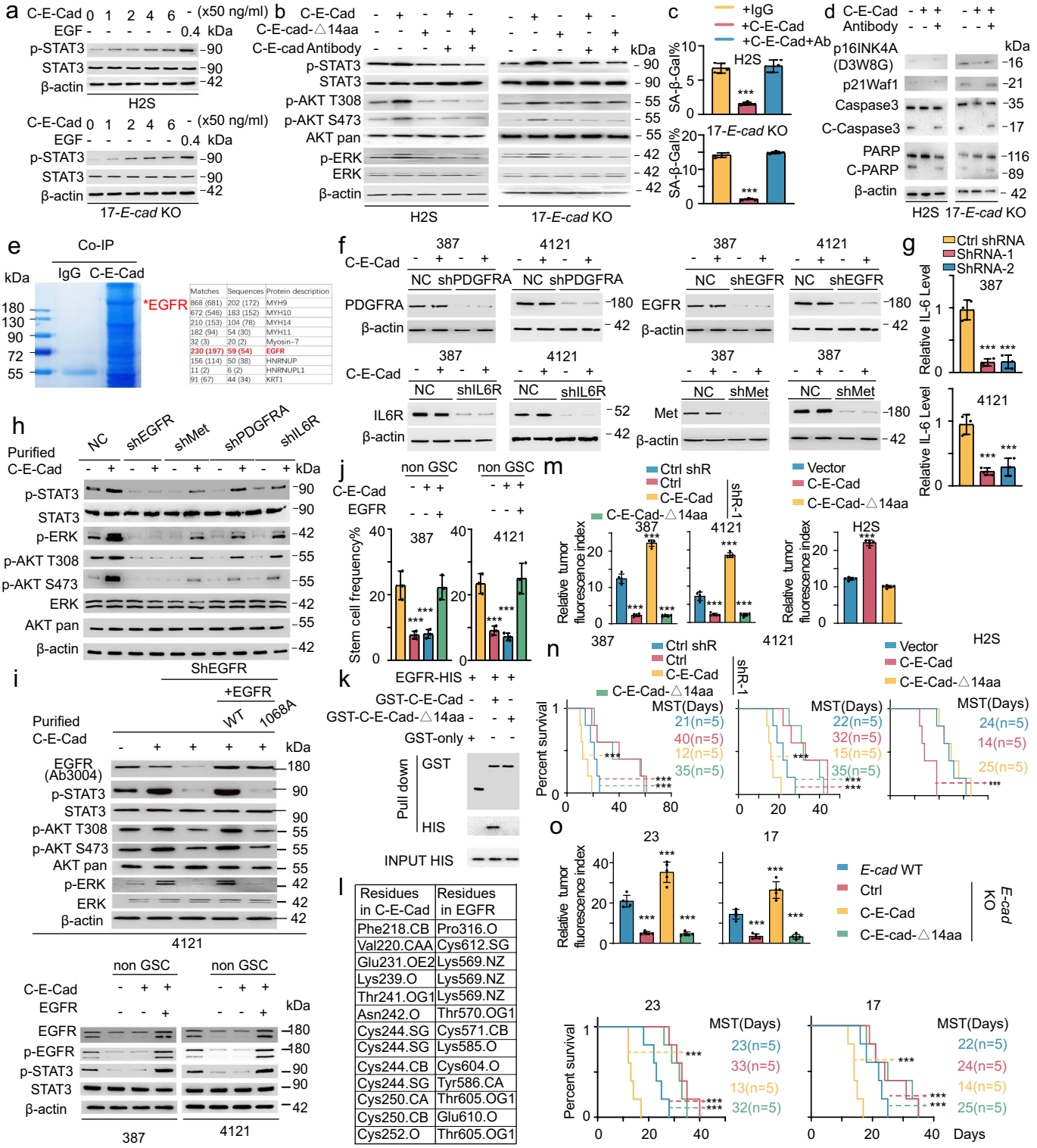
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | C-E-cad, but not circ-E-cad RNA, promotes GSC self-renewal and survival. (Related to Fig. 3). a, IB of indicated proteins in GSC4121 with indicated modifications. **b**, LDA assay and percentage of EdU-positive of GSC4121. **c**, The relative invasion depth of GSC4121. **d**, The percentage of SA-β-Gal positive cells of GSC387, 4121, H2S with indicated modifications. **e**, IB of indicated proteins in GSC387, 4121, and H2S with indicated modifications. **f**, Upper, BLI images of in vivo tumorigenicity using GSC4121 with indicated modifications. Lower left, tumor volumes. Lower right, Two-sided, Log-rank analysis of mice intracranially implanted with GSC4121 with indicated modifications. **g**, Left, a shRNA-resistant linearized C-E-cad vector and a mutated circ-E-cad RNA vector. Right, an adenine was inserted after the ATG start codon. **h**, qPCR of relative circ-E-cad RNA level in GSC387, 4121 and H2S with indicated modifications. **i**, IB of indicated proteins in GSC387 and 4121 with stable circ-E-cad RNA KD and re-expression of a linearized C-E-cad (resistant to junction shRNA), and in GSCH2S overexpressing a linearized C-E-Cad or a mutated circ-E-cad RNA (insertion A). **j**, LDA assay of GSC387 and 4121 with indicated modifications. **k**, Brain slice invasion of GSC387, 4121, and H2S with indicated modifications. **l**, Upper, tumor volumes of GSC387, 4121 and H2S with indicated modifications. Lower, Two-sided, Log-rank analysis of mice injected with GSC387, 4121 and H2S with indicated modifications, ***$p < 0.001$. In **a-e,h-k**, $n = 3$ independent experiments. In **f,l**, $n = 5$ biologically independent samples. In **b-d**,**f**,**h-l**, data was presented as mean ± SD. Two-sided t test, ***$p < 0.001$.



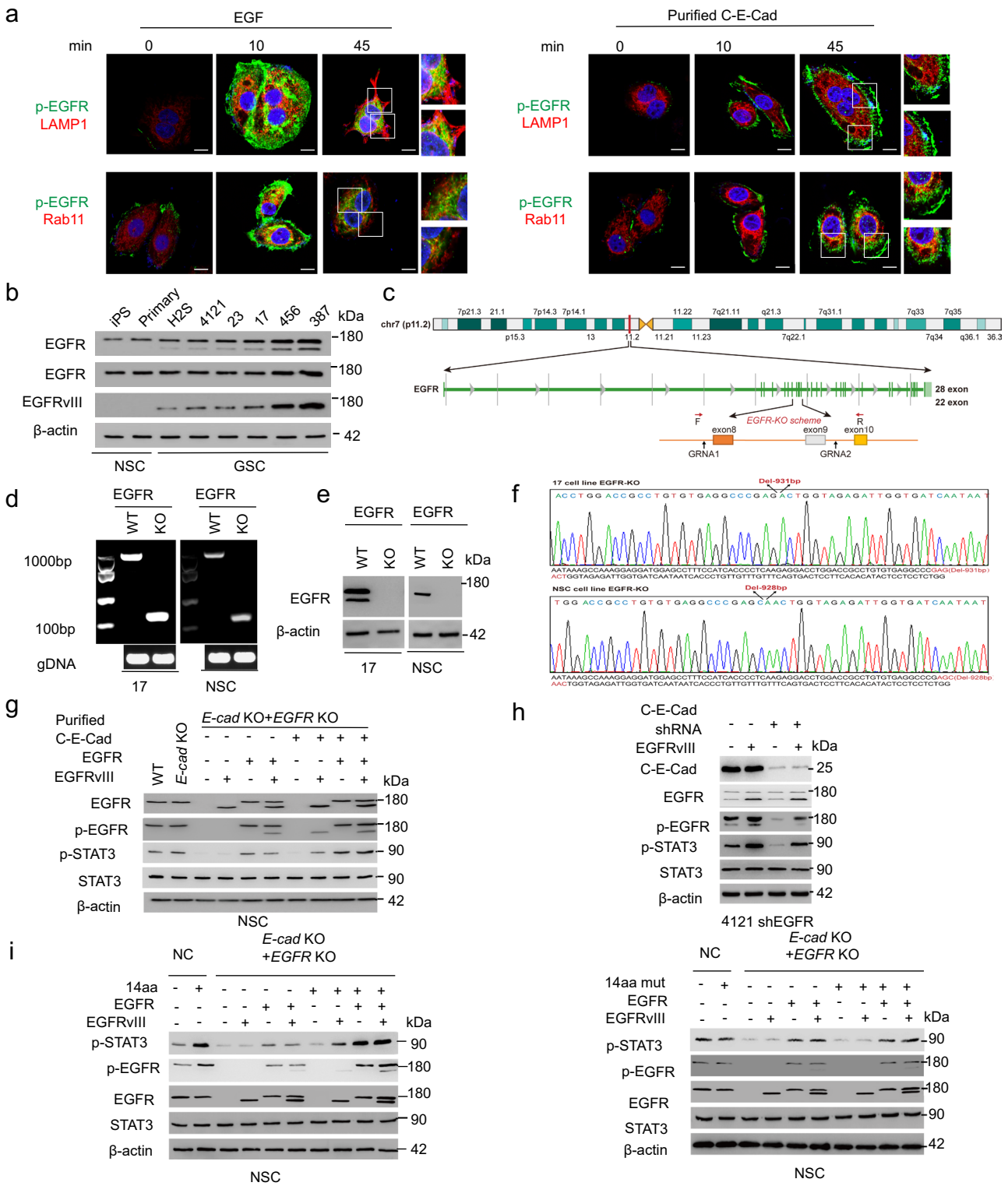
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Flanking ALU sequences are required to form circ-E-cad RNA; C-E-Cad activates STAT3, AKT and ERK signalling and independent functions of C-E-cad and E-cadherin. (Related to Figs. 3 and 4). a, The illustration of side flanking ALU sequences of circ-E-cad RNA and the CRISPAR/Cas9 strategy for KO downstream ALU sequences **b**, qPCR of circ-E-cad RNA in ALU KO GSC387 and 4121 with indicated modifications. **c**, IB in ALU KO GSC387 and 4121 with indicated modifications and controls. **d**, LDA assay in GSC387 and 4121 with indicated modifications.e. KEGG pathway enrichment analysis of circ-E-cad RNA stable KD GSC4121. **f**, IB of GSC4121 with indicated modifications. **g**, IF. p-STAT3 localization was determined in C-E-cad stable KD GSC387 4121 and in a circ-E-cad RNA or a linearized C-E-cad-overexpressing GSCH2S. scale bar, 20μm. **h**, LDA assay in GSC387, 4121 and GSCH2S with indicated modifications. **i**, IB in in GSC387, 4121 and GSCH2S with indicated modifications. **j**, Upper, Strategy of E-cadherin KO by CRISPAR/Cas9 system in GSC23, 17, and NSC. Lower, qPCR of Circ-E-Cad and E-Cadherin in GSC23, 17, and NSC with indicated modifications. **k**, IB in NSC-WT, NSC-E-Cadherin KO cells. Circ-E-Cad RNA or E-cadherin was re-expressed in the indicated cells. **l**, LDA assay, Edu assay and brain slice invasion assay in NSC cell lines with indicated modifications. **m**, The percentage of SA-β-Gal positive cells of GSC387, 4121, and H2S with indicated modifications. **n**, IB of senescence markers p16, p21 and apoptosis related Caspase3 and PARP of GSC387, 4121, and H2S with indicated modifications.In **a-d, f-n**, n = 3 independent experiments. In **b,d,h,j,l,m**, data were shown as mean ± SD, two-sided t test,***$p < 0.001$.

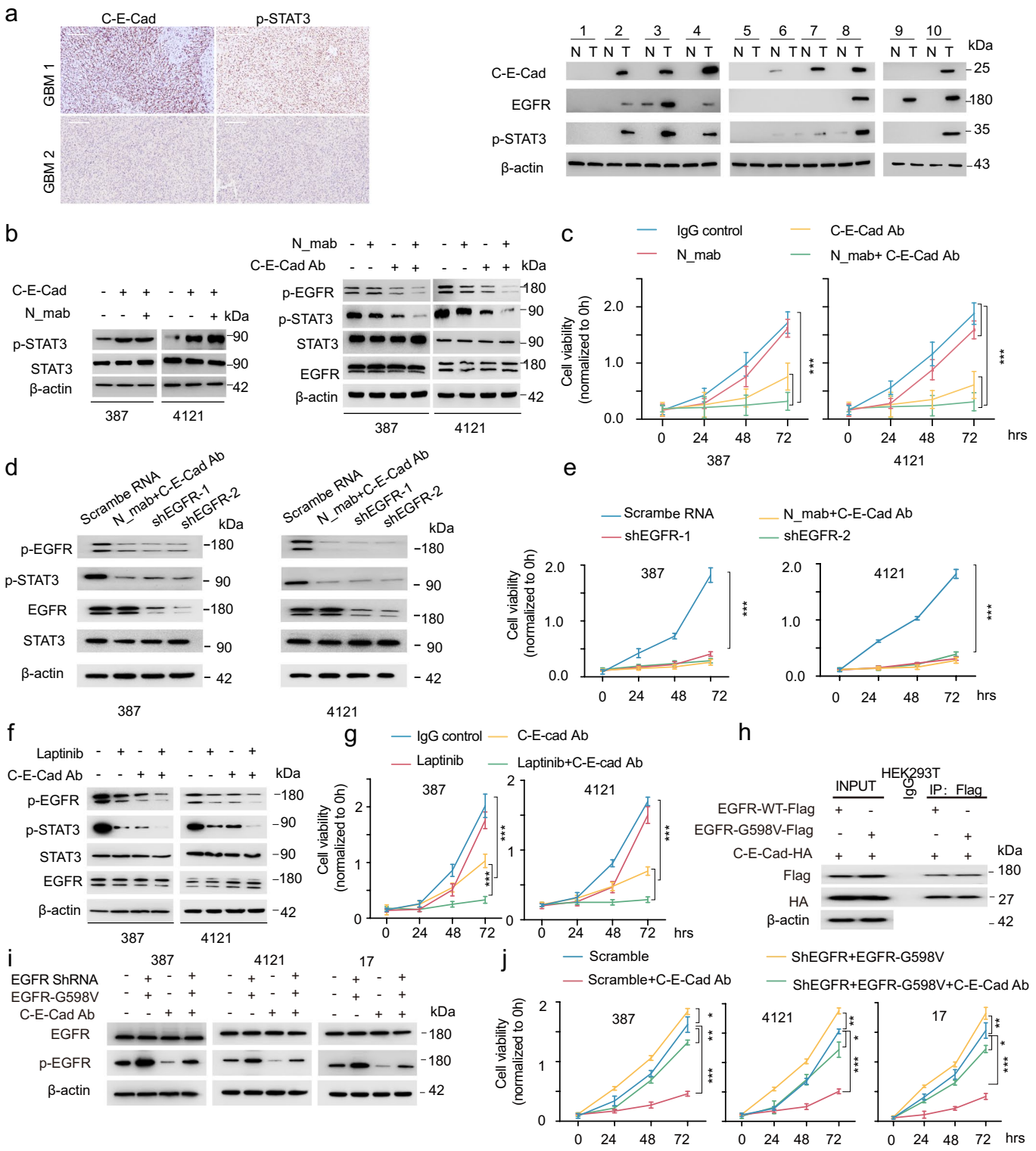


Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Recombinant C-E-cad activates EGFR. (Related to Fig. 5). a, IB. rC-E-cad with different concentrations were used to treat indicated cells. EGF was used as a positive control. **b**, IB for GSCH2S and GSC17 with indicated modifications. **c**, The percentage of SA-β-Gal positive cells of GSCH2S and GSC17 E-cadherin KO with indicated modification. **d**, IB of GSCH2S and GSC17 E-cadherin KO with indicated modification. **e**, Mass spectrometry analysis identified EGFR peptide sequences that was pulled down by the anti-C-E-cad antibody. (IB for GSC387 and 4121 with indicated modifications were separately treated with or without rC-E-cad proteins.). **g**, ELISA of IL-6 level in GSC387 and 4121 with indicated modifications. **h**, IB. Left, GSC4121 with stable KD of EGFR, Met, PDGFRA, or IL6R and treated with or without rC-E-cad (200 ng/ml). Right. GSC4121 with stable EGFR KD were re-expressed with WT EGFR or EGFR with a Y1068A mutation, then treated with rC-E-cad. **i**, IB. non-GSC387 and 4121 were modified or treated as indicated. **j**, LDA assay for non-GSC387 and 4121 with indicated modification or treatments. **k**, IB. EGFR-HIS, GST-C-E-Cad and GST-C-E-Cad-△14aa was purified and GST pull down assay was applied. **l**, Docking analysis of C-E-cad and EGFR CR2 domain. **m**, Tumor volume in GSC387, 4121 and H2S with indicated modifications. **n**, Two-sided, Log-rank analysis of mice intracranially implanted with GSC387, 4121 and H2S with indicated modifications, ***$p < 0.001$o. Tumor volumes and Two-sided, Log-rank analysis of mice bearing indicated GSC tumor xenografts with 5 mice per group. ***$p < 0.001$In **a-k**, $n = 3$ independent experiments. In **m-o**, $n = 5$ independent experiments. In c,g,j,m.o, data were shown as mean ± SD, two-sided t test,***$p < 0.001$.

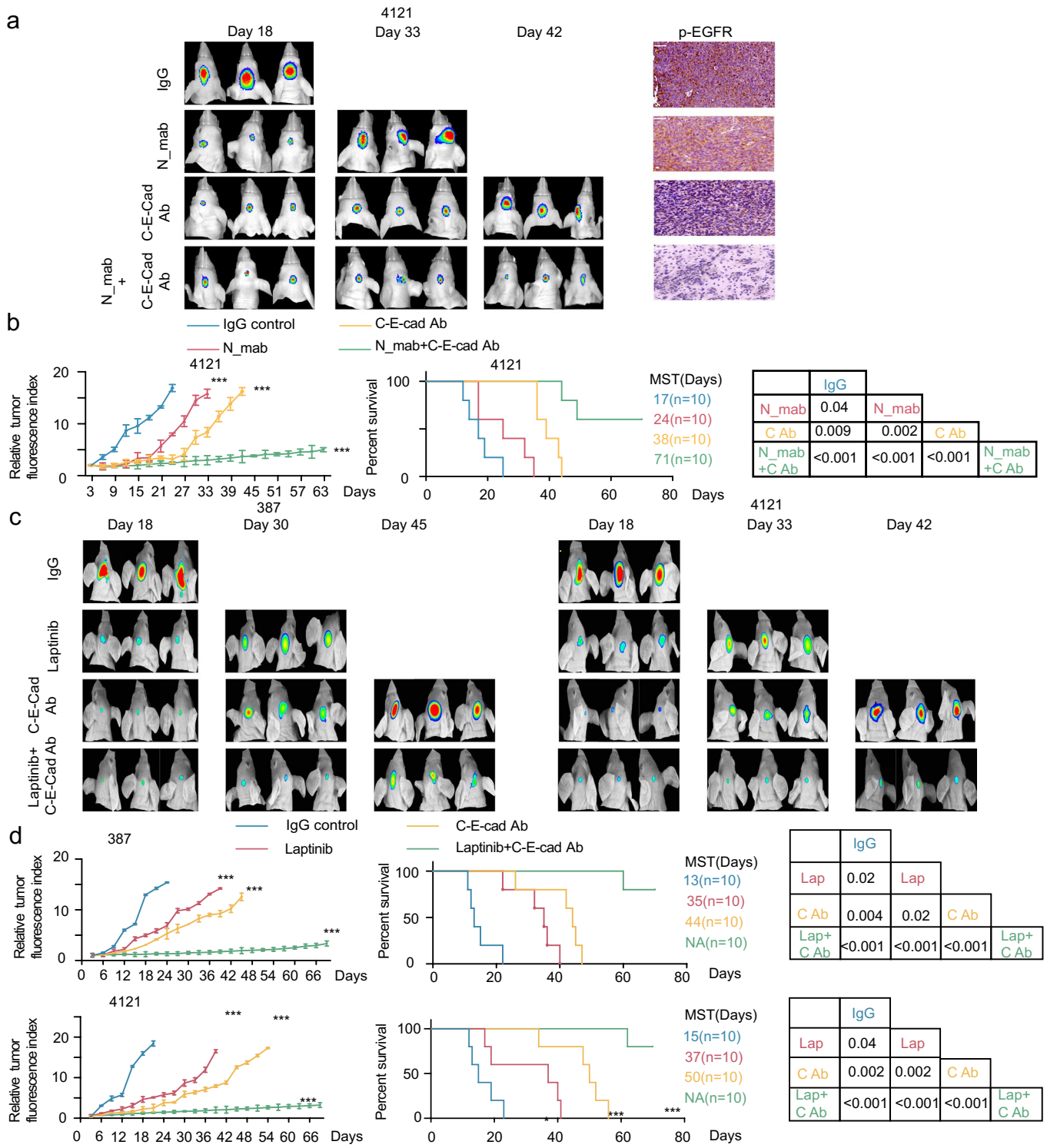


Extended Data Fig. 6 | C-E-cad interacts and activates EGFR. (Related to Fig. 7). a, IF. Colocalizations of p-Y1068-EGFR and LAMP1 or Rab11 were detected in GSC 17-E-cad KO treated with EGF or recombinant C-E-cad at indicated timepoints. Scale bar, 20 μm. **b**, IB of EGFR and EGFRvIII expression in indicated cell lines. **c**, CRISPAR/Cas9-induced EGFR KO strategy. **d**, DNA gel electrophoresis of RT-PCR for EGFR in GSC17-WT, 17-KO, NSC-WT, and NSC-KO cells. **e**, IB of EGFR in GSC17-WT, 17-KO, NSC-WT, and NSC-KO cells. **f**, Sanger sequencing of CRISPAR/Cas9-induced EGFR KO in indicated cells. **g**, IB of p-STAT3, p-EGFR, STAT3, and EGFR in WT and E-cadherin KO together with EGFR KO and EGFR WT or EGFRvIII re-expressed NSCs. Recombinant C-E-cad treatment (200 ng/ml) is indicated. **h**, IB. p-STAT3, p-EGFR, STAT3, and EGFR in GSC 4121 with stable EGFR KD and with stable C-E-cad KD and EGFRvIII re-expression. **i**, IB of p-STAT3, p-EGFR, STAT3, and EGFR in WT and E-cadherin KO together with EGFR KO and EGFR WT or EGFRvIII re-expressed NSCs treated with a synthetic C-E-cad C-terminal 14 a.a. peptide (200 ng/ml) (left) or C-E-cad C-terminal mutant 14 a.a. peptide (200 ng/ml) (right). In **a-b,d,e,g-i**, n = 3 independent experiments.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Clinical implication of the anti-C-E-cad antibody in vitro. (Related to Fig. 8). a, IHC and IB. Representative C-E-cad and p-STAT3 levels were determined in GBM samples by IHC (left) and IB (right). Scale bar, 250 μm. **b**, Left, IB of p-STAT3 in GSC387 and 4121 cells treated with purified C-E-Cad and Nimotuzumab (N_mab). Right, IB of p-STAT3 and p-EGFR in GSC387 4121 cells treated with Nimotuzumab (N_mab) and C-E-Cad antibody. **c**, Cell viability was detected in GSC387 and 4121 cells treated with Nimotuzumab (N_mab) and C-E-Cad antibody, ***$p < 0.001$. **d**, GSC387 and 4121 cells treated with Nimotuzumab + C-E-Cad antibody or transfected with shRNAs targeting EGFR. The levels of p-STAT3 and p-EGFR were detected. **e**, Cell viability was detected in GSC387 and 4121 cells with indicated modifications, ***$p < 0.001$. **f**, IB of p-STAT3 and p-EGFR in GSC387 and 4121 cells treated with Laptinib and C-E-Cad antibody. **g**, Cell viability was detected in GSC387 4121 cells treated with Laptinib and C-E-Cad antibody, ***$p < 0.001$. **h**, IP-IB of indicated proteins. HEK293T cells were co-transfected with EGFR-WT/EGFR-G958V and C-E-Cad-HA, Cells were then harvested and subjected to IP assay. **i**, GSC387, 4121, and 17 were transfected with shRNAs targeting EGFR and re-expressed with EGFR-G958V. Cells were treated with or without C-E-Cad antibody. IB of p-EGFR was detected. **j**, Cell viability in GSC387,4121 and 17 cells described in (i). In 387,*$p = 0.026$,**$p = 0.006$,***$p < 0.001$,in 4121, *$p = 0.01$,**$p = 0.001$,***$p < 0.001$,in 17, *$p = 0.022$,**$p = 0.009$,***$p < 0.001$.In **a-j**. n = 3 independent experiments. In **c,e,g,j**, data were shown as mean ± SD, two-sided t test.



Extended Data Fig. 8 | Clinical implication of the anti-C-E-cad antibody in vivo. (Related to Fig. 8). a, In vivo tumorigenicity assay using GSC4121 treatment with Nimotuzumab (N_mab), anti-C-E-cad antibodies or in combination, (n = 10 animals). BLI images (Left) indicated brain BITC tumor xenografts. Right, IHC of p-EGFR expression in indicated GSC tumours, scale bar, 250μm (n = 3 animals). **b**, Left, Relative intensity of fluorescent index of BLI that indicate tumor growth of GSC4121 brain tumor xenografts in animals treated with Nimotuzumab (N_mab), anti-C-E-cad antibodies or in combination. Data were presented as mean ± SD, two-sided t test, ***$p < 0.001$. Middle, Two-sided, Log-rank analysis of mice treated with indicated therapeutic strategies. Right, P values were calculated, n = 10 animals. **c**, In vivo tumorigenicity assay using GSC387 and 4121 and treatment with Laptinib, anti-C-E-cad antibodies or in combination. BLI images at different time point, (n = 10 animals). **d**, Left: Relative intensity of fluorescent index of BLI that indicate tumor growth of GSC387 and 4121 brain tumor xenografts in animals treated with Laptinib (1 μg/μl, 3 μl), anti-C-E-cad antibodies or in combination, n = 10 animals. Data were presented as mean ± SD, two-sided t test, ***$p < 0.001$.Middle, Two-sided, Log-rank analysis of mice treated with indicated therapeutic strategies. Right, P values were calculated, (n = 10 animals).

nature research

Corresponding author(s): Nu Zhang

Last updated by author(s): Jan 18, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	software RSEM (version 1.2.19),PhotoSound PAFT/256, TopHat2 (version 2.1.1),Olympus image FV100,ClusPro server4-8,Microsoft Excel,GraphPad Prism
Data analysis	R language (version 3.2.19), (https://www.r-project.org/)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA–seq data that support the findings of this study have been deposited in the NCBI PRJNA525736, SRA714646, PRJNA525736.
The human GBM data were derived from the TCGA Research Network: http://cancergenome.nih.gov/. The data-set derived from this resource that supports the findings of this study is available in http://gepia.cancer-pku.cn/detail.php?gene. And the CGGA data were derived from 301 data-set available in http://gepia.cancer-pku.cn/detail.php?gene. All other data supporting the findings of this study are available from the corresponding author on reason.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size were not predetermined by any statical methods. To ensure sufficient statical power, sample size,n= 5 or 10 were choosen.
Data exclusions	No data was excluded.
Replication	all experiments were applied with at least 3 biological replications with similiar results.
Randomization	For in vivo study, mice were randomly divided into each group without any self-selections.
Blinding	The investigators was blinded on the groups allocations in all the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

The antibodies were listed in the Methods section and Table5
Anti-C-E-Cad (mouse), 1:500 This study N/A
Anti-E-Cadherin, 1:1000 Cell Signaling Technology Cat#14472
Anti-β-actin, 1:5000 Cell Signaling Technology Cat#4970
Anti-Sox2, 1:1000 Cell Signaling Technology Cat#23064
Anti-GFAP, 1:1000 Cell Signaling Technology Cat#12389
Anti-Tuj-1, 1:1000 Cell Signaling Technology Cat#5568
Anti-C-Myc, 1:1000 Cell Signaling Technology Cat#5605
Anti-Nestin, 1:1000 Cell Signaling Technology Cat#33475
Anti-STAT3, 1:1000 Cell Signaling Technology Cat#9139
Anti-p-STAT3, 1:1000 Cell Signaling Technology Cat#9145
CD133, 1:1000 Abcam Ab19898
Anti-EGFR, 1:1000 Abcam Ab3004
Anti-EGFR, 1:1000 Abcam Ab32077
Anti-EGFRvIII, 1:1000 Nouvus DH8.3
Anti-p-EGFR (Y1068), 1:1000 Abcam Ab40815
Anti-C-E-cad （blocking peptide）, 1:500 This study N/A
Anti-p-AKT308, 1:1000 Cell Signaling Technology Cat#4056
Anti-p-AKT473, 1:1000 Cell Signaling Technology Cat#4060
Anti-AKT pan, 1:1000 Cell Signaling Technology Cat#2920
Anti-p-ERK, 1:1000 Cell Signaling Technology Cat#4370
Anti-ERK, 1:1000 Cell Signaling Technology Cat#4695
Anti-IL-6R, 1:1000 Abcam ab128008
Anti-Met, 1:1000 Abcam ab51067
Anti-PDGFRA, 1:1000 Abcam ab203491

Alexa Fluor 488, 1:2000 Thermo Fisher Scientific Cat# A-11029
Alexa Fluor 594, 1:2000 Thermo Fisher Scientific Cat# A-21207
Anti-mouse IgG, 1:1000 Vector Laboratories Cat# AI-2000
Anti-rabbit IgG, 1:1000 Vector Laboratories Cat#AI-1000
Anti-mouse IgG, 1:1000 Abcam Ab6785
Anti-rabbit IgG, 1:1000 Abcam Ab97051

Validation

The C-E-Cad antibody was validated based on the RNA level, blocking peptide and KO cell lines as negative control. other antibodies were validated according to the manufacturer validation.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

456 Dr. Professor Jeremy Rich N/A
4121 Dr. Professor Jeremy Rich N/A
387 Dr. Professor Jeremy Rich N/A
H2S Dr. Professor Jeremy Rich N/A
17 Dr. Professor Nu Zhang N/A
23 Dr. Professor Nu Zhang N/A
NSC Thermo Fisher Scientific Cat# A15654
iPS-NSC Dr. Professor Peng Xiang N/A

HEK293T ATCC Cat# CRL-3216

Authentication

cells were authenticated by STR method.

Mycoplasma contamination

all cells were tested negatively for mycoplasma infection.

Commonly misidentified lines
(See ICLAC register)

Cells in ICLAC were not used in this research.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Athymic (Ncr nu/nu) female mice at 6-8 weeks of age were purchased from Nanjing University Farms. The details were shown in Methods section.

Wild animals

no wild animals were used in the study

Field-collected samples

no field collected samples were used in the study

Ethics oversight

this study was approved by the Institutional Animal Care and Use Committee (IACUC) of Sun Yat-sen University

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

No population characteristics were covariate relevant to the conclusion of this study and no analysis of the characteristics were applied.All the patients were diagnosed as GBM and receive surgical operation. The characteristics of CGGA and TCGA database were available in the database http://cancergenome.nih.gov/. http://gepia.cancer-pku.cn/detail.php?gene accordingly, the characteristics of the total 197 patients of cohort1-3 were detailed as below:
ID Gender Age
1 Female 51
2 Female 42
3 Male 33
4 Male 43
5 Female 25
6 Female 38
7 Male 45
8 Female 40
9 Male 43
10 Male 31
11 Female 26
12 Male 42
13 Male 34
14 Male 37
15 Male 39
16 Male 34
17 Female 36

18 Female 45
19 Male 45
20 Male 39
21 Male 20
22 Female 61
23 Male 41
24 Male 35
25 Male 36
26 Male 40
27 Male 55
28 Male 49
29 Male 28
30 Male 47
31 Male 57
32 Male 35
33 Female 30
34 Male 56
35 Female 73
36 Male 76
37 Female 41
38 Female 40
39 Male 37
40 Male 29
41 Male 40
42 Female 69
43 Male 58
44 Female 50
45 Male 55
46 Male 45
47 Male 21
48 Female 24
49 Male 69
50 Male 47
51 Female 57
52 Female 63
53 Female 43
54 Male 61
55 Male 59
56 Male 36
57 Male 56
58 Female 40
59 Female 41
60 Male 35
61 Female 60
62 Female 57
63 Female 65
64 Male 27
65 Male 65
66 Female 46
67 Female 70
68 Male 23
69 Female 60
70 Female 58
71 Male 29
72 Male 40
73 Male 60
74 Male 57
75 Male 47
76 Male 41
77 Female 32
78 Male 40
79 Male 40
80 Male 43
81 Male 56
82 Male 41
83 Female 19
84 Male 44
85 Male 46
86 Female 64
87 Male 55
88 Female 63
89 Male 63
90 Female 30
91 Female 40
92 Female 39
93 Female 31

94 Male 48
95 Male 67
96 Male 30
97 Male 34
98 Female 36
99 Male 26
100 Female 30
101 Female 56
102 Female 44
103 Male 65
104 Male 32
105 Male 32
106 Male 25
107 Male 58
108 Male 24
109 Female 43
110 Male 39
111 Male 43
112 Male 41
113 Female 45
114 Male 42
115 Male 42
116 Male 45
117 Female 66
118 Male 44
119 Male 52
120 Male 34
121 Female 46
122 Male 45
123 Male 50
124 Male 30
125 Female 55
126 Male 27
127 Male 59
128 Female 72
129 Female 45
130 Male 27
131 Female 51
132 Male 59
133 Male 33
134 Male 31
135 Male 31
136 Male 45
137 Female 32
138 Male 43
139 Female 52
140 Female 53
141 Male 47
142 Female 21
143 Female 58
144 Male 58
145 Female 41
146 Male 37
147 Female 46
148 Male 66
149 Male 36
150 Female 28
151 Male 33
152 Male 42
153 Male 42
154 Female 61
155 Female 37
156 Male 35
157 Male 31
158 Female 47
159 Female 56
160 Male 29
161 Male 26
162 Female 41
163 Female 47
164 Male 41
165 Male 45
166 Female 47
167 Male 28
168 Female 42
169 Male 41

170 Female 44
171 Female 69
172 Female 34
173 Male 27
174 Male 34
175 Male 40
176 Male 42
177 Male 23
178 Female 44
179 Male 50
180 Male 40
181 Female 60
182 Female 15
183 Female 19
184 Male 62
185 Male 40
186 Female 66
187 Male 52
188 Female 46
189 Female 40
190 Female 28
191 Female 40
192 Female 34
193 Male 32
194 Male 42
195 Female 52
196 Female 44
197 Male 52

Recruitment	Patients were recruited randomly without any potential self selected.
Ethics oversight	informed and signed content were obtained form each patients approved by the ethical institution's Institutional Review Board of The First Affiliated Hospital of Sun Yat-sen University, Sun Yat-sen University Cancer Center, First Affiliated Hospital of Nanjing Medical University and Beijing Neurosurgical Institute, Capital University.following the protocol approved by the ethical institution's Institutional Review Board of The First Affiliated Hospital of Sun Yat-sen University, Sun Yat-sen University Cancer Center, First Affiliated Hospital of Nanjing Medical University and Beijing Neurosurgical Institute, Capital University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.