Single-cell analyses reveal YAP/TAZ as regulators of stemness and cell plasticity in glioblastoma

Martina Castellan1, Alberto Guarnieri1, Atsushi Fujimura1, Francesca Zanconato1, Giusy Battilana1, Tito Panciera1, Hanna Lucie Sladitschek1, Paolo Contessotto1, Anna Citron1, Andrea Grilli2, Oriana Romano2, Silvio Bicciato2, Matteo Fassan3, Elena Porcù4, Antonio Rosato5,6, Michelangelo Cordenonsi1,8 and Stefano Piccolo1,8

Glioblastoma (GBM) is a devastating human malignancy. GBM stem-like cells (GSCs) drive tumor initiation and progression. Yet the molecular determinants defining GSCs in their native state in patients remain poorly understood. Here we used single-cell datasets and identified GSCs at the apex of the differentiation hierarchy of GBM. By reconstructing the GSCs’ regulatory network, we identified the YAP/TAZ coactivators as master regulators of this cell state, irrespectively of GBM subtypes. YAP/TAZ are required to install GSC properties in primary cells downstream of multiple oncogenic lesions and are required for tumor initiation and maintenance in vivo in different mouse and human GBM models. YAP/TAZ act as main roadblock of GSC differentiation, and their inhibition irreversibly locks differentiated GBM cells into a nontumorigenic state, preventing plasticity and regeneration of GSC-like cells. Thus, GSC identity is linked to a key molecular hub integrating genetics and microenvironmental inputs within the multifaceted biology of GBM.

GBM is the most frequent and lethal form of brain cancer. GBM is characterized by a high degree of intratumoral cellular heterogeneity and plasticity, contributing to therapeutic resistance and recurrence. At the cellular level, many of the malignant traits of GBM have been interpreted through the biology of a stem-like cell population (GSCs), endowed with the ability to self-renew, to initiate tumors in vivo and to give rise to a hierarchy of more-differentiated progeny. These attributes establish operational definitions and corresponding bioassays for GSCs, retrospectively identifying GSCs in primary GBM. However, although these operational definitions are effective at revealing bona fide GSCs and allowing their experimental manipulation, they also necessarily introduce experimental biases that are hardly compatible with the need of molecularly characterizing GSCs as they exist in vivo. For example, GSCs are typically sorted from other GBM cell populations through surface markers of disputed and incomplete specificity; moreover, GSCs need to be expanded ex vivo under the confounding influence of artificial culturing conditions, clearly different from those existing in vivo (for example in term of growth factors availability, oxygen or mechanical gradients and more). Then, GSCs are typically cultured ex vivo as neurospheres, where they coexist with their more-differentiated progeny. Thus, the molecular portrait of GSCs that one can obtain through these procedures may differ substantially from one of tumor-resident, native GSCs. In turn, these confounding caveats have so far ultimately limited our understanding of the molecular underpinning of the native GSC state, its key determinants and associated vulnerabilities. Addressing this gap is essential for the design of more effective therapies.

Single-cell analyses have the potential to overcome the above limitations, allowing for a less-biased identification of GSC-like cells in their native environments. Identifying the molecular features of native GSCs in single-cell data is the subject of intense investigation, with some studies focusing on plasticity, some on hierarchies and some on subtype-specific stem cells. Here we started from single-cell data to identify gene regulatory networks of native GSCs, leading to the discovery of YAP/TAZ as key molecular engines at the heart of GSC biology.

Results
A gene expression program identifying native GSCs. To identify native GSC-like cells we aimed at visualizing the natural trajectory of differentiation of GBM cells. For this, we used single-cell (sc) RNA-seq profiles from 32 patients with IDH wild-type GBM and then validated our conclusions in large patient cohorts of TCGA and REMBRANDT datasets.

We started our investigation using Monocle, an algorithm allowing the ordering of cells in a trajectory on the basis of ‘pseudotime’, a quantitative measure of the progress along a biological process defined by changes in cellular transcriptional programs. By applying this method to a scRNA-seq study of primary GBM samples (Darmanis dataset), we found that neoplastic cells organized along a tripartite trajectory (Fig. 1a and Extended Data Fig. 1a). To gain more insights on the nature of the cells that are at the opposite ends of this pseudotime trajectory (Fig. 1b), we compared their transcriptomes. Taking advantage of signatures derived from scRNA-seq studies of human neural development (Supplementary Table 1), we first found that neoplastic cells endowed with the highest pseudotime value were enriched for markers of differentiated neural cells (astrocytes and neurons) or of committed precursors, both oligodendrocyte progenitor cells (OPCs) and intermediate...
neuronal precursors (INPs), as defined by gene set enrichment analysis (GSEA) (Extended Data Fig. 1b). Thus, these neoplastic cells likely represent a population of differentiated glioblastoma cells (DGCs). Conversely, cells with the lowest pseudotime value were enriched for markers of neural stem cells (NSCs) and of early neural progenitors, such as the outer radial glia (oRG), a primate-specific neural progenitor cell type endowed with migratory properties11 (Extended Data Fig. 1b,c). These cells expressed higher levels of markers previously associated with GSCs and NSCs, such as nestin12, vimentin11, integrin-α6 (ref. 13) and SOCS3 (ref. 14) and lower levels of INP markers in comparison to DGCs (Extended Data Fig. 1c). All together, these data support the identification of the cell population at the start of the pseudotime trajectory as prospective native GSCs.

Although Monocle is one of the commonly used tools for delineating cell differentiation trajectories, it should be noted that such analyses do not provide a sense of direction. For this, we complemented the above analyses with ‘RNA velocity’15, a tool allowing the reconstruction of the differentiation trajectory of sole neoplastic cells. Based on the dynamic of mRNA processing and degradation, this algorithm returns, for each cell, a vector (visualized as an arrow) that ultimately indicates the direction of the differentiation. As shown in Extended Data Fig. 1d, these vectors identify a differentiation trajectory that starts from GSCs (red dots) and gradually transits into DGCs (blue dots), as such, solidifying the results previously obtained with Monocle.

At this point, we characterized the gene expression program typifying this GSC-like cell population. We derived a gene signature, hereafter G-STEM, consisting of the genes more significantly upregulated in this cell population when compared to DGCs (Fig. 1c and Supplementary Table 2). By Monocle analyses, the G-STEM signature identifies cells at the start of the pseudotime trajectory in each individual patient (Extended Data Fig. 2a,b). Moreover, to avoid the risk of any polarization in cell trajectory caused by the presence of scRNA of normal cells, we repeated the Monocle analyses using only neoplastic cells (Darmanis dataset), finding that, also in these conditions, G-STEM identifies cells at the start of the differentiation trajectory (Extended Data Fig. 2c).

As revealed by gene ontology analysis (Extended Data Fig. 3a,b and Supplementary Table 3), G-STEM contains genes coding for proteins involved in ECM organization, cell–ECM adhesion, promotion of cell migration, control of cell proliferation and survival,
recruitment of innate immune cells, protection against immune responses and transduction of various extracellular signals. Overall, this molecular profile is consistent with the view that native GSCs entertain mutual relationships with their microenvironment. In contrast, the gene expression program of DGCs (DGC signature; Supplementary Table 2) mostly contains genes coding
for factors involved in the early steps of neuronal differentiation (Extended Data Fig. 1c), indicating that DGCs encompass neuronal precursor-like cell states.

GBMs can be classified in at least three transcriptional subtypes, defined as proneural, classical and mesenchymal; proneural and mesenchymal GBMs represent two extremes in terms of molecular marker expression and patient survival1. Subsequent studies have shown that the GSCs of proneural and mesenchymal GBMs are characterized by expression of subtype-specific cell surface markers32, raising the possibility that different GBM subtypes may originate from biologically distinct GSCs. To tackle this hypothesis, we used an independent scRNA-seq dataset from Neftel et al.33, containing 28 tumors including all GBM subtypes. Applying Monocle to the sole neoplastic cells of the Neftel dataset retrieved a complex pseudotime trajectory with many branches (Fig. 1d), consistently with the previously reported heterogeneity of GBM cells and associated plasticity. Nonetheless, at the start of the pseudotime trajectory we could identify tumor cells characterized by the high level of the G-STEM signature (Fig. 1e). Notably, by GSEA, this cell population resembled human NSCs and early neural progenitors (Extended Data Fig. 4a). Collectively, these analyses identified ostensible GSC-like cells in the largest GBM scRNA-seq collection to date. When we focused on the distinct proneural, classical or mesenchymal GBM subtypes, the G-STEM was invariably enriched in the cells at the start of the pseudotime trajectory (Extended Data Fig. 4b). Together, these data indicate that the GSC-like cells across all GBM subtypes share the G-STEM signature.

Data presented so far provided a molecular characterization of GSCs; next, we tested whether this correlated with the biological properties of prospective GSCs within GBM. Consistent with the role of GSCs in tumor initiation, aggressiveness and relapse, a number of studies highlighted that the GSC representation in tumor samples represents a considerable prognostic factor for poor clinical outcome in patients with GBM34–36, raising the possibility that the G-STEM signature may correlate with outcome in patients. To test this prediction, we applied the G-STEM for Kaplan–Meier survival analysis of patients with GBM from two large datasets, the TCGA and the REMBRANDT projects23,24. Patients were stratified in two groups, according to high or low expression of the G-STEM signature. As shown in Extended Data Fig. 4c, high levels of G-STEM expression are indeed predictive of worse outcome in both datasets. We conclude from this collective set of results that the G-STEM signature represents a new transcriptional program that identifies native GSCs at the top of the differentiation hierarchy in patients with GBM.

Identification of candidate master transcriptional regulators of the GSC state. Next, we aimed to attain a deeper understanding of the molecular nature of native GSCs, whose gene expression and functional attributes ultimately depend on transcriptional regulators (TRs) regulating each other and their downstream target genes (‘regulons’), as such defining gene regulatory networks (GRNs)37. Here we aimed at the identification of master TRs of the GSC state. This quest was challenged by the fact that computational tools able to unbiasedly infer master TRs from scRNA-seq data are currently under-developed; indeed, the intrinsic characteristics of scRNA data (high variability per gene detected between cells and high rates of zero values (dropouts) inherent to mRNA undersampling38), present great technical challenges for TR inference. To overcome these limitations, we combined the established algorithms of ARACNE and VIPER in a multistep computational pipeline, named ‘Rhabdomant’ (Extended Data Fig. 5a and Methods), which first reconstructs a coarse GRN from single-cell gene expression profiles and transcription factors active in GBM (step 1), then ‘prunes’ this GRN by anchoring inferred regulatory interactions to putative direct target genes of TRs (step 2) and finally prioritizes master TRs from the differential enrichment of their regulons in different cell states (step 3; Extended Data Fig. 5a).

Of the list of TRs we obtained from the Rhabdomant pipeline (Supplementary Table 4), we decided to focus our attention on the 27 TRs with the largest regulons (Extended Data Fig. 5b), controlling more than 95% of the GRN (1,409 out of 1,465 genes). Of these TRs, 15 were candidate ‘master’ TRs of the GSC state, indeed cumulatively controlling a large part (96%) of the G-STEM signature. Several of these factors (SALL2, NOTCH2, ETV5 and FOXO1/3) are known regulators of stemness properties in both GSCs and NSCs alike39–41. As for the DGC state, we retrieved seven TRs, most of which (TCF12, NFIA, NFIB, SOX9 and SOX4) are in fact
known to be involved in NSC differentiation toward neuronal or glial fates\textsuperscript{32–34}. Together these results nicely validate the ability of our computational approach to identify biologically meaningful TRs of cell states out of single-cell transcriptomic data.

Scoring the list of the most significant (false discovery rate (FDR) < 0.0001) candidate master TRs of the GSC state, we focused our attention on the transcriptional coactivators YAP1 and WWTR1 (also known as TAZ). In epithelial tumors, YAP/TAZ are essential for tumor initiation and progression by inducing stemness, proliferation and chemoresistance\textsuperscript{35}. Notably, YAP/TAZ activation has recently emerged as hub for tumor–stromal interaction\textsuperscript{35}, integrating multiple inputs, including mechanical signaling and hypoxia, which are indeed profoundly dysregulated in the GBM microenvironment and are associated with GBM recurrence\textsuperscript{1,16,18}. YAP/TAZ have been reported to be regulated by CD109 (ref. 19), a marker of mesenchymal GBM, although the functional significance of this regulation remains undefined. Moreover, elevated YAP/TAZ expression levels have been noted in GBM and correlated with shorter survival of patients with glioma\textsuperscript{36,37}, but their functional involvement in GSCs remains unexplored. Our interest on YAP/TAZ was further motivated by the fact that AP-1 family members also scored at the top of the candidate master TRs of the GSC state in our analyses. Indeed, a series of recent findings have revealed that AP-1 is a pervasive transcriptional partner of YAP/TAZ on a vast number of cis-regulatory elements and that AP-1 is functionally required for YAP/TAZ responses\textsuperscript{38–40}. Collectively, these considerations prompted us to focus on YAP/TAZ activity as a candidate overarching factor in defining the GSC state.

**YAP/TAZ activity is associated with the GSC state.** We addressed more directly whether YAP/TAZ are indeed specifically active in native GSCs. We thus zoomed into the GRN architecture and asked to what extent the GSC state, as defined by G-STEM, could be explained by placing YAP/TAZ at the center of a gene regulatory cascade controlling progressive layers of downstream TRs and their targets (Fig. 2a). The YAP/TAZ-controlled GRN structure contains FOXO1 as immediate YAP/TAZ regulated downstream TR; FOXO1 regulates FOS and SALL1 that in turn regulate FOXO3, BCL6 and ERF. Collectively the target genes of these TRs account for a remarkable 73% of the whole G-STEM gene list, with

-1.9 0 1.9

YAP/TAZ was further motivated by the fact that AP-1 family members also scored at the top of the candidate master TRs of the GSC state in our analyses. Indeed, a series of recent findings have revealed that AP-1 is a pervasive transcriptional partner of YAP/TAZ on a vast number of cis-regulatory elements and that AP-1 is functionally required for YAP/TAZ responses\textsuperscript{38–40}. Collectively, these considerations prompted us to focus on YAP/TAZ activity as a candidate overarching factor in defining the GSC state.

**YAP/TAZ activity is associated with the GSC state.** We addressed more directly whether YAP/TAZ are indeed specifically active in native GSCs. We thus zoomed into the GRN architecture and asked to what extent the GSC state, as defined by G-STEM, could be explained by placing YAP/TAZ at the center of a gene regulatory cascade controlling progressive layers of downstream TRs and their targets (Fig. 2a). The YAP/TAZ-controlled GRN structure contains FOXO1 as immediate YAP/TAZ regulated downstream TR; FOXO1 regulates FOS and SALL1 that in turn regulate FOXO3, BCL6 and ERF. Collectively the target genes of these TRs account for a remarkable 73% of the whole G-STEM gene list, with

-1.9 0 1.9

YAP/TAZ was further motivated by the fact that AP-1 family members also scored at the top of the candidate master TRs of the GSC state in our analyses. Indeed, a series of recent findings have revealed that AP-1 is a pervasive transcriptional partner of YAP/TAZ on a vast number of cis-regulatory elements and that AP-1 is functionally required for YAP/TAZ responses\textsuperscript{38–40}. Collectively, these considerations prompted us to focus on YAP/TAZ activity as a candidate overarching factor in defining the GSC state.

**YAP/TAZ activity is associated with the GSC state.** We addressed more directly whether YAP/TAZ are indeed specifically active in native GSCs. We thus zoomed into the GRN architecture and asked to what extent the GSC state, as defined by G-STEM, could be explained by placing YAP/TAZ at the center of a gene regulatory cascade controlling progressive layers of downstream TRs and their targets (Fig. 2a). The YAP/TAZ-controlled GRN structure contains FOXO1 as immediate YAP/TAZ regulated downstream TR; FOXO1 regulates FOS and SALL1 that in turn regulate FOXO3, BCL6 and ERF. Collectively the target genes of these TRs account for a remarkable 73% of the whole G-STEM gene list, with
YAP/TAZ alone directly contending one-third of this network (Supplementary Table 5).

Next, we addressed experimentally to what extent YAP/TAZ are in fact required for G-STEM expression and, more broadly, for GSC biology in vivo. For this, we triggered YAP/TAZ knockout in pre-established, full blown GBM-like neoplasms derived from transformed cells bearing different oncogenic insults (KRasG12V/shp53, HER2CA or shNF1/shp53); as detailed below, these transformed cells display GSC-like properties, including the ability to generate orthotopic and subcutaneous tumors driven by YAP/TAZ and displaying several features of human GBM. To assess the effects of YAP/TAZ depletion in pre-established tumors, we opted for subcutaneous models33, as this setup allows careful monitoring of tumor growth and unambiguous retrieval of neoplastic cells. While YAP/TAZ wild-type tumors kept expanding, YAP/TAZ knockout halted tumor growth (Extended Data Fig. 6a–e), a result consistent with loss of GSCs in tumorigenic cell populations12,34. From YAP/TAZ wild-type tumors (four out of four mice), we could retrieve gliomasphere-forming cells that could be passaged ex vivo and that are able to re-initiate tumorigenesis in vivo (Extended Data Fig. 6f).

In contrast, no gliomasphere-forming cells could be retrieved from YAP/TAZ-knockout tumors (n = 5). Notably, at the molecular level, YAP/TAZ inactivation in tumors caused loss of G-STEM expression (Fig. 2b) and also led to the collapse of the TR architecture identified in our GRN as operating downstream of YAP/TAZ (with loss of FOXO1, SALL1, FOS, FOXO3, BCL6 and ERF expression; Fig. 2c).

Collectively the findings indicate that YAP/TAZ are required to preserve the GSC state, which is identified by a YAP/TAZ-dependent G-STEM signature in vivo.

To further validate the connection between YAP/TAZ activity and GSCs, we quantified by immunohistochemistry (IHC) active TAZ (nuclear and stabilized) in human GBM sections, revealing massive activation in the perinecrotic areas and a progressively more salt-and-pepper, heterogeneous staining toward the tumor periphery (Fig. 2d); these findings are nicely consistent with previous reports suggesting that the GBM perinecrotic areas are indeed enriched in GSC representation18. Moreover, we also detected the highest levels of G-STEM in perinecrotic areas of the Ivy Atlas42, a molecular pathology atlas providing gene expression data from human GBM after laser microdissection and RNA-seq of different histologically defined tumor areas (Fig. 2c). Thus, G-STEM expression peaks in the same tumor areas where active TAZ peaks.

YAP/TAZ are required for oncogenes to confer GSC properties to normal neural cells. Given the activation of YAP/TAZ transcriptional programs in prospective GSCs, we next asked whether...

Fig. 4 | YAP/TAZ control GBM cell plasticity. a, Schematic representation of the experimental setup used to promote differentiation of HuTu cells and then revert them back to a dedifferentiated state. b–e, Effects of TAZ depletion on the plasticity of HuTu10 and HuTu13 cells subjected to the differentiation/de-differentiation protocol depicted in a. Representative GFAP and TAZ staining (scale bars, 100 μm) (b). Quantifications of the percentage of cells showing predominantly nuclear ‘N’ or predominantly cytoplasmic ‘C’ TAZ localization (c). Data are representative of at least 200 cells for each condition. Western blot analysis for GFAP and YAP/TAZ; GAPDH serves as loading control (d). P values were the highest and lowest values. For each condition, experiments were repeated four times with two independent replicas. All data are plotted.

Data are representative of at least 200 cells for each condition. Western blot analysis for GFAP and YAP/TAZ; GAPDH serves as loading control (d). P values were the highest and lowest values. For each condition, experiments were repeated four times with two independent replicas. All data are plotted. P values were determined by one-way ANOVA with Dunnett’s T3 multiple comparisons.
YAP/TAZ activity may represent an early addiction in gliogenesis. Compelling evidence indicate that the cells of origin of GBM include various cell types, including astrocytes\textsuperscript{43,44}, OPCs\textsuperscript{45–47} and, as recently described in human GBMs, cells of the subventricular zone (SVZ)\textsuperscript{48} that display mixed astrocyte-like and NSC-like features, such as expression of GFAP, NESTIN and SOX2 (ref. \textsuperscript{48} and

---

**Fig. 5 | YAP/TAZ are required to prevent GSC differentiation.** a–e, Gliomaspheres derived from PDGF\textsubscript{a}CA (a), shNF1/shp53 (b), EGFRC\textsubscript{a} (c), HER2CA (d) or KRasG12V/shp53 (e) transformed R26\textsuperscript{CaG-CreERT2};Yap\textsuperscript{fl/fl};Taz\textsuperscript{fl/fl} newborn astroglial cells were treated with either ethanol (YAP/TAZ wt) or 4OH-TAM (YAP/TAZ KO). Shown are representative images (left; scale bar, 100 µm) and quantifications (right; mean and individual data points of two independent experiments, each performed with two replicas) of the number of gliomaspheres per cm\textsuperscript{2} in vehicle versus TAM-treated samples. See also Extended Data Fig. 9a for a specificity control, showing that, in absence of CREER-expression, 4OH-TAM does not induce gliomasphere disaggregation. f, Analysis of RNA-seq data from gliomaspheres derived from KRasG12V/shp53-transformed R26\textsuperscript{CaG-CreERT2};Yap\textsuperscript{fl/fl};Taz\textsuperscript{fl/fl} newborn astroglial cells and treated either with vehicle or with 4OH-TAM as described above. The graph shows average log\textsubscript{2} gene expression changes of signatures for the indicated cell types (astrocytes, \(n=44\) genes; NSCs, \(n=89\) genes; INPs, \(n=119\) genes; neuroblasts, \(n=37\) genes) or for proliferating neural progenitors (proliferation, \(n=66\) genes) in 4OH-TAM-treated (YAP/TAZ KO) KRasG12V/shp53 gliomaspheres compared to vehicle-treated (YAP/TAZ wt) KRasG12V/shp53 gliomaspheres. Data are shown as mean±s.e.m. Positive and negative values indicate upregulation and downregulation, respectively of the indicated signatures after YAP/TAZ KO. \(P\) values were determined by Brown–Forsythe and Welch one-way ANOVA test with Dunnett’s T3 multiple comparisons of the distribution of log\textsubscript{2} gene expression changes of each signature with the distribution of log\textsubscript{2} gene expression changes for all expressed genes (\(n=12,211\)).
Articles

Extended Data Fig. 7d). Typical oncogenic lesions in GBM entail overactivation of the RTK/RAS pathway through genetic amplification or activating mutations of RTK-coding genes (PDGFRα, EGFR and HER2) or mutation of the RAS inhibitor NF1 (ref. 40).

To recapitulate, at least in part, the early stage of gliomagenesis, we introduced activated oncogenes in astroglial cells from newborn mice that are highly similar to the SVZ cell population29, expressing mixed astrocyte-like (GFAP) and NSC markers (NESTIN, SOX2) (Extended Data Fig. 7b,c). Newborn astroglial cells were transduced with lentivector vectors coding for the activated forms of PDGFR-α, EGFR or HER2, with vectors coding for oncogenic KRasG12V plus short hairpin (sh)RNA against p53 or shRNAs against NF1 and p53; this list includes established drivers of human GBM40. After approximately 3 weeks, gliomaspheres emerged from all oncogene-expressing primary astroglial monolayers and never from controls (Fig. 3a). Gliomaspheres could be dissociated as single cells and expanded over several passages in suspension (Extended Data Fig. 7d); when transplanted orthotopically (intracranially) in immunocompromised mice, dissociated gliomasphere cells gave rise to tumors well-recapitulating key histological features of human GBM, also resembling giant cell glioblastoma29 and displaying elevated levels of active TAZ in the perinecrotic regions (Extended Data Fig. 7e). Similar tumors emerged after subcutaneous cell transplantation, where they also included areas of necrosis surrounded by pseudopalisading cells, also common features of human GBM (Extended Data Fig. 7f).

Upon limiting-dilution transplantation, we found that 1,000 dissociated gliomasphere cells were sufficient to seed tumor formation (Extended Data Fig. 7g). When explanted and re-cultured ex vivo, tumor cells remained able to form gliomaspheres that could be expanded again over several passages in culture, retaining the capacity of tumor initiation upon serial transplantation in vivo (Extended Data Fig. 7h). As control, we could not detect any formation of spheroids from control newborn astroglial cells ex vivo (Fig. 3a); then, no outgrowth whatsoever could be detected when these parental cells were injected in recipient mice. In other words, overexpression of RAS/RTK oncogenes in otherwise normal primary cells specifically converts them into cells endowed with typical properties of GSCs, such as tumorigenic potential and the ability to undergo self-renewal, in vitro and in vivo. In accordance, by RNA-seq, tumors displayed elevated expression of the G-STEM signature and low expression of the DGC signature when compared with the original astroglial cell cultures (Extended Data Fig. 7i).

Are YAP/TAZ activated by oncogenic mutations during acquisition of a GSC-like state? To address this question, we first verified that YAP/TAZ were activated during the early phases of reprogramming of newborn astroglial cells into GSC-like cells by oncogenes. For this, we transduced early postnatal astroglial cells with 8xGTIIC-RFPDD, a YAP/TAZ-responsive lentiviral reporter41, and then monitored the expression of red fluorescent protein (RFP) during oncogene-mediated reprogramming. As shown in Fig. 3b, oncogenes specifically induced activation of the reporter within 24 h of reprogramming; RFP-positive cells acquired elongated shapes reminiscent of that of the radial glia, spreading and migrating, typically converging toward one cell cluster (Extended Data Fig. 8a). Overall, these data indicate that YAP/TAZ activation is an early event during reprogramming of normal astroglial cells into tumorigenic GSC-like cells.

We next asked whether YAP/TAZ are required for oncogene-mediated reprogramming of normal cells into GSC-like cells. To this end, we derived astroglial cells from newborn R26<sup>creERT2</sup>, Yap<sup>fl/fl</sup>, Taz<sup>fl/fl</sup> mice, in which genetic ablation of YAP/TAZ was induced after CRE activation by 4OHT-tamoxifen (4OH-TAM) treatment. As shown in Fig. 3c, gliomasphere formation by various RTK oncogenes, either transduced alone or in combination with shRNA-targeting tumor suppressors, was completely abolished by YAP/TAZ deletion.

Next, we investigated at the molecular level how YAP/TAZ contribute to oncogenic reprogramming of normal neural cells into GSCs. We performed transcriptomic analyses by RNA-seq of astrocytes from Yap<sup>fl/fl</sup>,Taz<sup>fl/fl</sup> mice transduced with lentivector vectors encoding CreERT2 and a doxycycline-inducible HER2CA, in absence or presence of 4OH-TAM at the incipit of cell transformation. By hierarchical clustering of samples using genes whose expression is significantly altered by HER2CA, we found that the transcriptome of YAP/TAZ-knockout cells remained remarkably similar to that of control/nontransformed cells in spite of HER2 overexpression (Fig. 3d). Specifically, HER2CA expression in newborn astroglial cells induces YAP/TAZ-dependent downregulation of astrocyte markers and upregulation of NSCs markers (compare lanes 1 and 2 of Fig. 3e with lanes 1 and 2 of Fig. 3f). Notably, 45 out of 46 NSCs marker genes upregulated by HER2CA in control cells are dependent on YAP/TAZ (Fig. 3g); however, at the incipit...
of oncogenic transformation, YAP/TAZ seem dispensable for the expression of genes involved in proliferation (Fig. 3f; lane 4). Overall, these results indicate that the main function of YAP/TAZ during GBM initiation is to repress differentiation and promote the acquisition of NSC-like properties.

**YAP/TAZ are required for the intrinsic differentiation plasticity of GBM cells.** One of the most lethal properties of GBM cells is their intrinsic plasticity, allowing for interconversion between GSC and non-GSC states depending on a number of factors, as such promoting tumor relapse. We thus tested the role of YAP/TAZ in GBM plasticity. For this, we took advantage of a previously reported differentiation/de-differentiation protocol for human GBM cells (Fig. 4a). Specifically, we used two independent primary GBM cell lines, HuTu10 and HuTu13 (ref. ), corresponding, by transcriptomics, to proneural and mesenchymal subtypes, respectively.

When cultured in serum-free, growth-factor-rich conditions, both cell types display no expression of the astrocyte marker GFAP; however, after exposure to serum and BMP2-containing medium, cells en masse differentiate into GFAP-positive cells with a stellate morphology, reminiscent of normal astrocytes (Fig. 4b,d,e; compare lanes 1 and 2). This event is accompanied by a dramatic decrease in the ability to form gliomaspheres (Fig. 4f,g; compare lanes 1 and 2), indicating that acquisition of astrocyte-like properties is associated with loss of stemness properties. Of note, this process is also accompanied by relocalization of TAZ in the cytoplasm, indicating that YAP/TAZ are inactivated during differentiation (Fig. 4b,c; compare lanes 1 and 2). The return to a GSC-like, less-differentiated phenotype can be induced by placing cells back into serum-free, growth-factor-rich culture conditions (de-differentiation), promoting YAP/TAZ re-entry in the nucleus, progressive disappearance of GFAP and re-acquisition of gliomasphere-forming abilities.
YAP/TAZ preserve the GSC state by preventing differentiation.

Having shown above the role of YAP/TAZ at inducing GSC-like properties in otherwise normal cells at the incipit of oncogenic transformation or in differentiated tumor cells, we next asked whether YAP/TAZ are required to preserve GSC-like cells after their establishment. For this, we first monitored the effects of YAP/TAZ ablation within pre-established gliomaspheres. YAP/TAZ depletion induced progressive disaggregation and demise of gliomaspheres, irrespectively of the type of oncogenic lesion that drove gliomasphere emergence (Fig. 5a–e). After single-cell replating, YAP/TAZ-knockout cells were unable to sustain any outgrowth/passaging (Extended Data Fig. 9b), indicating that YAP/TAZ are essential for self-renewal.

To investigate at the molecular level the consequences of YAP/TAZ ablation, we first compared the transcriptomes of gliomasphere cultures arising from KrasG12V/shp53-transformed R26CAGCreERT2;Yapfl/fl;Tazfl/fl astroglial cells, either treated with vehicle control or with 4OH-TAM. Notably, loss of YAP/TAZ in gliomaspheres was accompanied by upregulation of proliferation markers (Fig. 5f; lane 4), indicating that processes other than loss of proliferation are relevant for gliomasphere demise after YAP/TAZ inactivation. As shown in Fig. 5f, upon YAP/TAZ ablation, NSCs markers were strongly downregulated, whereas markers of neuroblasts were induced (including Dlx2, Stmn1, TUJ1 and many others; Supplementary Table 6), suggesting that YAP/TAZ are primarily required in GSCs to prevent their differentiation along the neuronal lineage. Consistently, YAP/TAZ ablation in gliomaspheres also caused upregulation of markers of INPs (Fig. 5f; lane 3), that is, proliferating neuroblast precursors.

YAP/TAZ control GBM initiation and differentiation in vivo. In keeping with the relevance of YAP/TAZ at preserving GSCs-like cell population in vitro, we next verified whether YAP/TAZ are required for one of the cardinal features of GSCs in vivo, tumor initiation. We tested this idea by injecting GBM cells orthotopically, in the brain of immunocompromised mice. We first used transformed Yapfl/fl;Tazfl/fl cells carrying either shNF1/shp53 or KRasG12V/shp53 and dual luciferase-GFP expression vectors, allowing for a non-invasive readout of tumor growth. Mice injected intracranially with parental GBM cells invariably formed large tumor masses with invasive behavior, as revealed by bioluminescence and by histological analyses (Fig. 6a–d and Extended Data Fig. 10a–c). In contrast, YAP/TAZ-knockout cells failed to form any outgrowth (Fig. 6a–c and Extended Data Fig. 10a–c). Similar results were obtained in immunocompromised mice by orthotopically injected control versus YAP/TAZ-depleted human GBM cells (HuTu13 cells, Extended Data Fig. 10d–f).

Next, we validated these findings using some of the most established models of mouse GBM, the mouse glioma cell lines GL261 and CT2A. After orthotopic transplantation in syngeneic mice, these cells form tumors displaying several characteristics of human GBMs, including intratumoral heterogeneity, pseudopalisading necrosis, radio resistance and chemoresistance. To study the role of YAP/TAZ in these models, we injected GL261 or CT2A expressing anti-YAP/TAZ doxycycline-inducible shRNAs in the brain of immunocompetent syngeneic mice. Control cells formed large tumor masses, whereas upon doxycycline treatment, YAP/TAZ-depleted cells did not (Fig. 6e–g and Extended Data Fig. 10g–i).

However, in some brain sections we could still detect residual YAP/TAZ-depleted cells tumor cells, thus allowing in vivo investigation of their differentiation state by immunofluorescence. As shown in Fig. 6h for GL261 cells, YAP/TAZ depletion causes a strikingly wholesale differentiation toward the neuronal lineage, being essentially all cells positive for TUJ1. In contrast, tumors generated by parental GL261 cells are almost invariably negative for TUJ1, with the exception of a minority of cells (about 3% of tumor cells typically found as small clusters).

We then confirmed that tumor cell differentiation is also at the root of halted tumor growth after YAP/TAZ inactivation in pre-established lesions in vivo (as in Extended Data Fig. 6). As mentioned above, loss of YAP/TAZ causes depletion of GSCs and downregulation of the G-STEM signature. Remarkably, this is accompanied by massive upregulation in the expression of early markers of neuronal differentiation (for example, Ascl1, Tau, TuJ1, Ncam, Stathmin and Sox11) (Fig. 6i). By IF, YAP/TAZ-knockout tumor cells acquired the expression of TUJ1, confirming that YAP/TAZ prevents GSC differentiation, with their ablation skewing the fate of GBM cells toward a neuronal-like fate (Fig. 6j and Extended Data Fig. 10j). We conclude from this collective set of results that YAP/TAZ are required for tumorigenesis and to prevent differentiation in vivo in multiple cellular and experimental contexts, in primary astroglial/SVZ-like cells transformed with different activated oncogenes and in classic GBM models growing in a syngeneic context.

Discussion

In this work, we advance on the molecular foundations of the GSC state, identifying YAP/TAZ as the transcriptional determinants that define GSC populations in their native state. YAP/TAZ activation occurs downstream of classic oncogenic drivers of GBM to induce GSC-containing tumorigenic cell populations; within such populations, YAP/TAZ remain key for self-renewal of gliomaspheres in vitro and for both tumor initiation and maintenance in vivo, as shown in different mouse and human GBM cellular contexts. Consistently, YAP/TAZ activity, as monitored by G-STEM signature, can identify prospective GSCs in heterogeneous GBM cell populations, at least in IDH wild-type GBMs.

A recent seminal study by Neftel et al. presented a model in which cellular heterogeneity of IDH wild-type GBM reflects the coexistence, within each individual tumor, of four cellular states, able to interconvert into each other. Still unclear is whether these cell subtypes are connected to a shared GSC state or whether different stem cell populations exist in different GBM subtypes. The present identification of the G-STEM only in part advances on these open issues. Our analyses of the scRNA-seq data of Neftel et al. reveal GSC populations earmarked by the G-STEM transcriptional program in different GBM subtypes. However, projecting a shared molecular signature into cell populations should be interpreted with caution, as this may imply either of two scenarios: one in which the G-STEM indeed identifies a specific GSC population common to all GBM subtypes or alternatively, that distinct stem cells in different GBM may share part of their anti-differentiation mechanisms and that one of this overlapping program may be highlighted by elevated YAP/TAZ activity as denoted by G-STEM.

Notably, we found that these native GSCs display a hybrid phenotype between the astrocyte-like and mesenchymal-like cell states of Neftel et al. For example, we found that native GSCs are enriched in molecular markers typical of both astrocyte-like state (such as HOPX, GFAP and MLC1) and of the mesenchymal-like state (such as VIM, CD44 and LGALS3) (Supplementary Table 2). This conclusion is apparently in contrast with our observation that GSCs acquire astrocyte-like features after in vitro differentiation induced by serum. However, normal astrocytes do not express mesenchymal markers and seem transcriptionally distinct from GSCs in our
analyses. Rather, it is tempting to speculate that native GSCs may resemble fetal-like cells, such as early neural progenitors, which also coexpress markers of astrocyte (such as HOPX and GFAP) and mesenchymal-like states (such as VIM and LGALS3)11.

Restoring differentiation capacity of GBM might represent a therapeutic option54, although this is complicated by the ability of differentiated cells to revert back to a GSC state55. Here we advance in these directions by showing that YAP/TAZ activity peaks in GSCs and that targeting YAP/TAZ is instrumental to cause their irreversible conversion into committed neural progenitors and more-differentiated neural cell types. Of note, this occurs independently of proliferation control, providing a departure from current models envisioning a central role of YAP/TAZ as regulators of cell-cycle progression in cancer56. We propose that an anti-YAP/TAZ therapy has the potential to be more effective than current chemotherapeutic regimens targeting cell proliferation, whose efficacy in patients with GBM is in fact very limited. Strategies aimed at blunting YAP/TAZ activity in vivo have been recently proposed, including inhibitors of Brd4, a YAP/TAZ co-activator that represents an addiction of GBM although only in vivo57,60, consistently with the identification of YAP/TAZ activity at the core of the native GSC state.

In conclusion, we have here advanced on a key molecular underpinning of the GSC native state, thus unveiling a core vulnerability and addiction of GBM, all in all hinting at new perspectives to ameliorate treatment of a devastating malignancy.

Methods

Analysis of scRNA-seq data from primary glioblastomas. Darmans dataset. We analyzed scRNA-seq data of primary glioblastoma samples from Darmans et al.54, retaining the cell annotation provided by the authors. Raw reads were downloaded from the Gene Expression Omnibus (GEO) under accession code GSE44645 and mapped to the human reference genome GRCh38 using STAR58. Raw gene counts were obtained using the featureCounts function of the Rsubread R package59 and the Gencode release 25 (GRCh38.p7) basic gene annotation. Quality controls and normalization were carried out using Seurat60 (v2.3.1) with default parameters. We retained for subsequent analyses 3,188 cells (out of 3,588 cells) with (1) number of unique detected genes between 500 and 8,000; (2) total number of detected molecules between 1 x 10⁶ and 1.5 x 10⁷; and (3) fraction of reads mapping to the mitochondrial genome ≤0.2. Pseudotime trajectories on neoplastic and normal neural cells were constructed using Monocle2 (ref. 54; v2.8.0); we used the unsupervised `dpFeature` procedure to order cells based on genes that differ between clusters and the ‘DDRTree’ algorithm for the dimensionality reduction step. Based on the distribution of pseudotime values, we select two populations of neoplastic cells, the first with pseudotime values below the first quartile (<Q1), and the second with pseudotime values above the first quartile (>Q1). We compared the trajectories of the two populations of neoplastic cells calculating the log, fold change of the expression level for each gene. We applied the GSEA software in preranked mode to the gene list ranked on log, fold change to evaluate the functional enrichment for the same gene sets tested in the Darmans dataset (Supplementary Table 1). Gene sets were considered significantly enriched at FDR ≤0.05 when using classic enrichment statistics and 1,000 permutations of gene sets.

Signature scores. Signature scores have been calculated as the average expression of the genes comprised in each signature.

All analyses were performed in R 3.5.0.

Collection and processing of GBM transcriptomes from the TCGA and REMBRANDT studies. Gene expression data of TCGA and REMBRANDT were obtained from the GEO (Supplementary Tables 8 and 9). For the TCGA dataset, related clinical and molecular subclass data were obtained from the Rembrandt database. For the TCGA dataset, we processed gene expression signals of all samples and as G-STEM signature low and vice versa. This classification was applied to expression values of TCGA and REMBRANDT glioblastomas. To evaluate the prognostic value of the G-STEM signature, we applied the Kaplan–Meier method on the patients’ survival data to estimate the probabilities that patients classified as ‘G-STEM high’ and ‘G-STEM low’ would survive. To confirm these findings, Kaplan–Meier curves were compared using the log-rank (Mantel–Cox) test. P values were calculated according to the standard normal asymptotic distribution. Survival analysis was performed in GraphPad Prism.

Identification of candidate master transcriptional regulators. To identify candidate master transcriptional regulators we assembled a three-step computational workflow named ‘Rhabdom’. In the first step, to recover the GCRM, we initially defined a list of transcription factors active in GBM, defined as TRs potentially associated with chromatin in GBM tumors. For this, we took advantage of the epigenetic analyses provided by a recent large-scale ATAC-seq profiling of several human tumor types, including GBM51. We carried out a DNA-binding motif enrichment analysis using the MEME algorithm open over chromatin regions of GBMs and then selected the list of transcription factors whose DNA-binding motifs were highly enriched (FDR <0.0001) in these genomic regions. We manually implemented this list with partner transcriptional effectors (Supplementary Table 10) for a total list of 151 TRs.

Next, we applied the reverse-engineering algorithm ARACNe—aP52 to gene expression signals of neoplastic cells of the Darmans scRNA-seq dataset, allowing to map the interactions between GBM-specific transcription factors and their co-regulated genes. To limit the effects of scRNA-seq data sparsity, we removed
genes expressed only in a limited number of cells (normalized counts >0 in fewer than 100 cells).

In the second step, in order overcome spurious TR-target gene association intrinsic to the noise generated by scRNA-seq data, we pruned the GRN by retaining only candidate direct targets genes of each TR. For this, we took advantage of the association map of between each ATAC-seq peak and its target genes, as provided by Corces et al. Supplementary Table 11. We then intersected this map with the gene interactomes obtained in step 1, in so doing retaining in the GRN only the lists of target genes (regulons) associated to a binding motif for a candidate TR in their cis-regulatory elements.

In the third step, we interrogated the GRN with the VIPER algorithm to identify candidate master TRs of GSCs and DGCs, namely, transcriptional determinants whose regulons were enriched of genes activated in one of the two opposite cell populations of the pseudotime trajectory of neoplastic cells, that is the G-STEM and DGC signatures. Based on VIPER analysis, we defined as candidate opposite cell populations of the pseudotime trajectory of neoplastic cells, that is the Ivy Glioblastoma Atlas Project (Supplementary Table 12). Gene expression RNA-seq data from six (Supplementary Table 5). The statistical significance of the overlaps was calculated using the featureCounts function of the Rsubread R package.

To compare G-STEM signature and master TR regulomes. To compare the G-STEM signature with the GRM regulomes, we calculated the enrichment of G-STEM signatures in the regulomes of YAP/TAZ TRs. Specifically, we considered the three layers of the YAP/TAZ gene regulatory network and identified six TRs (FOXO1, FOS, SALL1, FOXO3, BCL6 and ERF); then we calculated the intersection between the 313 G-STEM signature genes contained in the GRM regulons resulting from step 2 and the targets positively interacting (mode of action >0) with YAP/TAZ or one of their downstream TRs (Supplementary Table 5). The statistical significance of the overlaps was calculated in R using the fisher.test function of the stats package.

Comparison of samples from the Ivy Glioblastoma Atlas Project. RNA-seq data from six GBM were downloaded from the Anatomical Structures RNA-Seq repository of the Ivy Glioblastoma Atlas Project (Supplementary Table 12). Gene expression was quantified in R 3.3.1 using the featureCounts function of the Rsubread R package26 and the UCSC gene annotation (GRCh37/hg19). Data normalization was performed using the edger package (v.3.20.0); briefly, raw counts were normalized to counts per million (CPM) mapped reads and to fragments per kilobase per million mapped reads (FPKM). Gene expression data were then standardized between the samples composing each tumor. Signature scores were calculated as the average expression of the genes comprised in each signature. For each area of each tumor, the signature scores were calculated as the average of the signature scores of the different samples of the same area from the same tumor.

TAZ immunohistochemistry. Archived frozen GBM specimens were collected at the Azienda Ospedaliera di Padova. For IHC, 4-μm thick sections were obtained from tumor samples. IHC was performed with rabbit polyclonal anti-TAZ (Sigma, HPA007415; 1:50 dilution) as previously described73. For mouse tissues, IHC was performed with rabbit polyclonal anti-TAZ (Sigma, HPA007415; 1:50 dilution) as previously described73.

For the experiments depicted in Fig. 3d–g, astroglial cells from embryonic day 18 (E18) female C57BL/6 mice were used for intracranial injections.

Transgenic lines used in the experiments were provided by D. Pan (Yapfrt6/6 mice); D. Sauv and J. Siveke (R26-CAG-FGF-CreER mice) and P. Bonaldo (CMV-Fliplike mice). Double Yapfrt6/6, Tazfrt6/6 conditional KO mice were as described elsewhere. To obtain R26-CAG-Fliplike mice with CMV-Fliplike mice, R26-CAG-Creb mice were intercrossed with Yapfrt6/6, Tazfrt6/6 mice.

Glomiaspersion preparation. Primary newborn astroglial cells were isolated and maintained as previously described. For glomiaspersion preparation, astroglial cells were plated at 20–30% confluence in six-well plates in 2 ml astrocyte medium (DMEM medium supplemented with 10% FBS, glutamine and antibiotics). The next day, cells were transduced with lentiviral vectors coding for oncogenes or with empty vector as negative controls. After 24 h (day 3), transduced astrocytes were switched to NSC medium (DMEM/F12 supplemented with 100X N2, 20ng ml−1 EGF, 20ng ml−1 bFGF, glutamine and antibiotics). Spheres arising from the cell monolayer were evident after approximately 3 weeks. Sphere passaging was performed as described previously. To evaluate self-renewal properties (Extended Data Figs. 7d and 9b), glomalospheres were dissociated to single cells and replated in Ultra-Low-Attachment 24-well plates (Corning), at the concentration of 2,000 cells per well, fully formed glomalospheres were counted by visual inspection.

NATURE CANCER | VOL 2 | FEBRUARY 2021 | 174–188 | www.nature.com/natcancer

For doxycycline-inducible downregulation of mouse YAP and TAZ in GL261 and CT2A cell lines, we used Tet-pLKO-puro (a kind gift from G. Del Sal) lentiviral vectors expressing doxycycline-inducible sh-mouseTAZ in combination with Tet-pLKO-hygro lentiviral vectors (obtained by substituting the puromycin resistance cassette into Tet-pLKO-puro with the hygromycin resistance cassette from pBABEBhygro) expressing doxycycline-inducible sh-mouseYAP. Tet-pLKO-puro shControl lentiviral vectors were used as control.

For inducible transgenes, cells were treated with 2 μg ml−1 doxycycline for the whole duration of the experiments.

RNAseq performed with ipilimumab RNA-MAX (Thermo Fisher Scientific) in antibiotic-free medium according to manufacturer instructions. Sequences of siRNAs are provided in Supplementary Table 13. Lentiviral particle preparation and cell culture infections were as described elsewhere.

Mice. Animal experiments were performed adhering to our institutional and national guidelines as approved by OPBA (Padova) and the Ministry of Health of Italy. The housing conditions comprised a diet with 28% protein. A maximum of five adult mice weighing up to 20 g were housed in a single cage, maintaining the ambient temperature at 19–23 °C, the humidity at 55 ± 10% and a 12-h light/12-h dark cycle.

Female 6–8-week-old NOD-SCID mice (Charles River) were used for subcutaneous injections and female 6–8-week-old NSG or C57BL/6 mice (Charles River) were used for intracranial injections.

To obtain R26-CAG-Fliplike mice with CMV-Fliplike mice, R26-CAG-Creb mice were intercrossed with Yapfrt6/6, Tazfrt6/6 mice.

Glomiaspersion preparation. Primary newborn astroglial cells were isolated and maintained as previously described. For glomiaspersion preparation, astroglial cells were plated at 20–30% confluence in six-well plates in 2 ml astrocyte medium (DMEM medium supplemented with 10% FBS, glutamine and antibiotics). The next day, cells were transduced with lentiviral vectors coding for oncogenes or with empty vector as negative controls. After 24 h (day 3), transduced astrocytes were switched to NSC medium (DMEM/F12 supplemented with 100X N2, 20 ng ml−1 EGF, 20 ng ml−1 bFGF, glutamine and antibiotics). Spheres arising from the cell monolayer were evident after approximately 3 weeks. Sphere passaging was performed as described previously. To evaluate self-renewal properties (Extended Data Figs. 7d and 9b), glomalospheres were dissociated to single cells and replated in Ultra-Low-Attachment 24-well plates (Corning), at the concentration of 2,000 cells per well, fully formed glomalospheres were counted by visual inspection.

Monitoring YAP/TAZ activity during gliomagenesis. Newborn mouse astroglial cells were plated at 20–30% confluence in six-well plates in 2 ml of astrocyte medium. At 24 h after seeding, cells were infected with the 8xGTTIC-RFP-DFD lentiviral vector. Transduced cells were infected with oncogene-expressing or empty lentiviral vectors and, the day after, switched to NSC medium supplemented with trimethoprim (10 μM). Bright-field and fluorescent images were acquired daily using a Leica DMIL LED microscope equipped with a Leica DFC 3000G camera using LAS AF version X software.

For the experiments depicted in Fig. 3d–g, astroglial cells from Yapfrt6/6, Tazfrt6/6 mice were transduced with LV-CreERT2 and then with FUdeltaGW-tRTA and doxy-inducible HER2CA lentiviruses. After 24 h, infected newborn astroglial cells were switched to NSC medium containing 2 μg ml−1 doxycycline to induce

For the experiments depicted in Fig. 3d–g, astroglial cells from Yapfrt6/6, Tazfrt6/6 mice were transduced with LV-CreERT2 and then with FUdeltaGW-tRTA and doxy-inducible HER2CA lentiviruses. After 24 h, infected newborn astroglial cells were switched to NSC medium containing 2 μg ml−1 doxycycline to induce
HER2CA or 1 µM 4OH-TAM to induce YAP/TAZ knockout. Medium was replaced every 3–4 d, till day 19, when cells were collected for RNA extraction. Negative controls were provided by astroglial cells transduced with empty vector.

**Testing YAP/TAZ requirement for gliomastere maintenance.** Single glioblastoma cells obtained from R2G<sup>CreERT2</sup>, Yap<sup>fl/fl</sup>, Tat<sup>dox</sup> P2 gliomaspheres were seeded to form P3 gliomaspheres. Fully formed P3 gliomaspheres were treated with either vehicle (ethanol) or 1 µM 4OH-TAM. Fresh 4OH-TAM was added every 3 d and sphere morphology and size was evaluated 3 d and 1 week after the first treatment. Bright-field images were acquired with a Leica DMIL LED microscope equipped with a Leica DFC 3000G camera using LAS version X software.

**Subcutaneous tumor experiments.** Oncogene-induced P2/P3 gliomaspheres were dissociated into single cells with Tryple, resuspended in ice-cold Matrigel and injected into the flank of NOD/SCID mice (750,000 cells/200 μl Matrigel). Tumor growth was followed over time and masses were collected for histological analyses when they reached 1–2 cm in diameter, before any apparent ulceration of the skin.

To test the requirement of YAP/TAZ on the maintenance of subcutaneous tumors, we injected in the flank of NOD/SCID mice single cells dissociated from oncogene-induced R2G<sup>CreERT2</sup>, Yap<sup>fl/fl</sup>, Tat<sup>dox</sup> P2–P3 gliomaspheres (750,000 cells/200 μl Matrigel). At the appearance of palpable masses, a group of mice received a TAM400/CreER diet (Envigo) to promote YAP/TAZ KO in tumor cells (YAP/TAZ KO), whereas a second group, serving as control, continued to be fed with normal diet (YAP/TAZ wt). Tumor growth was followed over time and masses were collected for histological analyses when YAP/TAZ wt tumors reached 1–2 cm in diameter.

For serial passaging experiments, cell cultures were obtained from primary tumor masses and expanded in vitro as gliomaspheres in NSC medium for three passages. Gliomaspheres were then dissociated into single cells and subcutaneously transplanted (10,000 cells/200 μl Matrigel) for secondary tumor formation. This procedure was repeated for tertiary tumor formation of Exttar tumours for experiments described in Extended Data Fig. 7b.

For the experiments in Extended Data Fig. 7g, P3 KRasG12V/shp53 gliomaspheres were dissociated with Tryple, resuspended at different cell dilutions in 200 μl ice-cold Matrigel (100,000; 10,000; 1,000; 100 cells) and injected into the flank of NOD/SCID mice.

Tumor growth was followed over time and masses were collected for histological analyses when they reached 1–2 cm in diameter, before any apparent ulceration of the skin.

**Brain tumor experiments.** For experiments depicted in Fig. 6a–d and Extended Data Fig. 8a–c, single cells obtained from oncogene-induced Yopg<sup>CreERT2</sup>, Tat<sup>dox</sup> gliomaspheres were transduced with adenoviruses encoding for CRE recombinase (Ad-Cre) to induce YAP/TAZ KO or for GFP (Ad-GFP) as negative controls. Cells were then orthotopically injected into NSG mice (300,000 cells/2 μl PBS).

For the experiments in Extended Data Fig. 10d–f, HuTu13 cells were transduced with shControl or shYAP/TAZ for 48 h and then orthotopically injected into 6–8-week-old NSG mice (300,000 cells/2 μl PBS). For the experiments in Fig. 6e and in Extended Data Fig. 10g–i, CT2A or GL261 shControl and shYAP/TAZ lines were generated by infecting glioma cells either with the pLKO-hygro-puro-Tet-On-shControl or with pLKO-hygro-puro-Tet-On-shYAP and pLKO-puro-Tet-On-shTat lentiviral particles, respectively; cells were then transduced by hygromycin/puromycin-containing medium for 10 d to select transduced cells. GL261 and CT2A cells were then treated for 48 h with 2 μg/ml doxycycline and then orthotopically injected into the brain of 6–8-week-old C57BL/6J mice (300,000 cells/2 μl PBS). To sustain YAP/TAZ depletion after injection, doxycycline was added to the drinking water of all mice.

Before injection, all cells were transduced with a lentiviral construct coding for eGFP and firefly luciferase (GFP/Luc)<sup>+</sup>. The procedure and the coordinates used for the injection were as described previously<sup>18</sup>. Brain tumor growth was monitored by in vivo luciferase assay, by intraperitoneal injection of 150 μg·kg<sup>−1</sup> of XenoLight D-luciferin in PBS (PerkinElmer) and detecting brain luminescence equipped with a Leica DFC 3000G camera using LAS version X software.

**HuTu cell plasticity assay.** HuTu cells were classified on the basis of their transcriptional profiles (centroids are reported in Supplementary Table 14). Details of the procedure are provided as Protocol Exchange at https://doi.org/10.21203/rs.3.pek-1218/v1.

HuTu cells were plated on fibronectin-coated six-well plates and cultured in stem medium (DMEM/F12, 10% B27, 25 ng/ml bFGF, 25 ng/ml bFGF, glutamine and antibiotics). Differentiation was established by switching to differentiation medium (DMEM/F12, 10% FBS, BMP2 50 ng/ml<sup>−1</sup>, glutamine and antibiotics) for 15 d. De-differentiation was induced through medium switch to stem medium. When indicated, siRNA transfections were performed the last day of differentiation, before starting de-differentiation. De-differentiation was maintained for 5 d before collecting for western blot. For immunofluorescence, at the end of the de-differentiation process cells were plated on fibronectin-coated glass slides for 24 h and then fixed with 4% PFA.

For experiments in Fig. 4l, cells were transfected with control siRNA or with two independent YAP/TAZ siRNA mixes after 15 d of differentiation. Cells were then plated in low-attachment 24-well plates (2,000 cells per well) in HuTu stem medium; growing spheres were counted after 4 d.

**Gene expression analyses by RNA-seq.** RNA extraction from cells was performed with Nucleospin 8 RNA Core kit (Macherey-Nagel) according to the manufacturer’s instructions, using an automated system (Freedom EVO, Tecan). Preparation and sequencing of RNA-seq libraries are described elsewhere<sup>19</sup>. Raw reads were mapped to the mouse reference genome (GRCm38) using STAR<sup>20</sup>. Raw gene counts were obtained using the featureCounts function of the Rsubread R package<sup>21</sup>; only genes with CPM > 1 in at least one sample (or two samples when replicates were available) were further retained for differential analysis. Differential gene expression analysis was performed using the exactTest function of the edgeR package<sup>22</sup>.

Hierarchical clustering of Fig. 5d was performed using the Hierarchical Clustering of the Multi-Experiment Viewer (MeV 4.8.4) package with Pearson correlation as distance metric and average linkage clustering, using the row-wise standardized FPKM of genes significantly (FDR < 0.05) upregulated (fold change ≥ 1.33; FPKM ≥ 1 in YAP/TAZ wt HER2CA-expressing astrocytes) or downregulated (fold change ≤ 0.75; FPKM ≥ 1 in control newborn astroglial cells) in HER2CA-expressing astrocytes expressing HER2CA compared to control newborn astroglial cells. Gene expression heat maps have been generated in GraphPad Prism 8.0.2 software using row-wise standardization of the expression values.

Average log, gene expression changes were calculated as the standardized average log, fold change of signature genes in all samples and plotted as mean ± s.e.m. For Fig. 5e, displaying log, gene expression changes between HER2CA or tetracyclin and control astrocytes, we considered only genes expressed either in HER2CA-expressing astrocytes for upregulated genes or in control astrocytes for downregulated genes, that is, genes displaying FPKM ≥ 1 in HER2CA astrocytes for genes with log fold change >0 and genes displaying FPKM ≥ 1 in control astrocytes for genes with log fold change <0. For Fig. 5f, displaying log, gene expression changes between YAP/TAZ KO and YAP/TAZ wt gliomaspheres, we considered only genes expressed either in YAP/TAZ KO gliomaspheres for upregulated genes or in YAP/TAZ wt gliomaspheres for downregulated genes, that is, genes displaying FPKM ≥ 1 in YAP/TAZ KO gliomaspheres for genes with log fold change >0 and genes displaying FPKM ≥ 1 in YAP/TAZ wt gliomaspheres for genes with log fold change <0.

**Quantitative real-time PCR.** Real-time PCR was performed as described previously<sup>23</sup>, using System thermal cycler and analyzed with QuantStudio software (Thermo Fisher Scientific; v.1.4.3). Expression levels were normalized to GAPDH. PCR oligonucleotide sequences are listed in Supplementary Table 15.

**Immunofluorescence.** Immunofluorescence on PFA-fixed cells and tissue samples was performed as previously described<sup>24</sup>. Primary and secondary antibodies and their working dilutions are described in Supplementary Table 16. Slides were mounted with Fluoroshield Mounting Medium with DAPI (F6057, Sigma). Images were acquired with Leica TCS SP5II confocal microscope equipped with a CCD camera using LAS AF 2.7.3.9723 software and analyzed using Volocity software 6.0 (PerkinElmer).

**Western blot.** Immunoblots were performed as previously described<sup>25</sup>. Chemiluminescence was digitally acquired by ImageQuant LAS 4000 1.2 (GE healthcare). Primary and secondary antibodies and their working dilutions are described in Supplementary Table 16.

**Statistics and reproducibility.** Data are mean ± s.d. or s.e.m. as indicated in figure legends. Statistical tests (Student’s t-test, ANOVA and Kaplan–Meier survival analyses) are indicated in the figure legends and were performed with GraphPad Prism 8.0.2 software. Sample sizes for each experiment are stated in the corresponding figure legends. No statistical method was used to predetermine sample size. All tested animals were included in the analysis. All experiments were reproducible. Every figure states how many times each experiment was performed with similar results. Mice were randomly allocated to experimental or treatment groups. Investigators were not blinded to mouse grouping. Pathological examination of histological sections was carried out by M. Fassan (a professional pathologist), who was blind to animal treatments.

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

**Data availability**

All RNA-seq and microarray raw data generated for the present study, along with counts matrices and metadata for each sample, are publicly available in GEO under...
accession code GSE133471. The scRNA-seq data of primary glioblastoma samples from Darmanis et al. were downloaded as raw reads from GEO (GSE84465). The expression matrix and metadata of the Nefelt dataset were downloaded from the Single Cell Portal of the Broad Institute (https://singlecell.broadinstitute.org/single_cell/study/SCP393/single-cell-rna-seq-of-adult-and-pediatric-glioblastoma#study-summary). Raw gene expression data (CEL files) of the GBM TCGA cohort were downloaded from GEO (GSE83130). Raw gene expression (CEL files) and clinical data of the REMBRANDT study were downloaded from GEO (GSE10194). BAM files of Ivy Atlas GBM samples were downloaded from the Anatomic Structures RNA-Seq repository of the Ivy Glioblastoma Atlas Project (http://glioblastoma.alleninstitute.org/ncsreseq.bam.csv). Source data for Figs. 1e, 3b–f, 4c,f,g, 5a–f, 6h, and Extended Data Figs. 1c, 2b,c, 3b, 4a–d, 7d, 8b, 9a,b and 10b,ef have been provided with the paper. All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

Code availability

All code used R 3.5.0 and publicly available packages cited in the paper. No custom functions were written for the analysis. STAR is available in GitHub (https://github.com/alexdobin/STAR). ARACNE-AP is available in GitHub (https://github.com/califano-lab/ARACNE-AP). HOMER is available at http://homer.ucsd.edu/homer/. MultiExperiment Viewer is available at http://mev.tm4.org.

Received: 6 August 2019; Accepted: 28 October 2020; Published online: 7 December 2020

References

1. Chen, J., McKay, R. M. & Parada, L. F. Malignant glioma: lessons from brain cancer patients. *Nat. Rev. Cancer*, 6, 83–94 (2006).

2. Singh, S. K. et al. Identification of human brain tumour initiating cells. *Science*, 306, 840–843 (2004).

3. Gally, J. A. et al. Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cell Res.*, 16, 701–7021 (2004).

4. Bao, S. et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature*, 444, 756–760 (2006).

5. Gimple, R. C., Bhargava, S., Dixit, D. & Rich, J. N. Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer. *Genes Dev.*, 33, 591–609 (2019).

6. Pollard, S. M. et al. Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens. *Cell Stem Cell*, 4, 568–580 (2009).

7. Darmanis, S. et al. Single-cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. *Cell Rep.*, 21, 1399–1410 (2017).

8. Nefelt, C. et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell*, 178, 835–849 e821 (2019).

9. Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*, 344, 1396–1401 (2014).

10. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods*, 14, 979–982 (2017).

11. Pollen, A. A. et al. Molecular identity of human outer radial glia during cortical development. *Cell*, 163, 55–67 (2015).

12. Chen, J. et al. A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature*, 488, 522–526 (2012).

13. Lathia, J. D. et al. Integrin-a6 regulates glioblastoma stem cells. *Cell Stem Cell*, 6, 421–432 (2010).

14. Bertolini, I. A. et al. Mapping the global chromatin connectivity network for Sox2 function in neural stem cell maintenance. *Cell Stem Cell*, 24, 462–476 (2019).

15. La Manno, G. et al. RNA velocity of single cells. *Nature*, 560, 494–498 (2018).

16. Miroshnikova, Y. A. et al. Tissue mechanics promote IDH1-dependent HIF1α-dependent CαC feedback to regulate glioblastoma aggression. *Nat. Cell Biol.*, 18, 1336–1345 (2016).

17. Hambardzumyan, D., Gutmann, D. H. & Kettenmann, H. The role of microglia and macrophages in glioma maintenance and progression. *Nat. Neurosci.*, 19, 20–27 (2016).

18. Lathia, J. D., Mack, S. C., Mulkearns-Hubert, E. E., Valentim, C. L. & Rich, J. N. Cancer stem cells in glioblastoma. *Genes Dev.*, 29, 1203–1217 (2015).

19. Minata, M. et al. Phenotypic plasticity of invasive edge glioma stem-like cells in response to ionizing radiation. *Cell Rep.*, 26, 1893–1905 (2019).

20. Laks, D. R. et al. Neurosphere formation is an independent predictor of clinical outcome in malignant glioma. *Stem Cells*, 27, 980–987 (2009).

21. Pallini, R. et al. Cancer stem cell analysis and clinical outcome in patients with glioblastoma multiforme. *Clin. Cancer Res.*, 14, 8205–8212 (2008).

22. Zeppernick, F. et al. Stem cell marker CD133 affects clinical outcome in glioblastoma patients. *Clin. Cancer Res.*, 14, 123–129 (2008).

23. Brennan, C. W. et al. The somatic genomic landscape of glioblastoma. *Cell*, 155, 462–477 (2013).

24. Gusev, Y. et al. The REMBRANDT study; a large collection of genomic data from brain cancer patients. *Sci. Data*, 5, 180158 (2018).
57. Pistollato, F. et al. Molecular mechanisms of HIF-1α modulation induced by oxygen tension and BMP2 in glioblastoma derived cells. *PLoS ONE* **4**, e6206 (2009).

58. Oh, T. et al. Immunocompetent murine models for the study of glioblastoma immunotherapy. *J. Transl. Med.* **12**, 107 (2014).

59. Zanconato, F. et al. Transcriptional addiction in cancer cells is mediated by YAP/TAZ through BRD4. *Nat. Med.* **24**, 1599–1610 (2018).

60. Miller, T. E. et al. Transcription elongation factors represent in vivo cancer dependencies in glioblastoma. *Nature* **547**, 355–359 (2017).

61. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

62. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

63. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

64. Yao, Z. et al. A single-cell roadmap of lineage bifurcation in human ESC models of embryonic brain development. *Cell Stem Cell* **20**, 120–134 (2017).

65. Nowakowski, T. J. et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).

66. Zhong, S. et al. A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018).

67. Kanton, S. et al. Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature* **574**, 418–422 (2019).

68. Iritizy, R. A. et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264 (2003).

69. Corces, M. R. et al. The chromatin accessibility landscape of primary human cancers. *Science* https://doi.org/10.1126/science.aav1898 (2018).

70. Lachmann, A., Giorgi, F. M., Lopez, G. & Califano, A. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. *Bioinformatics* **32**, 2233–2235 (2016).

71. Alvarez, M. J. et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat. Genet.* **48**, 838–847 (2016).

72. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

73. Cordenonsi, M. et al. The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* **147**, 759–772 (2011).

74. Panciera, T. et al. Induction of expandable tissue-specific stem/progenitor cells through transient expression of YAP/TAZ. *Cell Stem Cell* **19**, 725–737 (2016).

75. Zhang, N. et al. The Merlin/NF2 tumor suppressor functions through the YAP oncoprotein to regulate tissue homeostasis in mammals. *Dev. Cell* **19**, 27–38 (2010).

76. Schonhuber, N. et al. A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat. Med.* **20**, 1340–1347 (2014).

77. Azzolin, L. et al. YAP/TAZ incorporation in the β-catenin destruction complex orchestrates the Wnt response. *Cell* **158**, 157–170 (2014).

78. Schilde, S., Bohrer, C., Beck, K. & Schachtrup, C. Isolation and culture of mouse cortical astrocytes. *J. Vis. Exp.* https://doi.org/10.3791/50079 (2013).

79. Pocci, E. et al. BMP9 counteracts the tumorigenic and pro-angiogenic potential of glioblastoma. *Cell Death Differ.* **25**, 1808–1822 (2018).

80. Totaro, A. et al. YAP/TAZ link cell mechanics to Notch signalling to control epidermal stem cell fate. *Nat. Commun.* **8**, 15206 (2017).

**Acknowledgements**

We thank I. Verma, J. Massagué and L. Naldini for plasmids and colleagues sharing their plasmids through Addgene (M.-C. Hung, L. Pedersen, C. Counter, C. Cepko, K. Hochedlinger and M. Meyerson). We thank D.J. Pan, D. Saur, J. Siveke and P. Bonaldo for gifts of mice, G. Basso for HuTu cells, G. Zaccolotto for the GFP/Luc-expressing lentiviral construct, V. Barbieri for in vivo experiments, V. Guzzardo for histology, S. Bresolin for microarrays, G. Leo for TAZ IHC analysis and M. Forcato for comments.

M. Castellan was supported by a FIRC-AIRC fellowship for Italy; O.R. is supported by Fondazione Umberto Veronesi (Post-Doctoral Fellowship 2020). The research leading to these results has received funding from AIRC 5x1000 2018 ‘Metastasis as mechanodisease’ (ID, 22759) grant to S.P.; from AIRC IG Grant 2019 (ID, 23307) to S.P.; from the Italian Ministry of Education, University and Research under a MIUR-FARE grant to S.P. and a MIUR-PRIN Bando 2017 grant to S.P. (cod. 2017HWTP2K); from the European Research Council under the European Union’s Horizon 2020 research and innovation program (DENOVOSTEM grant agreement No 670126) to S.P.

**Author contributions**

M. Castellan performed most of the in vitro and in vivo experiments and contributed to writing. A.F. carried out the initial experiments of this study. A. Guarnieri and G.B. carried out experiments with HuTu cells. T.P. optimized mouse astrocyte isolation and infection. F.Z. contributed to manuscript preparation. E.Z., H.L.S., P.C. and A.C. optimized technical procedures critical for experiments in vivo and in vitro; M.F. performed histology and histopathological evaluations. E.P. and A.R. performed brain tumor experiments. O.R., A. Grilli and S.B. performed bioinformatic analyses. S.P. and M. Cordenonsi conceived the initial hypothesis and experimental design, organized the work and wrote the manuscript.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Extended data is available for this paper at https://doi.org/10.1038/s43018-020-00150-z.

Supplementary information is available for this paper at https://doi.org/10.1038/s43018-020-00150-z.

Correspondence and requests for materials should be addressed to M.C. or S.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
Extended Data Fig. 1 | Identification of the gene expression program of GSCs. **a.** Single-cell differentiation trajectory of GBM cells reconstructed by Monocle2 using single-cell RNA-seq data of the indicated cell populations from primary GBM samples of the Darmanis dataset. **b.** Gene set enrichment analysis (GSEA) for association between the cell populations at the start and at the end of the pseudotime trajectory of the neoplastic cells of the Darmanis datasets (as depicted in Fig. 1c), and gene sets denoting the identity of specific cell types. Gene lists denoting early neural progenitor cells (RG: Radial Glia; oRG: outer Radial Glia; vRG: ventricular Radial Glia) or neural stem cells (NSC) are indicated in red; those identifying neurons, astrocytes or committed neuronal progenitors (OPC: Oligodendrocyte Progenitor Cells; INP: Intermediate Neuronal Progenitors) are, respectively, in blue, purple and blue-green colors; gene lists enriched in the putative GSC and DGC populations are highlighted in orange and in light blue, respectively. Signatures are available in Supplementary Table 1. GSEA calculated FDR adjusting for multiple comparisons; details of p-value and FDR calculation are described in the GSEA website (http://software.broadinstitute.org/gsea/index.jsp). Related to Fig. 1c. **c.** Log2 expression levels of the indicated oRG (top graphs), NSC and GSC (middle graphs) and INP markers (bottom graphs) in the subpopulations of neoplastic cells of the Darmanis dataset that are at the start (GSC, n = 221 cells) and at the end (DGC, n = 221 cells) of the pseudotime trajectory depicted in Fig. 1c. Data are presented as mean ± s.d. p-values were determined by unpaired two-tailed t test. **d.** RNA velocities (arrows) of neoplastic cells of the Darmanis dataset projected in the space of the first two principal components. Red and blue dots are the cells that are at the start (GSC) and at the end (DGC) of the pseudotime trajectory depicted in Fig. 1c.
Extended Data Fig. 2 | Validation of the G-STEM signature.  

a-b, Violin plots showing the expression of the G-STEM signature (right panels in (b)) on the cells at the start (Low; red dots in the left panels in (b)) of the pseudotime trajectories (a) of patient-specific cohorts of the Darmanis dataset, vs. the neoplastic cells that are on the opposite ends of the same trajectories (High; blue dots in the left panels in (b)). The p-values were determined by two-tailed Mann-Whitney test.  

C, Violin plots showing the expression of the G-STEM signature (right panel) on the cells at the start (Low; red dots in the middle panel) of the pseudotime trajectory (left panel) of the sole neoplastic cells of the Darmanis dataset, vs. the cells that are on the opposite ends of the same trajectory (High; blue dots in the middle panel). The p-values were determined by two-tailed Mann-Whitney test.
Extended Data Fig. 3 | Characterization of the G-STEM signature. a, Graphs depicting the most significant GO terms emerging from the Gene Ontology analyses of the genes composing the G-STEM and the DGC signatures. The full lists of significant GO terms of both signatures are in Supplementary Table 3. b, Log2 expression levels of the indicated components of the G-STEM signature in the subpopulations of neoplastic cells of the Darmanis dataset that are at the start (GSC, n = 221 cells) and at the end (DGC, n = 221 cells) of the pseudotime trajectory depicted in Fig. 1c. Data are presented as mean ± s.d. p-values were determined by unpaired two-tailed t test.
Extended Data Fig. 4 | Validation of the G-STEM signature in large datasets of GBM patients. **a**, Gene set enrichment analysis (GSEA) for association between the cell population at the start of the pseudotime trajectory of the neoplastic cells of the Neftel datasets (as depicted in Fig. 1e) vs. all the other neoplastic cells and gene sets denoting the identity of specific cell types. Abbreviations and color codes are as in Extended Data Fig. 1b. Signatures are available in Supplementary Table 1. GSEA calculated FDR adjusting for multiple comparisons; details of p-value and FDR calculation are described in the GSEA website (http://software.broadinstitute.org/gsea/index.jsp). Related to Fig. 1e. **b**, Violin plots showing the expression of the G-STEM signature (bottom panels) on the cells at the start of the pseudotime trajectory (GSC; red dots in the top panels) of small tumor cohorts of the Neftel dataset, pre-sorted according to the Proneural, Classical or Mesenchymal classification of GbMs, vs. all the other neoplastic cells of the same cohorts (NON GSC; light blue dots in the top panels). The p-values were determined by two-tailed Mann-Whitney test. **c**, Kaplan–Meier analysis representing the probability of survival in n=541 GBM patients from the TCGA dataset (left panel), n=210 GBM patients from the REMBRANDT dataset (middle panel), and n=390 GBM patients carrying wild-type IDH1 from the TCGA dataset (right panel), stratified according to high or low GSC-signature. The p-value of the Log-rank (Mantel-Cox) test reflects the significance of the association between GSC-signature ‘low’ and longer survival. G-STEM expression is prognostic for the vast majority of GBM, that is IDH1-wild type tumors (93%, of those annotated in the TGCa dataset; n=390 out of 419 IDH1-annotated samples).
Extended Data Fig. 5 | A computational procedure to identify candidate TRs controlling the gene expression program of GSCs. a, Overview of the experimental flow for inference of the master Transcriptional Regulators (TRs) of the GSC state using the Rhabdomant pipeline on the Darmanis sc-RNA-seq dataset of primary GBM samples. See Methods for details. b, List of candidate master Transcriptional Regulators (TRs) emerging from the analysis of the Darmanis dataset of scRNA-seq dataset with the Rhabdomant pipeline, ordered on the base of their normalized enrichment signal (NES). The Rhabdomant pipeline calculated FDR adjusting for multiple comparisons; see Methods for details about p-value and FDR calculation. The lists of candidate master TRs of the GSC and of the DGC state are highlighted in orange and in light blue, respectively. The most significant candidate master TRs of the GSC state are indicated in red.
Extended Data Fig. 6 | YAP/TAZ are required for GSC maintenance in vivo. a–c, Effects of YAP/TAZ knockout on the growth of established subcutaneous GBM-like lesions. Transformed cells were obtained by dissociation of gliomaspheres obtained from HER2CA- (a), shNF1/shp53- (b) or KRasG12V/shp53- (c) transformed R26CAGCreERT2; Yapfl/fl; Tazfl/fl newborn mouse astroglial cells (as in Fig. 3), and then injected in NOD-SCID mice. When subcutaneous tumors reached approximately 0.5 cm of diameter, mice were either fed with Tamoxifen food to induce YAP/TAZ knockout (YAP/TAZ KO), or maintained under normal diet (YAP/TAZ wt). Graphs are growth curves of YAP/TAZ wt (KRasG12V/shp53-, n = 4 mice; HER2CA, n = 6 mice; shNF1/shp53, n = 5 mice) and YAP/TAZ KO (KRasG12V/shp53-, n = 4 mice; HER2CA, n = 4; shNF1/shp53, n = 8 mice) tumors (average volume ± s.e.m.). d–f, Effects of YAP/TAZ knockout in tumors derived from KRasG12V/shp53 gliomaspheres, following the experimental setup described in a–c. d, Dot plot for tumor weight at sacrifice (YAP/TAZ wt, n = 8; YAP/TAZ KO, n = 6). Mean ± s.e.m. of the distribution are also shown. p-value was calculated by unpaired two-tailed t-test. e, Representative H&E stainings. Scale bar, 2.5 mm. N, necrotic area; *, Matrigel residue. f, Tabular results showing the number of NOD/SCID mice displaying subcutaneous tumor formation after injection of cells dissociated either from gliomaspheres derived from HER2CA-transformed primary newborn astroglial cells (Primary tumors), or from HER2CA-gliomaspheres derived from one of the Primary tumors (Secondary tumors).
Extended Data Fig. 7 | Ex-vivo reprogramming of normal neural cells into GSC-like cells. a, GFAP and SOX2 stainings (scale bars, 50 μm) of the mouse SVZ, representative of n = 3 mice. Nuclei were counterstained with DAPI. b, c, GFAP, NESTIN and SOX2 stainings (scale bars, 50 μm) in mouse newborn astroglial cells, representative of two independent experiments. d, Gliomaspheres emerging from newborn astroglial cell cultures transformed by the indicated oncogenes (P0 spheres) were dissociated to single cells and replated at clonal density for gliomasphere formation (P1 to P10 spheres). Results are representative of three experiments with n = 3 replicates each. Data are presented as scatter dot plots and bar graphs showing mean with s.d. e, Left panel: H&E staining of a lesion obtained after intracranial transplantation of shNf1/shp53-transformed astroglial cells. N, necrotic area. Scale bar, 2.5 mm. Middle panel: High magnification of the same tumor, showing large polynucleated cells (arrowheads). Right panel: TAZ IHC on the same tumor. Scale bars, 100 μm. Experiments were independently repeated on n = 10 mice, with similar results. f, H&E staining of subcutaneous tumors obtained by injecting cells dissociated from gliomaspheres carrying the indicated oncogenic lesions, representative of: KRasG12V/shp53, n = 4 tumors; HER2Ca, n = 6 tumors; shNf1/shp53, n = 5 tumors. N, necrotic areas. Scale bars, 250 μm. g, Number of mice displaying tumor formation after injection of cells dissociated from KRasG12V/shp53-gliomaspheres at the indicated cell dilutions. h, Schematic representation of the serial transplantation assay performed with HER2Ca-transformed cells (see Methods for details). Bottom, H&E staining (scale bars, 2.5 mm) of tumors obtained after each round of transplantation, representative of n = 4 primary tumors, n = 8 secondary tumors and n = 4 tertiary tumors, respectively. Numbers of mice developing tumors per numbers of transplanted mice are indicated in each picture. i, GSEA curves of the G-STEM and the DGC signatures in KRasG12V/shp53-tumors compared to the astroglial cells from which they derive. Signatures are available in Supplementary Table 7.
Extended Data Fig. 8 | Oncogenic insults activate YAP/TAZ in transformed primary astroglial cells. 

**a**, Bright-field and fluorescent pictures (representative of n = 5 independent samples each) of newborn astroglial cells transduced with lentiviral vectors encoding for the YAP/TAZ reporter 8xGTIIIC-RFP-6DD\(^6\), and with lentiviral vectors encoding for the indicated oncogenes or, as negative control, with empty vector, as in Fig. 3b. Images were taken 4 days after inducing oncogenic reprogramming by incubating cells in NSC medium. Scale bars, 50 μm. 

**b**, Compendium of Fig. 3c. Efficiency of yap\(^−\)/Taz\(^−\) downregulation in R26CAG-CreERT2; yap\(^−\)/Taz\(^−\) mouse newborn astroglial cells treated with either vehicle (Control) or 4OH-TAM (YAP/TAZ KO), as measured by qRT-PCR (mean ± s.d. of all independent samples of three experiments). p-values are calculated by two-way ANOVA with Sidak’s multiple comparisons.
Extended Data Fig. 9 | YAP/TAZ are required for GSC maintenance in vitro. a, Control experiment of Fig. 5a–e. Gliomaspheres derived from HER2CA-transformed Yap^{fl/fl}; Taz^{fl/fl} newborn astroglial cells, not expressing CREERT2, were treated with either ethanol (Vehicle) or 4OH-TaM (TAM). Panels are representative images (left; scale bar, 100 µm) and quantifications (right; mean ± s.d. of two independent experiments, each performed with two replicates) of the number of gliomaspheres/cm² in vehicle versus 4OH-TaM-treated samples. p-values were determined by two-way ANOVA with Sidak’s multiple comparisons test. In the absence of CREERT2 expression, treatment with 4OH-TaM is inconsequential for gliomasphere formation, indicating that gliomasphere disaggregation shown in Fig. 4a–e is specifically caused by yap/Taz deletion. b, P2 gliomaspheres derived from R26CAG-CreERT2; Yap^{fl/fl}; Taz^{fl/fl} newborn astroglial cells transformed with the indicated oncogenes were dissociated to single cells and replated at clonal density for P3 gliomasphere formation in presence of ethanol (YAP/TAZ wt), or of 4OH-TAM to induce YAP/TAZ knockout (YAP/TAZ KO). Data are presented as scatter dot plots (n = 3 replicates each) and bar graphs showing mean with s.d. The p-values were calculated by unpaired two-tailed t-test.
Extended Data Fig. 10 | See next page for caption.
Extended Data Fig. 10 | YAP/TAZ are required for GBM initiation in vivo. a–c, Immunocompromised mice were injected intracranially with KRasG12V/shp53-transformed Yapfl/fl;Tazfl/fl cells, also transduced with dual luciferase-GFP expression vectors. Control animals (n = 6) were injected with cells transduced with Ad-GFP, whereas YAP/TAZ KO animals (n = 5) were injected with cells transduced with Ad-Cre. a, Representative images of brain bioluminescence. b, Bioluminescence quantification shown as scatter dot plots and bar graphs showing mean with s.d.; p-value was calculated by unpaired two-tailed t-test. c, Representative H&E stainings. Scale bars, 2.5 mm in left panels and 250 μm in the magnification shown on the right. Arrowheads highlight the presence of large, polynucleated cells. d–f, Immunocompromised mice were injected intracranially with HuTu13 cells transduced with dual luciferase-GFP expression vectors, and transfected with siCo (Control; n = 5) or siYAP/TAZ (YAP/TAZ depleted; n = 5). d, Representative images of brain bioluminescence. e, Bioluminescence quantification shown as scatter dot plots and bar graphs showing mean with s.d.; unpaired two-tailed t-test p-values are shown. f, Representative H&E stainings. Scale bars, 2.5 mm in left panels and 250 μm in the magnification shown on the right. ‘N’ indicates necrosis. g–i, CT2a cells were transduced with dual luciferase-GFP expression vectors and injected intracranially in syngeneic mice. Control animals (n = 5) were injected with cells expressing anti-GFP shRNA, whereas YAP/TAZ-depleted animals (n = 5) were injected with cells expressing doxycycline-inducible YAP and TAZ shRNAs. g, Representative brain bioluminescences at one day and 14 days after injection. h, Bioluminescence quantification at three different time points shown as scatter dot plots and bar graphs showing mean with s.d.; unpaired two-tailed t-test p-values are shown. i, Representative H&E stainings. Scale bars, 2.5 mm in left panels and 250 μm in the magnification shown on the right. ‘N’ indicates necrosis. j, GFP and TUJ1 stainings in sections from YAP/TAZ-wt and YAP/TAZ-KO subcutaneous shNF1/shp53-induced tumors (representative of n = 3 independent samples each). Scale bars, 50 μm.
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 Authors & Referees 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

LAS AF 2.7.3.9723 software for Leica TCS SPS II was used to acquire confocal images. Bright-field images were acquired with a Leica DMIL LED microscope equipped with a Leica DFC 3000G camera using LAS version X software. NPPScan1.3 was used to acquire H&E images. ImageQuant LAS 3.0D 1.2 was used to acquire western blot images. QuantStudio Design & Analysis Software v1.4.3 was used to acquire qPCR data. HuTu cells expression matrices were obtained with Affymetrix Microarray profiling on HG-U133 Plus 2.0 arrays (Affymetrix). For in vivo luciferase measurements data analysis was performed using Living Image software version 4.7.2 [PerkinElmer]. RNA-seq libraries for gene expression profiling of mouse astroglial cells, gliomaspheres, and tumors were prepared with TruSeq Stranded mRNA Library Prep (Illumina), according to manufacturer’s instructions and sequencing was performed with Illumina HiSeq 2500. Detailed informations on data collection can be found in the Method section of the paper. TAZ IHC images were acquired using the D-Sight-F system for digital pathology (Menarini Diagnostics).

Data analysis

Confocal images were analyzed with Volocity software [PerkinElmer, version 6.0]. GraphPad Prism 8.0.2 for Mac was used for statistical analysis. Velocity software [PerkinElmer, version 6.0] was used to analyze IF images.

All the analyses of single-cell RNA-seq data, processing of GBM transcriptomics from the TCGA and REMBRANDT studies and gene expression profiling of HuTu cell lines have been performed in R 3.5.0. Pseudotime trajectories on neoplastic and normal neural cells were constructed using Monocle2 version 2.8.0. RNA velocity was estimated using velocyto.R [github.com/velocyto-team/velocyto.R]. Differential gene expression analysis between GSC and DGC populations have been performed using the Seurat package version 3.2.1. Results from the analysis of differential expression were functionally annotated using Gene Set Enrichment Analysis (GSEA, http://software.broadinstitute.org/gsea/index.jsp) and gene sets derived from previously published gene signatures. Gene signatures were annotated using the Enrichr website (https://maayanlab.cloud/Enrichr/). Glioblastoma regulatory networks was reconstructed using ARACNe-AP on the transcriptomes of the Darmanin scRNA-seq dataset. To identify master regulators of GSCs we applied the VIPER algorithm in R 3.5.0. Gene expression of transcriptomes from the Iwh Glioblastoma Atlas was quantified in R 3.3.1. STAR, featureCounts, edgeR package and R (version 3.3.1) were used to analyze RNA-seq data. Hierarchical clustering of Fig. 43d was performed using the Hierarchical Clustering of the MultiExperiment Viewer (MeV 4.8; http://mev.tm4.org) package with Pearson correlation; gene expression heatmaps have been generated in GraphPad Prism 8.0.2. Detailed informations on data analysis can be found in the Method section of
the paper. TAZ IHC images were analyzed using the D-Sight Viewer software [Menarini Diagnostics].

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/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

All RNA-seq and microarray raw data generated for the present study, along with counts matrices and metadata for each sample, are publicly available in Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession GSE133471. Single-cell RNA-seq data of primary glioblastoma samples from Darmanis et al. were downloaded as raw reads from GEO GSE84465. Expression matrix and metadata of the Netfell dataset were downloaded from the Single Cell Portal of the Broad Institute (https://singlecell.broadinstitute.org/single_cell/study/SCP393/single-cell-rna-seq-of-adult-and-pediatric-glioblastoma#study-summary). Raw gene expression data (.CEL files) of the GBM TCGA cohort were downloaded from GEO GSE83130. Raw gene expression (.CEL files) and clinical data of the REMBRANDT study were downloaded from GEO GSE108474. BAM files of Ivy Atlas GBM samples were downloaded from the Anatomic Structures RNA-Seq repository of the Ivy Glioblastoma Atlas Project (http://glioblastoma.alleninstitute.org/rnaSeq/bam.csv).

Source data for Fig. 1e, 3b, 3e-f, 4c, 4f-g, 5a-f, 6b, 6f, Extended Data Fig. 1c, 2b-c, 3b, 4b, 4a-d, 7d, 8b, 9a-b, 10b, 10e, 10f have been provided with the paper.

All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

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.
- [x] 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-list.pdf

Life sciences study design

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

| Sample size | Sample sizes for each experiment are stated in the corresponding figure legends. These numbers are based on previous studies (Ref. 38,43,59,73,79). In general, for each experiment at least 3 or more replicates/number of animals was used to account for variability in biological systems. No statistical method was used to predetermine sample size. |
| Data exclusions | All data were included for all the biological experiments and most the computational analyses. The only exception was about pseudotime analyses of single patients from the Darmanis scRNA-seq dataset, for which we pre-established to include only patients with sufficient sampling of cells (>200), as required for reconstruction of differentiation trajectories from single-cell RNA-seq data, as discussed by Kester and van Oudenaarden, Cell Stem Cell 2018. We thus excluded patient BT-56 that was characterized by a very limited number of sequenced cells. |
| Replication | All experiments were reproducible. Every figure states how many times each experiment was performed with similar results. |
| Randomization | Mice were randomly allocated to experimental or treatment groups. |
| Blinding | Investigators were not blinded to mouse grouping since animal experiments and tumor assessment were performed by the same researchers. Instead, pathological examination of histological section was carried out by M. Fassan [a professional pathologist] who was blind to animal treatments. |

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

- X Antibodies
- X Eukaryotic cell lines
- X Palaeontology
- X Animals and other organisms
- X Human research participants
- X Clinical data

### Methods

| n/a | Involved in the study |
|-----|-----------------------|
|     | - ChiP-seq            |
|     | - Flow cytometry      |
|     | - MRI-based neuroimaging |

### Antibodies

**Antibodies used**
- Primary antibodies:
  - anti-GFAP: Z0334, Dako (1:1000 for IF, 1:2000 for WB)
  - anti-SOX2: L1D6A2, Cell Signaling Technology (1:50 for IF)
  - anti-mNESTIN: MAB353, Millipore (1:300 for IF)
  - anti-hNESTIN: MAB5326, Millipore (1:250 for IF)
  - anti-Tuj1: Tuj133, Boehringer (1:300 for IF)
  - anti-GFP: 13970, Abcam (1:200 IF)
  - anti-YAP/TAZ: sc-101199, Santa Cruz Biotechnology (1:200 for IF, 1:1000 for WB)
  - anti GAPDH: MAB374, Millipore (1:50000 for WB)
  - anti-TAZ: HPA007415, Sigma (1:50 for IHC)

**Secondary antibodies:**
- Goat anti mouse: A28723, Invitrogen (1:200, for IF)
- Goat anti rabbit: A32731, Invitrogen (1:200, for IF)
- Goat anti Chicken: A28723, Invitrogen (1:200, for IF)
- Goat anti mouse: A12004, Invitrogen (1:200, for IF)
- Goat anti rabbit: A12004, Invitrogen (1:200, for IF)
- Anti Rabbit-HRP: A0545, Sigma, (1:1000 for WB)
- Anti Mouse-HRP: PN4201V, Sigma, (1:5000 for WB)

**Validation**
Pre-validated antibodies were purchased from reputable commercial sources. All antibodies were tested in the relevant application to detect recombinant protein without cross-reactivity with non-specific proteins.

### Eukaryotic cell lines

**Policy information about cell lines**

**Cell line source(s)**
HEK293T were from ATCC. HuTu10 and HuTu13 cells were a gift from G. Basso, University of Padova. CT2A and Gl261 mouse glioma cell lines were purchased from Millipore and from DSMZ, respectively.

**Authentication**
HEK293T were authenticated by DSMZ service. Genetic characteristics were determined by PCR-single-locus-technology. 21 independent PCR-systems Amelogenin, D3S1358, D1S1636, D6S1043, D13S317, Penta E, D165599, D18S51, D21S1338, CSF1PO, Penta D, TH01, vWA, D21S11, D7S820, D5S818, TPOX, D8S1179, D12S391, D19S433 and FGA were investigated (Promega, PowerPlex 21 PCR Kit).

**Mycoplasma contamination**
All cell lines were routinely checked for mycoplasma contamination and tested negative.

**Commonly misidentified lines**
None of the cell lines used in this study is present in the database of commonly misidentified cell lines.

### Animals and other organisms

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

**Laboratory animals**
Experimental mice [B6C3FCrIR12; Yapl/F; 1a2f/f] used to obtain astrocytes are described in methods section; C57BL/6j mice were from Charles River. Both male and female mice were used. C57Bl/6, NOD-SCID and NSG females used for in vivo experiments were from Charles River; they were 6-8 weeks old at the beginning of experiments.

**Wild animals**
No wild animals were used in this study. Field-collected

**Field-collected samples**
The study did not involve samples collected on the field.

**Ethics oversight**
Animal experiments were performed adhering to our institutional and national guidelines as approved by OPBA (Padova) and the Ministry of Health of Italy.

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