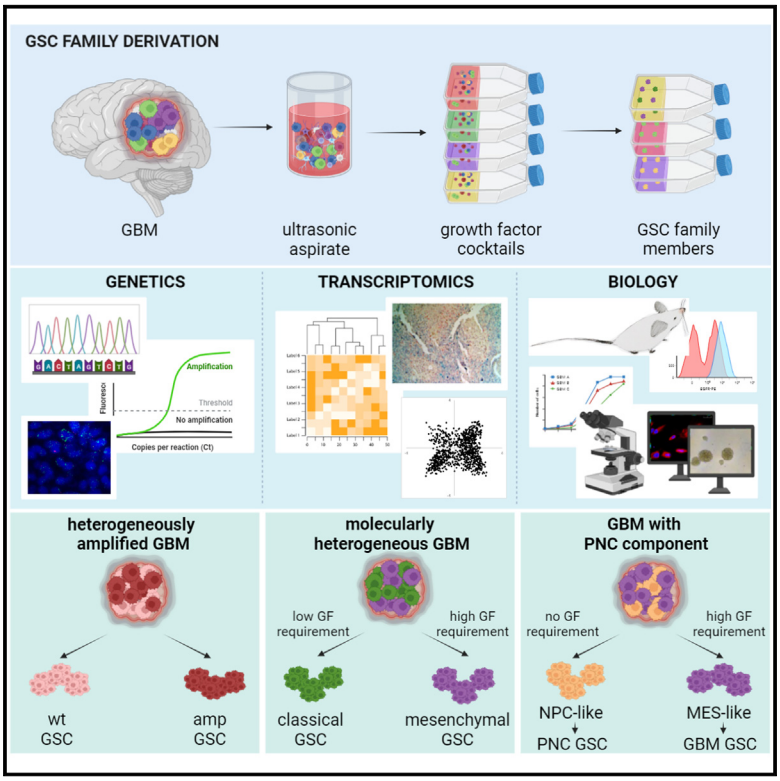
Cell Reports

Resource

Coexisting cancer stem cells with heterogeneous gene amplifications, transcriptional profiles, and malignancy are isolated from single glioblastomas

Graphical abstract



Highlights

- Distinct GSCs can be derived from single GBMs by different growth factor cocktails
- GSCs segregate different RTK expression and gene amplifications
- GSCs segregate distinct transcriptional profiles, each related to different malignancy
- GSCs with a neural progenitor profile recapitulate the GBM primitive neuronal component

Authors

Francesca De Bacco, Francesca Orzan, Giovanni Crisafulli, ..., Gaetano Finocchiaro, Pietro L. Poliani, Carla Boccaccio

Correspondence

carla.boccaccio@ircc.it

In brief

De Bacco et al. devise a methodology to propagate distinct glioblastoma stem-like cells (GSCs) from individual tumors, stably segregating different features coexisting in the original tumor. The GSCs' transcriptional profile is a superior indicator of their tumorigenicity compared with their genetic make-up, uncovering diverse phenotypes that drive intratumoral heterogeneity.

De Bacco et al., 2023, Cell Reports *42*, 112816
August 29, 2023
https://doi.org/10.1016/j.celrep.2023.112816

CellPress

Cell Reports

CellPress
OPEN ACCESS

Resource

Coexisting cancer stem cells with heterogeneous gene amplifications, transcriptional profiles, and malignancy are isolated from single glioblastomas

Francesca De Bacco,[1,2,13] Francesca Orzan,[1,13] Giovanni Crisafulli,[3] Marta Prelli,[1,2] Claudio Isella,[2,4] Elena Casanova,[1] Raffaella Albano,[5] Gigliola Reato,[1,2] Jessica Erriquez,[5] Antonio D'Ambrosio,[1] Mara Panero,[6] Carmine Dall'Aglio,[6] Laura Casorzo,[6] Manuela Cominelli,[7] Francesca Pagani,[7] Antonio Melcarne,[8] Pietro Zeppa,[8,9] Roberto Altieri,[9] Isabella Morra,[10] Paola Cassoni,[10] Diego Garbossa,[8,9] Anna Cassisa,[4] Alice Bartolini,[5] Serena Pellegatta,[11] Paolo M. Comoglio,[3] Gaetano Finocchiaro,[12] Pietro L. Poliani,[7] and Carla Boccaccio[1,2,14,*]

[1]Laboratory of Cancer Stem Cell Research, Candiolo Cancer Institute, FPO-IRCCS, 10060 Candiolo, Italy
[2]Department of Oncology, University of Turin, 10060 Candiolo, Italy
[3]IFOM ETS - The AIRC Institute of Molecular Oncology, 20139 Milan, Italy
[4]Laboratory of Oncogenomics, Candiolo Cancer Institute, FPO-IRCCS, 10060 Candiolo, Italy
[5]Core Facilities, Candiolo Cancer Institute, FPO-IRCCS, 10060 Candiolo, Italy
[6]Unit of Pathology, Candiolo Cancer Institute, FPO-IRCCS, 10060 Candiolo, Italy
[7]Pathology Unit, Department of Molecular and Translational Medicine, University of Brescia, 25123 Brescia, Italy
[8]Neurosurgery Unit, Città della Salute e della Scienza University Hospital, 10126 Turin, Italy
[9]Department of Neurosciences, University of Turin, 10126 Turin, Italy
[10]Department of Medical Sciences, University of Turin, 10126 Turin, Italy
[11]Unit of Immunotherapy of Brain Tumors, Fondazione IRCCS Istituto Neurologico C. Besta, 20133 Milan, Italy
[12]Department of Neurology, IRCCS Ospedale San Raffaele, 20132 Milan, Italy
[13]These authors contributed equally
[14]Lead contact
*Correspondence: carla.boccaccio@ircc.it
https://doi.org/10.1016/j.celrep.2023.112816

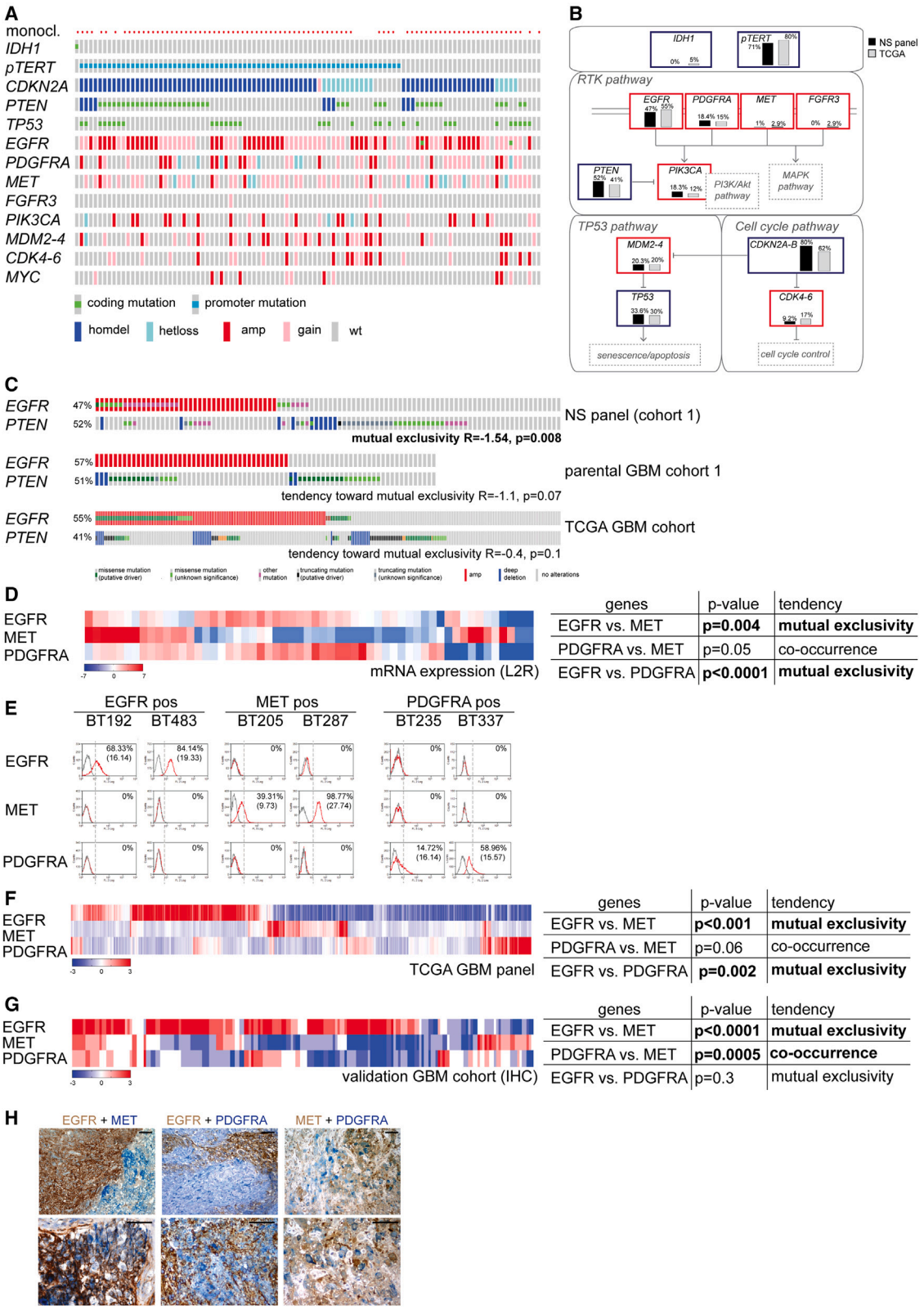
SUMMARY

Glioblastoma (GBM) is known as an intractable, highly heterogeneous tumor encompassing multiple subclones, each supported by a distinct glioblastoma stem cell (GSC). The contribution of GSC genetic and transcriptional heterogeneity to tumor subclonal properties is debated. In this study, we describe the systematic derivation, propagation, and characterization of multiple distinct GSCs from single, treatment-naive GBMs (GSC families). The tumorigenic potential of each GSC better correlates with its transcriptional profile than its genetic make-up, with classical GSCs being inherently more aggressive and mesenchymal more dependent on exogenous growth factors across multiple GBMs. These GSCs can segregate and recapitulate different histopathological aspects of the same GBM, as shown in a paradigmatic tumor with two histopathologically distinct components, including a conventional GBM and a more aggressive primitive neuronal component. This study provides a resource for investigating how GSCs with distinct genetic and/or phenotypic features contribute to individual GBM heterogeneity and malignant escalation.

INTRODUCTION

Glioblastoma (GBM) is one of the deadliest tumors, with a median overall survival <15 months and a 5-year survival rate <5%.[1] GBMs exhibit extensive intratumor heterogeneity, including areas and individual cells that display diverse and fluctuating transcriptomic profiles, as well as different genetic alterations, in particular mutually exclusive amplification of receptor tyrosine kinases.[2–9] GBM heterogeneity hampers therapeutic efficacy and lays the ground for tumor recurrence, as chemo-radiotherapy can promote the selection of resistant subclones and tumor genetic evolution.[3,10–14] However, the mechanisms that generate overall GBM heterogeneity are complex and only partly understood.

GBM is known to contain a cell hierarchy headed by a subset of stem-like cells (GSCs).[15,16] Until now, it has been underappreciated that the GBM heterogeneous subclones should originate from GSCs, which evolve through accumulation of genetic lesions responsible for the phenotypic hallmarks upon which the selective pressures from the environment, including therapies, act, ultimately promoting cancer genetic evolution.[17] Furthermore, GSCs are also a recognized source of phenotypic heterogeneity, as they generate a subset of pseudodifferentiating cells, which lose tumorigenic ability along with stem properties.[15,16] Evidence indicates that a likely GSC cell of origin is the neural stem cell of the brain subventricular zone,[18,19] a progenitor capable of differentiation into the multiple neural tissue cell subtypes.



genes	p-value	tendency
EGFR vs. MET	**p=0.004**	**mutual exclusivity**
PDGFRA vs. MET	p=0.05	co-occurrence
EGFR vs. PDGFRA	**p<0.0001**	**mutual exclusivity**

genes	p-value	tendency
EGFR vs. MET	**p<0.001**	**mutual exclusivity**
PDGFRA vs. MET	p=0.06	co-occurrence
EGFR vs. PDGFRA	**p=0.002**	**mutual exclusivity**

genes	p-value	tendency
EGFR vs. MET	**p<0.0001**	**mutual exclusivity**
PDGFRA vs. MET	**p=0.0005**	**co-occurrence**
EGFR vs. PDGFRA	p=0.3	mutual exclusivity

(legend on next page)

Alternatively, the cell of origin can be a more differentiated cell that regained stem features and differentiation multipotentiality after reprogramming by oncogenes[20] or by neurodevelopmental transcription factors.[21] These origins establish the conditions for the expression of diverse transcriptional programs during the generation of the overall tumor cell population, as reflected in the historical definition of glioblastoma as "multiforme."

Besides theoretical reasoning, data indicate that isolation of GSCs in culture as "neurospheres" (NSs) appears to be the most reliable methodology for obtaining long-term self-propagating cell clones that faithfully retain the genetic make-up of tumor (sub)clones, providing obvious advantages over the highly genetically drifted conventional cell lines.[22] Thus, GSCs are invaluable for functionally challenging the assumptions generated by single-cell analysis.[5] So far, however, NSs have typically been obtained from surgical samples of limited size, hardly representative of the entire variegated landscape of the tumor.[23]

In this study, we developed a methodology to isolate and propagate multiple, distinct GSCs from individual tumors, by applying different growth factor (GF) cocktails to whole GBM cell populations. We thus obtained and characterized a model to explore the GSC contribution to GBM genetic and phenotypic heterogeneity and malignant progression.

RESULTS

Conventional NSs segregate different genetic alterations and expression of receptor tyrosine kinases coexisting in GBM tissues

GBM genetic alterations were reported to be heterogeneously distributed within each tumor, in particular receptor tyrosine kinase (RTK) amplifications,[7] and *TP53* and *NF1* mutations.[2,11] GSCs propagated *in vitro* as NSs are known to retain the original genetic alterations,[22] but it is unknown whether each NS can recapitulate the tumor's genetic heterogeneity. By analyzing a retrospective panel of 98 NSs (Table S1), each obtained from a different GBM (IDH-wt) by conventional methodologies, i.e., by selection in standard medium containing epidermal growth factor (EGF) and fibroblast growth factor 2 (FGF2), and mostly from small tumor fragments, we found that each NS is likely monoclonal, at least as far as driver gene alterations are concerned. NS monoclonality is inferred by evaluation of tumor suppressor gene alterations, such as (1) mutations in *PTEN* and *TP53* (occurring in 65% of NSs), which are almost invariably hemizygous, and (2) *CDKN2A* deletion (occurring in 80% of NSs), which is always homozygous (Figures 1A, S1A, and S1B; Table S1). The overall analysis of GBM driver genes showed that these NSs faithfully retain the alterations of matched original tumors (parental GBMs, cohort 1), which occurred with frequencies similar to those of The Cancer Genome Atlas (TCGA) GBM cohort (Figure 1B; Table S1).[24] We found that NSs segregate genetic alterations such as EGF receptor (*EGFR*) amplification and *PTEN* biallelic inactivation (Figure 1C; Table S1). Such alterations tend to be mutually exclusive also in GBM tumor tissues, both in the cohort that originated this NS panel (cohort 1) and in TCGA GBM cohort, without reaching statistical significance (Figure 1C). This indicates that *PTEN* and *EGFR* alterations tend to occur alternately in different tumors and, if they coexist in subclones of the same tumor, they are alternatively isolated in NSs in a stochastic manner.

Interestingly, we observed that also non-genetic features of parental GBMs tend to segregate in NSs, in particular expression of EGFR, MET (also known as hepatocyte growth factor receptor [HGFR]) and platelet-derived growth factor receptor A (PDGFRA), which were mutually exclusive in NSs (Figures 1D, 1E, and S1C). In TCGA GBM cohort, *EGFR* mRNA expression is mutually exclusive either with *MET* or *PDGFRA* (Figure 1F). Given mutual exclusivity between *EGFR* amplification and *PTEN* biallelic inactivation, *MET* or *PDGFRA* expression tends to associate with the latter in both the NS panel and TCGA GBM cohort (Figures S1D and S1E). Immunohistochemical analysis of an ample independent cohort of GBM tissues (validation cohort, n = 70)[25] confirmed mutual exclusivity between EGFR and the other RTKs (Figure 1G). Interestingly, in 25% of cases, the three RTKs were coexpressed but localized in different cells, either intermingled or grouped in different tumor areas (Figure 1H). Altogether, these data indicate that different tumor subclones, characterized by mutual exclusivity of genetic alterations (e.g., *EGFR* amplification and *PTEN* loss) and/or RTK expression, often coexist in the same GBM. Using standard methodologies, these subclones are typically isolated individually as NSs, meaning that, in each NS, only a single tumor component is represented (Figure S1F).

GSC families can be isolated from GBM whole-cell resuspensions via different GF cocktails

To isolate GSCs representing the subclones coexisting in the same GBM, we designed a protocol combining the use of a comprehensive source of the whole tumor mass with selective pressures able to differentially act on subclones and support generation of established NSs (Figure 2A). As a comprehensive GSC source, we exploited "surgical ultrasonic aspirates"

Figure 1. Conventional NSs segregate different genetic alterations and RTK expression coexisting in GBM tissues

(A) Genetic alterations detected in cohort 1 NS panel (n = 98). Red dots: NS monoclonality inferred by analysis of *CDKN2A*, *PTEN,* and *TP53* genetic alteration.

(B) Gene alteration frequencies in the NS panel (black bars), and in TCGA GBM cohort (gray bars), according to cBioportal and, for p*TERT* mutations, to Brennan et al.[24]

(C) Oncoprint representing mutual exclusivity of *EGFR* amplification and *PTEN* biallelic inactivation in cohort 1 NS panel, and tendency toward mutual exclusivity in parental GBM cohort 1 (n = 74) and in TCGA GBM cohort (n = 273).

(D) Left: RTK mRNA expression in cohort 1 NS panel (qPCR). Right: correlation analysis of RTK expression.

(E) RTK expression measured by flow cytometry in six representative NSs. Percentages of positive cells and mean fluorescence intensity are reported.

(F) Left: RTK mRNA expression in the TCGA GBM cohort. Right: Correlation analysis of RTK expression.

(G) Left: RTK protein expression measured by immunohistochemistry (IHC) in the validation GBM cohort (n = 70). Right: Correlation analysis of RTK expression.

(H) RTK co-expression detected by IHC in different tumor areas (upper panels) or different intermingled cells (lower panels). Scale bar, 100 μm.

See also Figure S1; Table S1.

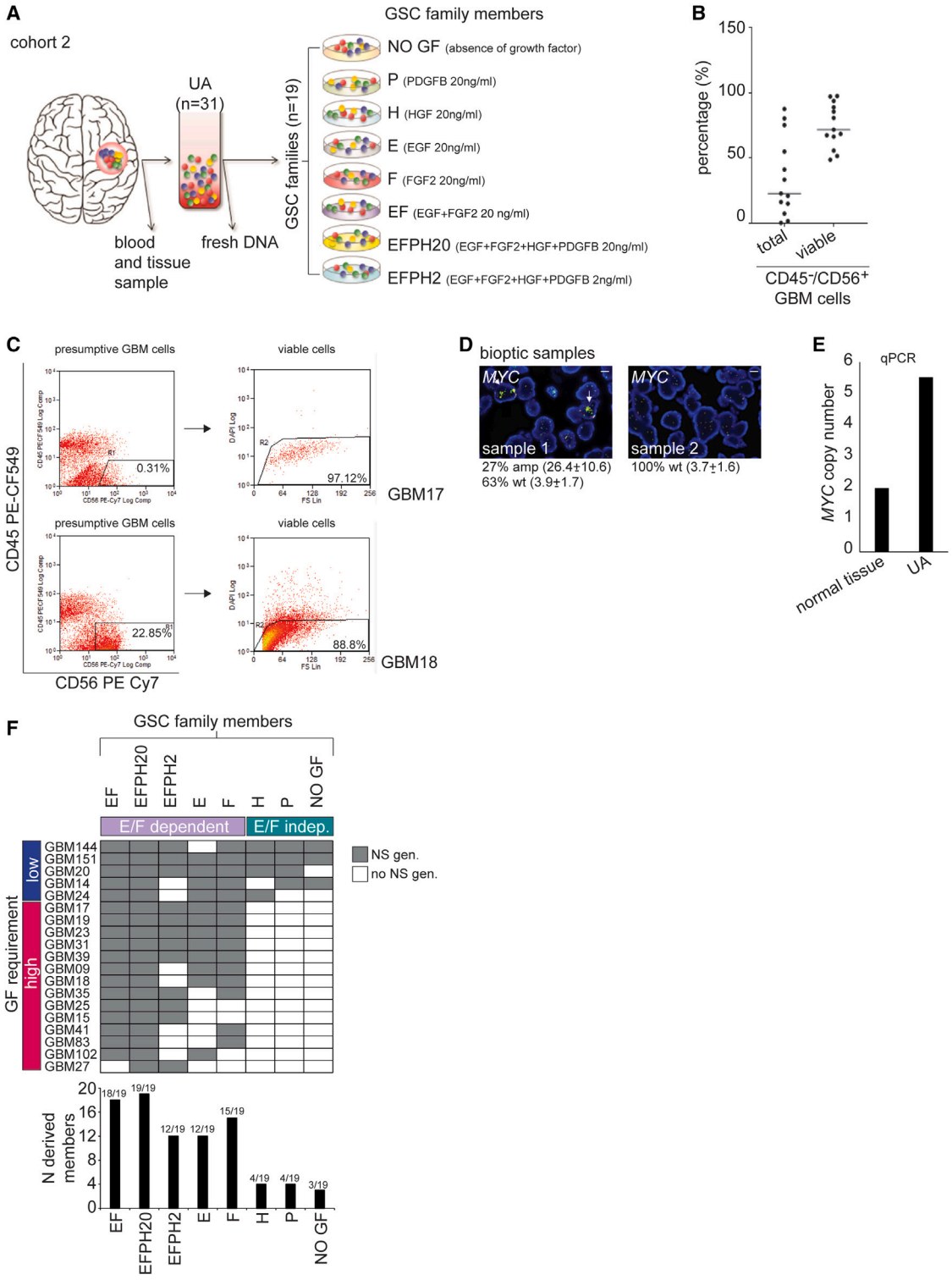


Figure 2. GSC families can be isolated from GBM whole-cell resuspensions (UAs) via different GF cocktails

(A) GSC family derivation protocol. GSC families (n = 19) are selected starting from ultrasonic aspirates (UAs), by supplying the indicated GFs of GF cocktails, each identifying a family "member."

(B) Percentage of total and viable $CD45^{neg}/CD56^{pos}$ GBM cells in a representative UA panel (n = 13). Median is indicated.

(C) Flow cytometry of two representative UAs (GBM17, low cellularity and high viability; GBM18, high cellularity and viability). Percentages of total GBM cells (left: $CD45^{neg}/CD56^{pos}$, gated) and $CD45^{neg}/CD56^{pos}$ viable GBM cells (right: $DAPI^{neg}$, gated) are reported.

(legend continued on next page)

(UAs), which allow recovering of the whole tumor mass, as opposed to conventional small biopsies (cohort 2, n = 31) (Figure 2A; Table S2).[26] In a representative UA subset (n = 13), the percentage of $CD45^{neg}/CD56^{pos}$ GBM cells[27] was variable (median = 22.8%), but cell viability was consistently high (median = 71.7%) (Figures 2B, 2C, S2A, and S2B). Importantly, UA retained the ensemble of gene amplifications, known to be heterogeneously distributed[7,28] and often at risk of being missed in single biopsies (Figures 2D and 2E).

As selective pressures, we used EGF, FGF2, PDGFB, and HGF, alone or in various combinations, resulting in a total of eight different conditions (Figures 2A and 2F). This choice was based on the following observations: (1) these GFs are highly expressed in the brain and GBM microenvironment (Figures S2C–S2E)[29]; (2) the corresponding RTKs are expressed in a mutually exclusive fashion in tumors and GSCs (Figures 1D–1H and S2C–S2E); and (3) these GFs can support GSC proliferation and long-term propagation (Figure S2F).[16,30,31] The combination of the four GFs was used either at a standard (20 ng/mL each: EFPH20) or a reduced (2 ng/mL each: EFPH2) concentration, the latter in order to prevent the counterselection of *EGFR* and *PDGFRA* amplification observed in conventional NSs (Figures S2G and S2H; Table S1).[32,33] Moreover, as highly aggressive cancer stem cells from other tumors can be selected in the complete GF absence,[34] we also used a medium devoid of any GF (Figures 2A and 2F).

The selection protocol was applied to 31 GBM UAs (cohort 2) (Figures 2A and 2F). The ensemble of established NSs selected from each GBM was defined as a GSC "family," and each NS as a GSC family "member." Nineteen GSC families were derived (efficiency ~62%), each including at least two members, indicating that not only the standard medium, but also the other conditions can select NSs. However, the presence of either EGF or FGF2 was usually required; the standard medium (EF) and the richest GF cocktail (EFPH20) displayed the maximum probability to select a member in all families (Figure 2F). Conversely, either PDGFB or HGF alone could select a member only in five of 20 families. Only within the latter families (three of five), a member could be selected in the absence of any GF as well (NO GF; Figure 2F). Overall, GBMs could be grouped into two major subsets, one giving families (n = 14) whose members required the presence of at least EGF or FGF2, the other giving families (n = 5) including also members independent from standard GFs (Figure 2F).

Correlations between the genetic landscape of original GBMs and GSC family generation

Genetic analysis of cohort 2 original GBMs (UAs, n = 31) revealed the presence of driver gene alterations at frequencies similar to those reported by TCGA,[24] except for *CDKN2A* deletion, which was underrepresented (Figure 3A; Table S2). This could be explained by the fact that the relatively small percentage of GBM cells present in UA (Figures 2B and S2B) decreased the chance of detecting *CDKN2A* deletion, although without compromising gene mutation or amplification analysis. In addition, in our cohort, *MYC* or *MYCN* amplification occurred in 13% of cases (Figure 3A; Table S2), a frequency significantly higher compared with TCGA GBM cohort (1.8%),[24] but consistent with recent reports highlighting focal *MYC* amplification in GBM tissues.[35]

Concerning correlations between the original tumor genetic make-up and the probability to derive GSC families, we observed that GBMs harboring *TP53* mutations displayed increased probability to generate NSs, consistently with previous evidence that *TP53* inactivation supports the cancer stem phenotype[36] (Figures 3B and S3A). Conversely, GBMs harboring *EGFR* amplification showed a significantly decreased ability to generate NSs (Figures 3B and S3A). No other GBM genetic alteration could be statistically significantly correlated with generation of GSC families nor was any alteration preferentially associated with the generation of NSs independent of EGF or FGF2 for their propagation (Figures 3A, 3B and S3A).

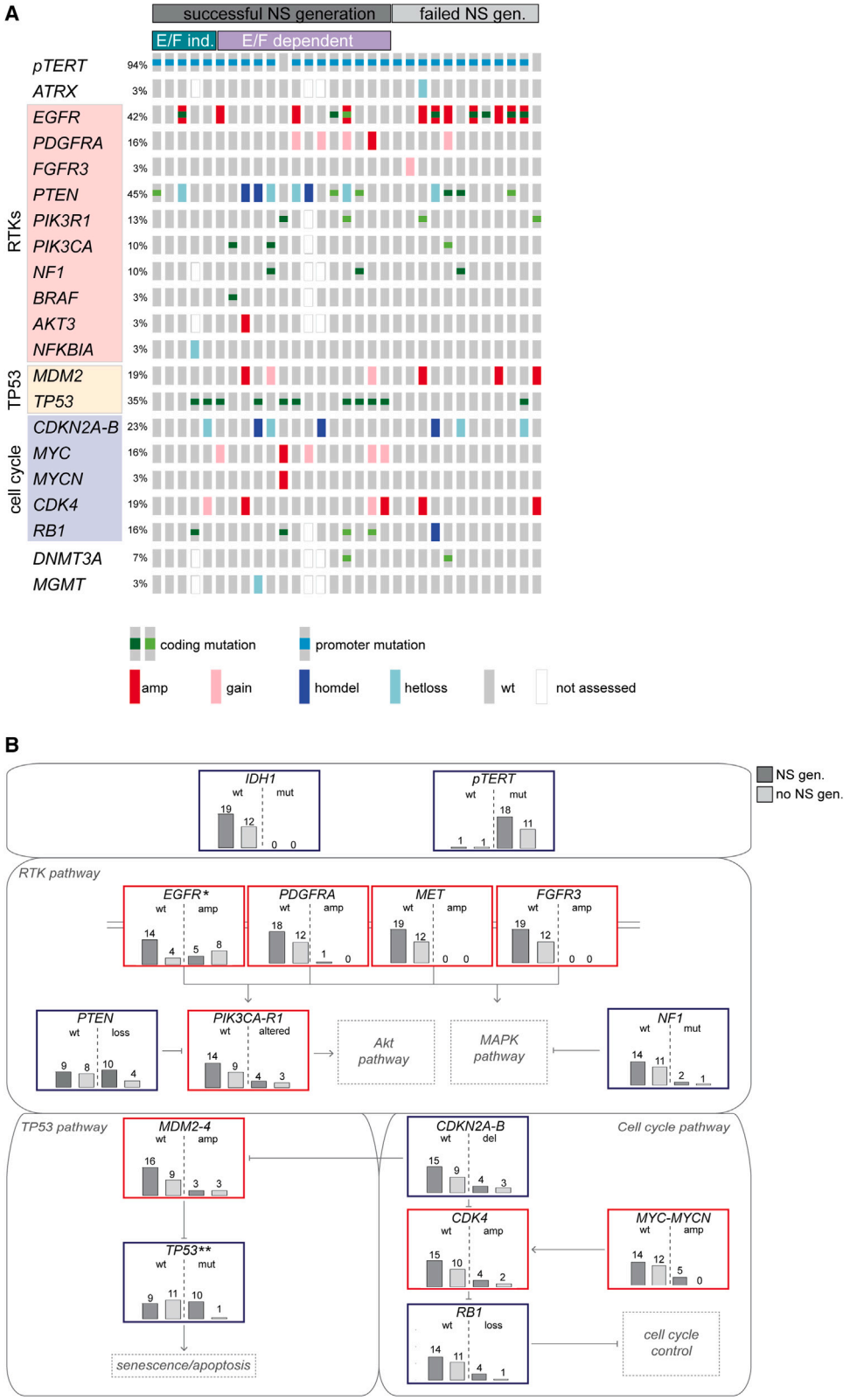
GSC family members share driver mutations and display heterogeneous and plastic gene amplifications

For comprehensive genetic and functional characterization, we chose four GSC families corresponding to original GBMs (UA) with different genetic landscapes, including possibly subclonal or heterogeneous alterations (Figure 3A; Table S2): (1) GBM20 displayed high *EGFR* amplification, *EGFR* mutation, and *PTEN* hemizygous deletion; (2) GBM18 displayed both *EGFR* and *PDGFRA* amplification, and *PTEN* hemizygous deletion; (3) GBM17 displayed *EGFR* amplification and *MYC* copy number (CN) gain; and (4) GBM14 was devoid of RTK amplifications or *PTEN* loss, but, as a peculiarity, displayed *RB1* mutation (Figure 3A; Table S5). The latter mutation could correlate with cell cycle escape from GF control, and indeed the corresponding GSC family included a member able to grow in the absence of GFs (NO GF) (Figure 2B).

In each family, all members were monoclonal, as shown by variant allele frequencies (VAFs) of pathogenic single nucleotide variants (SNVs), i.e., mutations, in GBM driver genes, like in the case of conventionally derived NSs (Figure 4A). Moreover, all driver gene mutations detected in the original tumors were shared among all members of the corresponding families (Figure 4A).

Conversely, driver gene CN gain or amplifications were more heterogeneously distributed among family members and displayed differences with the corresponding tumors (Figure 4A; Table S5).

(D) Fluorescence *in situ* hybridization (FISH) showing *MYC* amplification in two different biopsies from GBM09. The percentage of amplified and/or wild-type cells with gene CN±SEM is indicated. Scale bar, 25 μm.
(E) *MYC* CN measured in UA (GBM09).
(F) GSC family derivation from GBM UAs (n = 31). Rows: GBMs yielding GSC families (n = 19), grouped based on GF requirements for NS generation (high: EGF and/or FGF2 required; low: EGF and/or FGF2 dispensable). Columns: GSC family members (NS) derived in each GF or GF cocktail. E/F dependent: requiring EGF and/or FGF2; E/F indep.: not requiring EGF and/or FGF2. Histogram: N of GSC family members derived in each GF or GF cocktail vs. the total N of GSC families (n = 19).
See also Figure S2; Table S2.



(legend on next page)

In GBM20, all members, like the original tumor, displayed high *EGFR* amplification, as well as *PTEN* monoallelic inactivation (Figure 4A). However, *EGFR* amplification was always extrachromosomal and CN, although always >10 copies/cell, was variable across family members. At the time of NS stabilization (~10 passages, t_1), *EGFR* CN was relatively lower and uniform in members kept in high EGF concentration (EF, EFPH20, and E) (Figures 4B and 4C), consistently with evidence that EGF counterselects cells with high *EGFR* CN (Figure S2G; Table S1). At a later time point (~20 passages, t_2), *EGFR* CN remained substantially unchanged, or with a tendency to decrease in members kept with FGF2 or HGF alone, which were mostly displaying very high CN at an earlier time point (Figures 4B and 4C).

In GBM18, *EGFR* and *PDGFRA* amplifications of the original tumor could be segregated in different members, although both reduced to a CN gain (in particular *EGFR*), with members EF and E retaining *EGFR*, and member F retaining *PDGFRA*. While RTK CN decreased, *PTEN* biallelic loss became evident in all GBM18 members (Figure 4A).

In GBM17, original *EGFR* amplification was lost or strongly reduced, while hemizygous *PTEN* mutation (i.e., biallelic inactivation) became evident in all family members. In addition, *MYC* amplification, barely detectable in the original tumor, was detected in all family members (Figure 4A). Interestingly, *MYC* amplification was extrachromosomal and, at the time of NS stabilization (approximately passage 10, t_1), displayed both quantitative and qualitative differences (formation of double minutes and/or clusters) across family members, without clear correlation with culture conditions (Figures 4A, 4D, and 4E). However, we can observe that the richest GF cocktail (EFPH20) counterselected *MYC* amplification, while the cocktail with a lower GF concentration (EFPH2) primarily selected *MYC* amplification in clusters, which may be associated with higher transcriptional activity.[37] The *MYC* amplification features of each member tended to remain stable over time, with the exception of member E, where amplifications changed from double minutes to clusters (Figures 4D and 4E).

To investigate the ability of GSCs to change their amplification features across multiple cell generations, we performed single-cell cloning of GBM17 member EF, which displayed a high degree of *MYC* amplification heterogeneity (~54% of cells with *MYC* in clusters, ~32% with *MYC* in double minutes, ~14% without amplification). These single cells regenerated NSs that reconstituted the original *MYC* amplification heterogeneity (on average ~61% of cells with *MYC* in clusters, ~30% with *MYC* in double minutes, ~9% without amplification), suggesting stochastic distribution and plasticity of amplifications in the original NS cell population (Figures 4E and S3B).

Finally, in GBM14, *RB1* and *TP53* biallelic inactivation were retained by all members (Figure 4A).

In summary, within each family, different GSCs retained from original GBMs and fully shared the same driver gene mutations. The presumptive subclonal composition of the original tumor, suggested by coexistence of two different RTK amplifications, or RTK amplification(s) and *PTEN* deletion were segregated only in the case of *EGFR* and *PDGFRA*. However, the characteristics of gene amplification observed within GSC families, and among cells of the same GSC culture, suggest that, unlike SNV or deletion, gene amplification is not a fixed feature that unequivocally identifies a subclone, but, rather, a plastic feature that can be influenced by the relative abundance of exogenous GFs in the microenvironment.

Within each family, GSC members display overall similar mutational landscapes

The degree of overall genetic similarity among members of the same GSC family, and similarity between the original GBM (UA) and the GSC family, was reconstructed (Figure 4F). In all members of each family, the majority of SNVs displayed high VAF, further attesting NS monoclonality, and were shared among all members of the same family, mirroring the homogeneity of driver genes (Figures 4F and 4A). Besides such fully shared high-frequency SNVs, each GSC family member displayed a lower percentage of private or partially shared SNVs, possibly stochastically associated with different culture conditions. By comparing each GSC family with its original GBM, we observed that, with a few exceptions (GBM17), no *de novo*, private, and high-frequency SNVs appeared in family members with respect to the GBM (Figure 4F). Conversely, each original GBM displayed a significantly higher proportion of low-frequency SNVs, as compared with NSs. In a minor part, such SNVs became high-frequency in NSs, and in a greater part they were lost, indicating counterselection of subclonal tumor SNVs during NS generation (Figure 4F). However, no SNV exclusive of any member (either *de novo* or positively selected from the original GBM) had a presumptive pathogenic role.

GSC family members display heterogeneous transcriptional profiles

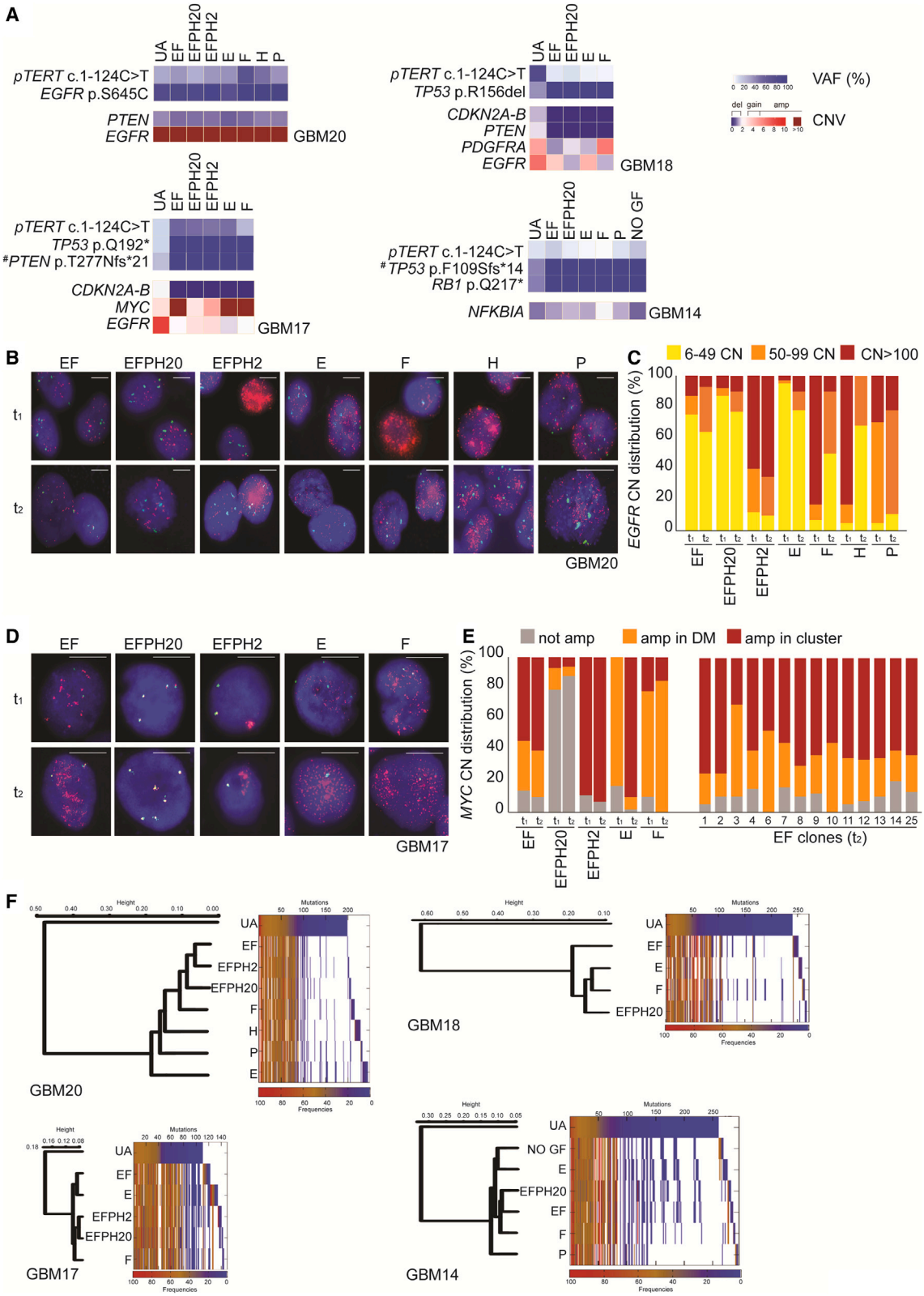
After transcriptional profiling, GSC family members were subtyped into mesenchymal, classical or proneural.[25,38] These subtypes have been previously associated with specific genetic alterations, in particular the classical subtype with *EGFR* and the proneural with *PDGFRA* amplification.[39] Consistently, we observed that retained *EGFR* amplification uniformly correlated with a classical profile in all members of GBM20 family (Figure 5A; Table S3). Moreover, unsupervised clustering analysis revealed that GBM20 members were highly homogeneous (Figure 5B). Conversely, in the other families, members displayed different subtyping and more heterogeneous global transcriptional profiles, causing member distribution in different clusters

Figure 3. Correlations between the genetic landscape of original GBMs and GSC family generation

(A) Genetic alteration frequencies in cohort 2 GBMs (n = 31). GBMs are grouped according to NS generation and dependence on EGF/FGF2.

(B) Genetic alterations and NS generation. Bars indicate the number of GBMs, either wild-type or altered in each gene, that generated NSs (dark gray bars) or not (light gray bars). *Anti-correlation between *EGFR* amplification and NS generation ($p = 0.05$); **Positive correlation between *TP53* mutation and NS generation ($p = 0.02$) (Fisher's exact test).

See also Figure S3A; Table S2.



(legend on next page)

(Figures 5A and 5B; Table S3). In GBM17 family, which lost the original *EGFR* amplification but retained heterogeneous *MYC* amplification (Figures 4A and 4D), member F was subtyped as classical, while the others were mesenchymal (Figure 5A; Table S3). In GBM18 family, which segregated *EGFR* and *PDGFRA* CN gains in different members, members E (*EGFR* CN gain) and F (*PDGFRA* CN gain) were both subtyped as classical, member EFPH20 (no RTK gain) as mesenchymal, and member EF (*EGFR* gain) remained unclassified. In GBM14 family (devoid of any gene amplification), members EF, EFPH20 and F were mesenchymal, member E was unclassified and the two GF-independent members (NO GF and P) were proneural (Figure 5A; Table S3).

The transcriptional profile of GSC family members could not be compared with original UA, where low tumor cellularity prevented reliable transcriptional characterization (Figure S2B). However, in GBM17 and GBM18 families, the transcriptional profile was analyzed at different time points during GSC propagation, showing that, at low passages ($\leq$8, t_0), each member tend to be more heterogeneous, being often (in 5/8 cases) assigned to 2 or 3 different subtypes (Figure 5C; Table S3). At the time of NS stabilization ($\sim$10–12 passages, t_1) each member displays a uniform subtype that is maintained for long time (up to >20 passages, t_2), in some cases tending to blur (Figure 5C; Table S3).

The transcriptional profile of each GSC family member seems independent from the type of GF(s) used for selection and propagation. However, although statistical analysis was affected by the small numbers, we could observe that mesenchymal GSCs were preferentially selected by GF cocktails, rather than by single GFs (Figure 5D).

Principal-component analysis (PCA) confirmed relative homogeneity of GBM20 family members, and relative member heterogeneity within the other families (Figure 5E). Of note, members of different families belonging to the same subtype (classical or mesenchymal) displayed a high degree of similarity, even greater than that shared with other members of the same family (Figure 5E), as observed also in the unsupervised clustering analysis (Figure 5B). This suggests that the transcriptional profile is relatively independent from the genetic make-up, and may heavily rely on other molecular factors, possibly related to epigenetic control and developmental stage of the cell of origin.[40]

In summary, within GSC families, members displayed different transcriptional profiles despite mutational homogeneity, and with inconstant correlation with gene amplifications. Moreover, GSC family members displayed transcriptional similarities across families, irrespectively of different genetic landscapes.

GSC family members display distinctive RTK expression

Within GSC families, transcriptional heterogeneity was mirrored by different RTK expression, (Figures 5F–5H). EGFR was mostly expressed by classical members, and MET (HGFR) by the mesenchymal, consistently with previous evidence that the two receptors are subtype markers,[30,33,39] while PDGFRA expression was more promiscuous, and, in general, displayed by a lower percentage of cells in each GSC family member (Figures 5G and 5H). Interestingly, the two GBM14 family members able to grow in the complete absence of GFs (NO GF) or PDGF alone (P), both subtyped as proneural, did not express *EGFR* or *MET* or *PDGFRA*, but only *FGFRs* (Figures 5F–5H and S3C). Global transcriptomic analysis of all the RTKs (*EGFR*, *FGFR* and *PDGFR* families, and *MET*) for the GFs used for GSC derivation showed that mesenchymal GSCs, beside *MET*, expressed statistically significantly higher levels of *PDGFR*s and *FGFR*s compared to classical or proneural GSCs (Figure S3C). This suggests that mesenchymal GSCs are more predisposed to respond to GF cocktails compared with the other GSC subtypes, consistently with the tendency to be selected by such media (Figure 5D).

Consistently with findings in GBM tissues (Figure 1H), flow cytometry showed also that RTKs were coexpressed in the original GBM (UA) but tended to segregate into different GSC family members, according to the NS subtype at the time of stabilization (tendency to mutual exclusivity between EGFR and MET: $p = 0.7$, between EGFR and PDGFRA: $p = 0.4$; tendency to co-occurrence between MET and PDGFRA: $p = 0.1$) (Figure 5H). These data indicate that heterogeneous aspects coexisting in GBM tissues are singularly recapitulated by distinct GSC family members, thanks to differential selective pressures exerted through critical subtype functional markers such as RTKs.[41]

GSC family members display heterogeneous phenotypes and malignancy mirroring their transcriptional subtypes

Within each GSC family, members displayed typical stem and tumorigenic properties. However, with the exception of the GBM20 family, such properties differed quantitatively among members, and tumorigenicity displayed a striking correlation with the transcriptional subtype. Stem cell frequency tended to be higher in classical compared with mesenchymal members

Figure 4. GSC family members display similar mutational landscapes and heterogeneous gene amplifications

(A) GBM driver gene pathogenic alterations (mutation variant allele frequency [VAF] and CN variation [CNV]) detected by whole exome sequencing (WES) and validated in four paradigmatic GBMs (UA) and corresponding GSC families. #VAF estimated by integrating WES and Sanger Sequencing data.

(B) FISH showing *EGFR* amplification levels in GBM20 family members at the time of NS culture stabilization (t_1: $\sim$10 passages) and at $\sim$20 passages (t_2). Scale bar, 25 μm.

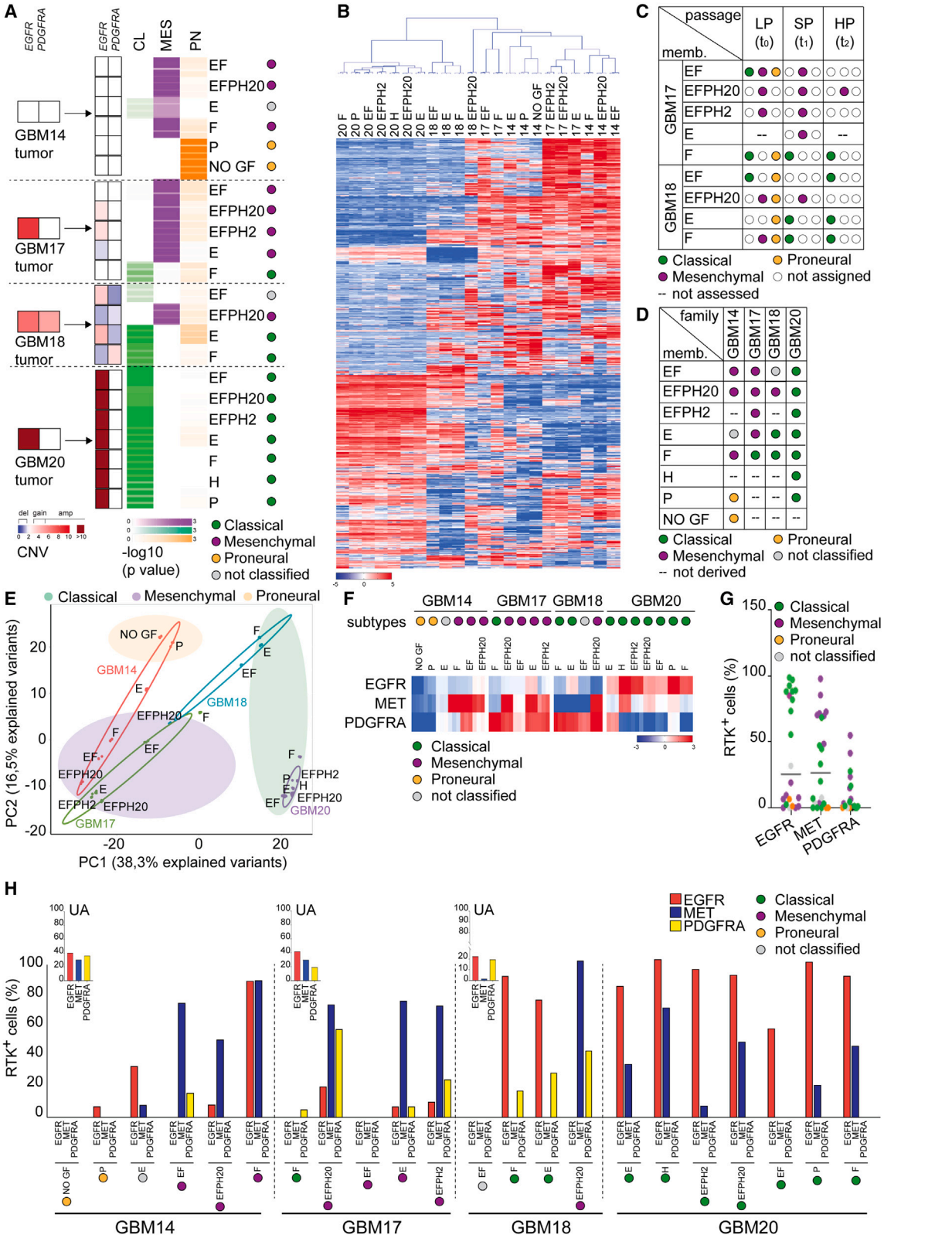
(C) Fraction of cells in each GBM20 family member with respective CN as in (B).

(D) FISH showing *MYC* amplification levels and modalities in GBM17 family members at the time of NS culture stabilization (t_1: $\sim$10 passages) and at $\sim$20 passages (t_2). Scale bar, 25 μm.

(E) Fraction of cells with or without *MYC* amplification in double minutes (DM) or clusters, in each GBM17 family member as in (D), and in NS subclones of the EF member.

(F) Genomic similarities among parental GBMs (UAs) and matched GSC families. In each graph, hierarchical clustering (left) shows distance between samples (height) and heatmap (right) shows the number of fully or partly shared, and private mutations, with respective VAFs (frequencies).

See also Figure S3B; Table S2.



(legend on next page)

(Figures 6A and S4A–S4D). The proneural members, present only in GBM14 family (NO GF and P), displayed a lower stem cell frequency compared with classical and mesenchymal (Figures 6A and S4A). *In vivo* tumorigenicity mostly reflected stem cell frequency, with classical GSCs displaying statistically significant earlier tumor onset ($p = 0.01$) and higher tumor take ($p = 0.005$), compared with mesenchymal GSCs, across all families (Figures 6B and S4E–S4H). As a notable exception, GBM14 proneural members, endowed with a relatively low *in vitro* stem cell frequency, displayed the highest tumorigenic potential compared with any other GSCs of the same or other families (Figures 6B and S4E). Kaplan-Meier curves were consistent with tumorigenic potential, showing the shortest survival in mice injected with proneural, intermediate survival in mice injected with classical, and longest survival in mice injected with mesenchymal GSCs (Figure 6C).

The apparent paradox of GBM14, where low stem cell frequency associated with the highest tumorigenic potential, could be explained by the proliferative autonomy (independence from exogenous GFs) displayed by GBM14 proneural members. Indeed, the latter could be propagated in the complete absence of GFs or in the presence of the sole PDGF, for which they lack the receptor (Figures 5F, 5H, and 6D). We previously associated cancer stem cell proliferative autonomy with exceedingly high tumor-initiating frequency.[34] In these GSC families, we found that tumorigenic potential mirrored not only the transcriptional subtype but also the relative sensitivity to GFs. Classical GSCs (all selected and propagated in the presence of GFs), when switched to a GF-deprived medium, overall displayed a more limited ability to autonomously proliferate compared with proneural; mesenchymal GSCs were overall unable to self-sustain proliferation (Figures 6D and S5A–S5D). Conversely, after exposure to the richest GF cocktail (EFPH20), the proliferative response was the highest in the mesenchymal, intermediate in the classical, and negligible in the proneural group (Figures 6E and S5E–S5H). These data indicate that mesenchymal GSCs, expressing the global highest level of RTKs (Figures 5G and S3C), vigorously respond to GFs and may be strongly dependent on the tumor microenvironment, which, in the immunocompromised host, can be depleted of essential cues.

In differentiation assays, classical and mesenchymal members in all families could be induced to pseudodifferentiate and expressed multiple differentiation markers, with no seemingly preferential association between GSC subtype and differentiation lineage (Figures 6F–6I and S6). Interestingly, the two GBM14 proneural members endowed with proliferative autonomy were resistant to differentiation, a feature that, again, can correlate with increased tumor aggressiveness (Figure 6F).

These findings indicate that the presence of high-level *EGFR* amplification correlates not only with high genetic and transcriptional homogeneity, but, consistently, with biological homogeneity of different GSC family members. In the absence of *EGFR* amplification, GSC transcriptional profiles and biological properties are more heterogeneous within GSC families. Across multiple families, the tumorigenic potential better correlates with the transcriptional rather than the genetic profile.

GBM14 family members segregate distinct pathological features of the original GBM-PNC tumor

Among the four families, GBM14 represents a paradigmatic example. Indeed, on the one hand, GBM14 family members display a relatively homogeneous genetic landscape, as they fully share genetic drivers and show the highest degree of overall genetic similarity among families (Figures 4A and 4F). On the other hand, GBM14 family members display marked biological differences: NO GF and P members (proneural subtype) concomitantly display autonomy from exogenous GFs and high tumorigenic potential, while the remaining members (mesenchymal or unclassifiable subtype) are dependent of exogenous GFs and less aggressive (Figures 5A, 6A–6E, S4A, S4E, S5A, and S5E; Table S3).

The original GBM14 tumor was diagnosed as a GBM with a primitive neuronal component (GBM-PNC), a rapidly lethal GBM variant recognized by the World Health Organization classification, featuring a distinctive methylation profile and frequent association with *TP53* and *RB1* inactivation (as found in GBM14, Figure 4A; Table S2).[42–44] Histopathologically, GBM-PNCs include areas defined as GBM component (positive for GFAP staining) and areas defined as PNC component (negative for GFAP and positive for synaptophysin staining) (Figure 7A).[43] Analysis of tumors formed by GBM14 family members showed

Figure 5. GSC family members display heterogeneous transcriptional profiles and distinctive RTK expression

(A) GSC family member classification into mesenchymal (MES), classical (CL), and proneural (PN) subtypes. Color intensity is proportional to probability of subtype assignment. Full color dots: significant subtype assignment for a single classifier ($p < 0.05$). Gray dots: lack of any subtype assignment. *EGFR* and *PDGFRA* CNV of each original GBM and GSC family member are also indicated.

(B) Unsupervised clustering of GSC family members (columns) based on expression of the 1,000 most differentially expressed genes (rows).

(C) GBM17 and GBM18 family member classification at different NS culture passages: low (LP, t_0: $\leq$8p in GBM17; $\leq$5p in GBM18); stabilization (SP, t_1: $10 < p < 15$); high (HP, t_2: $\geq$15p). Colored dots: significant subtype assignment ($p < 0.05$). White dots: not significant subtype assignment ($p \geq 0.05$). –: not assessed.

(D) Transcriptional subtypes selected by different GFs (rows) in each GSC family (columns). Colored dots: GSC transcriptional subtype as in (A). Preferential association of mesenchymal subtype with GF cocktails vs. single GFs ($p = 0.07$, Fisher's exact test).

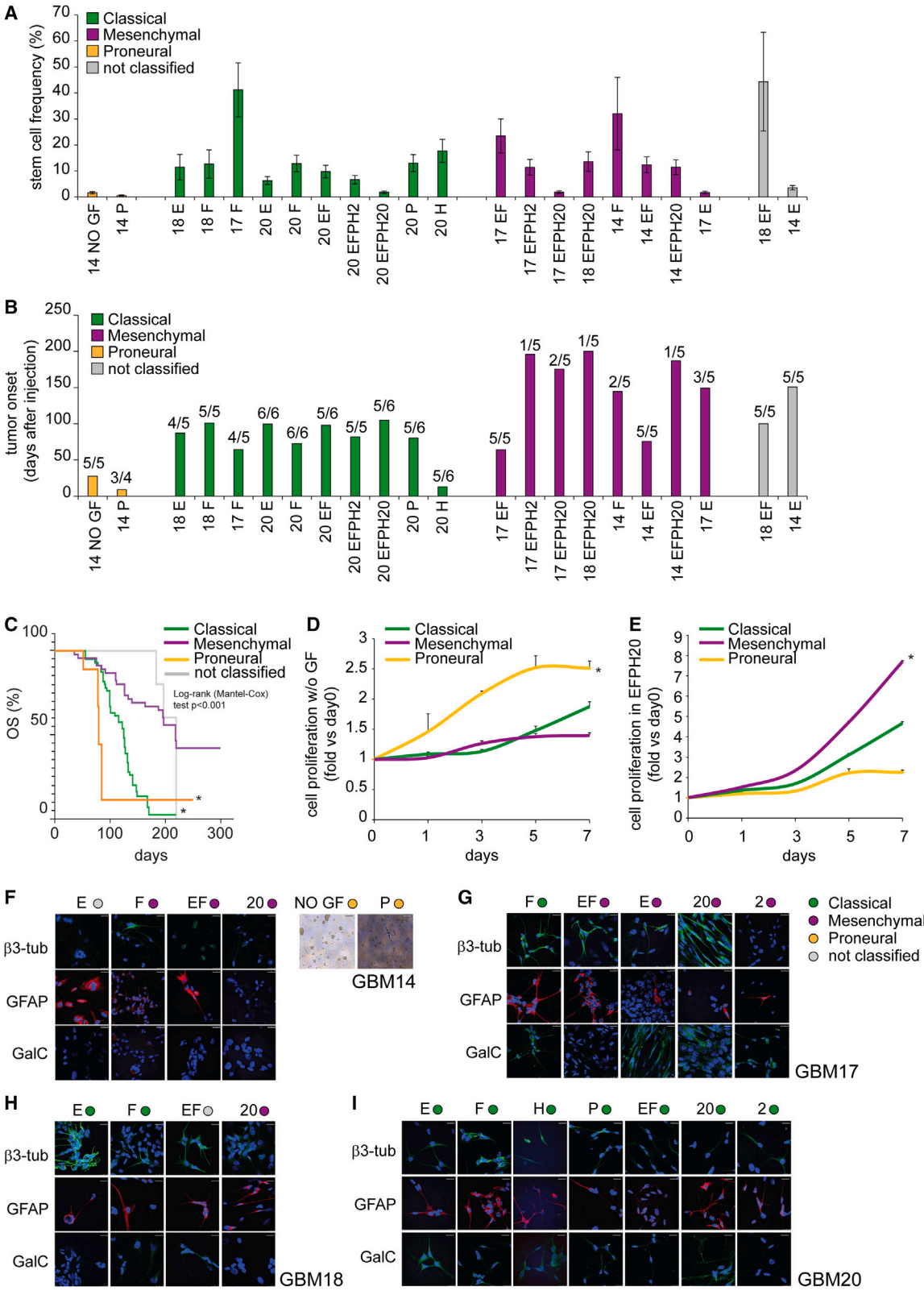
(E) Principal-component analysis (PCA) performed on most differentially expressed genes as in (B). Each dot represents a GSC family member. Colored areas identify transcriptional subtypes.

(F) RTK mRNA expression in GSC family members (RNAseq L2R). Transcriptional subtypes as in (A) are indicated.

(G) Percentage of cells expressing each RTK in overall GSC family members identified by transcriptional subtype, measured by flow cytometry. Median is indicated.

(H) Percentage of cells expressing each RTK in GBM UAs (insets) and GSC family members (flow cytometry representative data).

See also Figure S3C; Table S3.



(legend on next page)

that autonomous and aggressive NO GF and P members regenerated a pure PNC component, while the other members regenerated a pure GBM component (members F, EF) or a heterogeneous GBM-PNC (member E) (Figure 7A). The only tumor generated by the EFPH20 member (Figure S4E) was too small to be exhaustively analyzed. However, the close transcriptional similarity between EFPH20 and the members giving a pure GBM component (F and EF, Figure 5E) suggests that EFPH20 could generate a pure GBM as well.

GBM14 family members segregate distinct transcriptional programs related either to GBM or PNC

To better characterize the distinctive features of GSCs regenerating the one or the other GBM-PNC component, we investigated the most differentially expressed genes in the group of PNC-regenerating family members (P and NO GF) vs. those regenerating a pure GBM (EF and EFPH20, chosen for their reciprocal close transcriptional proximity and maximum distance from P and NO GF, Figure 5E). This analysis identified a GBM-PNC signature encompassing 2,401 genes (1,293 upregulated in PNC and 1,108 upregulated in GBM) (Figure 7B; Table S3). GBM-PNC signature gene ontology showed a striking enrichment in genes related to neurogenesis and neuronal fate in PNC-regenerating members, and in genes associated with typical mesenchymal GBM features in the GBM-regenerating members (Figure S7A; Table S3).[45] Application of the GBM-PNC signature to the entire GBM14 family highlighted that member F (regenerating a pure GBM) specifically upregulated genes associated with mesenchymal GBM, while member E (regenerating a GBM-PNC) displayed an overall modest activation of both GBM and PNC transcriptional programs (Figures S7B and S7C).

Among the most expressed genes in PNC-regenerating members, we noticed *DCX*, *CD24,* and *SOX11* (Figure 7B; Table S3), known to be tightly associated with the so-called ''neural-progenitor-like (NPC-like)'' state, identified in GBM by single-cell analysis.[5] Other highly expressed genes such as *RBFOX3* (encoding NeuN), *NEUROD1*, and *BTG2* were related to an early developmental neural stage as well.[46] Conversely, in the GBM component, typical mesenchymal markers such as *YAP1*, *VIM* (vimentin), *CD44*, and *MET* were specifically expressed (Figure 7B; Table S3).[39] Application of gene signatures identified by Neftel et al.[5] recognized PNC-regenerating members as ''NPC-like'' and GBM-regenerating members as ''mesenchymal-like (MES-like)'' (Figures 7C and 7D). The member able to regenerate both PNC and GBM components (E) was defined as ''astrocyte cell-like (AC-like)'' (Figures 7C and 7D). Specific, mutually exclusive expression of NPC-like or MES-like genes was confirmed in family members that regenerated the PNC (NO GF) or the GBM (EF) aspect of the tumor, respectively, while the member with bifunctional properties (E) displayed expression of MES-like together with NPC-like markers (Figures 7E and 7F).

In line with bulk transcriptomic profiling, single-cell analysis revealed that GSCs regenerating GBM only (member F) were homogeneously assigned to the MES-like subtype (>95% of cells), while those regenerating PNC only (member NO GF) were assigned to the NPC-like subtype (~45%) or to the MES-like (<30%), or remained unclassified (Figures 7G and S7D; Table S3). GSCs able to regenerate both GBM and PNC (member E), included a majority of cells (~65%) assigned to the MES-like subtype, but also a few cells assigned to the NPC-like (~2%) or AC-like (~0.5%) subtypes (Figures 7G and S7D; Table S3). Consistently, GSCs reproducing GBM or PNC only expressed high levels of MES-like or NPC-like signature markers, respectively, while GSCs reproducing both GBM and PNC displayed intermediate levels (Figure S7E). In addition, by applying the GBM-PNC signature to single cells (Figure S7B), we observed that, while member NO GF and member F mostly display sharp PNC or GBM features, member E includes a majority of cells with intermediate transcriptional features, which can correspond to GSCs in an earlier or more plastic transcriptional state (Figure S7F). Overall, single-cell analyses suggest that member F contains a homogeneous MES-like cell population, enabling selective regeneration of the GBM component of the original tumor upon transplantation. Member NO GF is more heterogeneous but includes a predominant population with sharp NPC-like features, which tend to take over upon transplantation, thus reproducing the PNC component of the tumor. Finally, member E includes heterogeneous cells that may be able to regenerate both aspects of the tumor because of the concomitant presence of NPC-like cells together with abundant MES-like cells, or because they retain a transitional state, open to evolve either into GBM or PNC.

Finally, by *in situ* RNA hybridization on the original GBM14 tumor, we found that the GBM or PNC component expressed

Figure 6. GSC family members display heterogeneous phenotypes and malignancy mirroring their transcriptional subtypes

(A) Stem cell frequency in GSC family members. Limiting dilution assay data are represented as mean ± SEM of at least two independent experiments.

(B) Timing of tumor onset after injection of GSC family members. Tumor take (n of mice developing tumors/n of injected mice) is indicated.

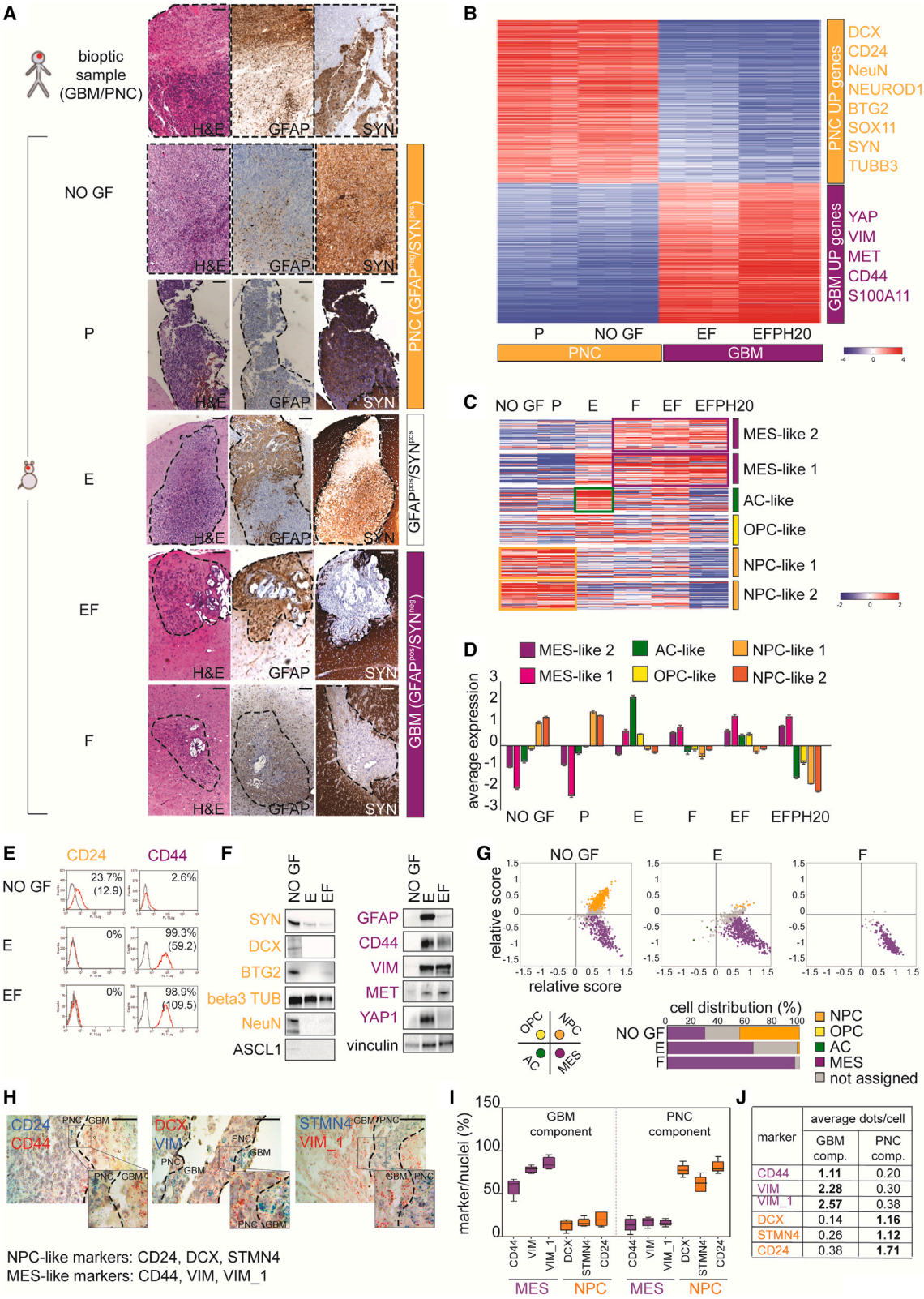
(C) Survival curves (overall survival [OS]) in mice groups. Each group includes mice injected with all members of the four GSC families with the same subtype (PN: two members, total mice: n = 9; CL: 10 members, total mice: n = 56; MES: eight members, total mice: n = 40; not classified: two members, total mice: n = 10) *Log rank Mantel-Cox test, $p < 0.001$.

(D) Proliferation curves in GSC family member groups measured in the absence of GFs. Groups include all members of the four GSC families with the same subtype (PN: n = 2; CL: n = 10; MES: n = 8). Data are shown as mean ± SEM of at least two independent experiments. *Student's paired t test, $p < 0.05$.

(E) Proliferation curves in GSC family members groups, measured in the presence of a full GF cocktail (EFPH20). Groups include all members of the four GSC families as in (D). Data are shown as mean ± SEM of at least two independent experiments. *Student's paired t test, $p < 0.05$.

(F–I) Expression of differentiation markers in GBM14 (F), GBM17 (G), GBM18 (H) and GBM20 (I) family members cultured in prodifferentiating conditions. GBM14 NO GF and P members did not adhere to the substrate (lack of differentiation, shown by light microscopy) and could not undergo immunofluorescence. beta3-tub, tubulin beta 3 (neuronal marker); GFAP, glial fibrillary acidic protein (astrocytic marker); GalC, galactosylceramidase (oligodendroglial marker). Scale bar, 25 μm.

See also Figures S4–S6.



(legend on next page)

preferentially and in a mutually exclusive manner the markers characterizing the MES-like or NPC-like state in GSC family members, respectively (Figures 7H–7J). These data indicate that different GSC family members such as member F and NO GF can consistently segregate the molecular determinants of distinct aspects of the original tumor, retain those determinants during *in vitro* propagation, and faithfully reproduce the related distinct tumor aspects upon *in vivo* transplantation.

DISCUSSION

In GSC families obtained through selective pressures represented by different combinations of GFs, we found that driver gene mutations and deletions were relatively homogeneous. Although it is expected that each GBM undergoes genetic evolution, it is unsurprising that, at clinical presentation, the clone harboring the majority of driver gene mutations has become prevalent, and/or it has the highest probabilities of being selected by any culture condition. However, we found that GSCs could segregate different RTK amplifications, which are known as late alterations unevenly distributed within each GBM.[7–9]

GSCs can contribute to elucidate how intratumor gene amplification heterogeneity can result from plasticity of gene amplifications residing in extrachromosomal DNA, which were recently highlighted in GBM and other tumor tissues.[28,35,37,47] In GSC families, different members display varying levels and/or patterns of extrachromosomal amplification of the same gene. Moreover, this heterogeneity is reproduced in single-cell subclones of the same GSC culture, allowing for the coexistence of cells with different levels of the same oncogene derived from the same mother cell. This protean, non-mendelian inheritance of extrachromosomal material can favor counterselection of RTK amplifications during GSC derivation, as observed in this study and in previous reports,[32,33] and it has significant implications for targeted therapies since RTKs are frequently amplified genes and top candidates for drug targeting.

The high *EGFR* amplification observed in all members of the GBM20 family, despite quantitative differences, consistently associated with a classical subtype, as expected.[39] However, GSCs can exhibit a classical profile even in the absence of *EGFR* amplification, suggesting that this profile may depend more strongly on the cell type or developmental stage at which transformation occurs, rather than on the genetic alteration itself.

Unlike in the *EGFR*-amplified GBM20 family, members of the other families, despite sharing largely common mutational landscapes, display heterogeneous transcriptional profiles, and, not surprisingly, different biological properties. Classical GSCs consistently exhibit higher tumorigenic potential compared with mesenchymal GSCs across multiple tumors. As GBMs with a bulk classical transcriptional profile do not display a different prognosis from those displaying a bulk mesenchymal or proneural profile, our data suggest that the inherent greater tumorigenicity of classical GSCs can result from the absence of an immune system in the mouse xenograft. Conversely, in the human microenvironment, mesenchymal GSCs, which, as we show, can respond more vigorously to exogenous GFs, may find appropriate cues and even prevail on classical subclones. Indeed, in recurrent GBMs, the tumor tissue consistently shows enrichment of growth factor-producing cells, while GBM cells tend to express mesenchymal programs,[48] indicating that the mesenchymal phenotype can be both selected and actively sustained by the microenvironment. Moreover, the ability to respond to multiple exogenous GFs can confer a double selective advantage, including increased proliferative drive and protection against therapeutic damage.[49]

As a paradigmatic case, we found that genetically homogeneous GSCs from the same family (GBM14) segregated two distinct histopathological aspects (GBM and PNC or "primitive neuronal component," respectively), which coexisted in the original tumor. The PNC component is known as a sign of hypermalignant progression and correlates with a harsher prognosis.[43,50] Interestingly, the GSCs regenerating the PNC component

Figure 7. GBM14 family members segregate distinct pathological features of the original tumor associated with distinct transcriptional programs

(A) Histopathology of parental GBM14 tumor and of representative experimental GBMs generated by GSC family members. Staining of sequential sections for specific GBM (GFAP, glial fibrillary acidic protein) or PNC (SYN, synaptophysin) markers, or hematoxylin and eosin (H&E). Dotted lines: tumor areas. Scale bar, 200 μm.

(B) Expression levels of genes identified as upregulated in GSCs generating PNC (P and NO GF) or GBM (EF and EFPH20) experimental tumors. Full gene list is in Table S3.

(C) Expression levels of gene signatures identified by Neftel et al.[5] in GBM14 family members.

(D) Metagenes calculated as the average L2R expression levels of gene signatures as in (C).

(E) Flow cytometry showing NPC-like (CD24) and MES-like (CD44) marker expression in GBM14 members generating PNC (NO GF), GBM (EF), or mixed GBM-PNC (E) experimental tumors. Percentage of positive cells and mean fluorescence intensity (MFI) are reported.

(F) Western blots showing expression of NPC-like (orange) and MES-like (purple) markers in GSCs as in (E). ASCL1, OPC-like marker; vinculin, loading control.

(G) Transcriptional subtyping of cells from GBM14 family members based on scRNA profiling. Bar graphs represent the percentage of cells significantly assigned to each subtype.

(H) *In situ* RNA hybridization of GBM14 original tumor tissue for NPC-like and MES-like markers. Dotted lines separate GBM and PNC components. Squares: higher magnifications. Scale bar, 100 μm.

(I) Nested box plot showing the percentage of cells positive for the indicated MES-like and NPC-like markers in the GBM or PNC component as in (H). MES-like markers are significantly more expressed than NPC-like markers in the GBM component (nested t test, $p < 0.002$), and vice-versa in the PNC component (nested t test, $p < 0.001$).

(J) Expression levels (average dots/cell) of the indicated markers in the GBM or PNC component as in (H). Bold: significantly higher value in one component vs. the other (nested t test, $p < 0.02$).

See also Figure S7; Table S3.

display, unlike the other GSC family members, the ability to self-sustain long-term propagation in the absence of exogenous GFs. This ability is emerging as a common trait of stem-like cells from highly aggressive tumors, such as ''cancers of unknown primary'' origin, which exhibit a differentiation-resistant phenotype, preventing identification of the tissue of origin, coupled with high metastatic ability.[34] Interestingly, GSCs regenerating the PNC component display an NPC-like (neural progenitor cell-like)[5] profile, typical of the early neural stem cell, suggesting propensity to remain in an undifferentiated state.

The NPC-like status has been associated with *CDK4* amplification,[5] which is absent in GBM14 tumor or family members. However, in GBM14, we detected *RB1* loss, which is characteristic of GBM-PNC[42,51] and shares properties with *CDK4* amplification, as both alterations can disrupt cell cycle restrictions and sustain autonomous proliferation. However, *RB1* loss alone is likely insufficient to impose the NPC-like transcriptional profile and the PNC histopathological phenotype, since syngeneic GSCs from the GBM14 family (all sharing *RB1* loss) can generate either the PNC or the GBM (mesenchymal) component. These data indicate that the appearance of a PNC component in GBM can occur without obvious genetic evolution but through the stabilization of a transcriptional program retaining GSCs in an early neuronal state. A relevant next question is to understand the molecular basis of this differentiation block. Intriguingly, the GBM14 family includes a member (E) that can reproduce both GBM and PNC aspects, and display mixed transcriptional features, providing an opportunity to investigate presumptive transitional states and mechanisms.

In conclusion, most GBMs consist of different functional subclones, each sustained by a GSC running a distinctive transcriptional program. However, the number of subclonal typologies appears to be limited, and the same typology seems to recur across different patients irrespective of the genetic landscape, which lays the foundation for recognizing shared vulnerabilities. The parallel derivation of GSCs by imposing the pressure of different GFs or GF cocktails allows the isolation of GSCs with distinctive properties from individual tumors, and helps us better understand the cell populations on which microenvironmental factors, including therapies, exert their selective pressure, thereby propelling GBM therapeutic resistance and recurrence.

Limitations of the study

The GSC families characterized in this study, although representative of the main GBM subtypes, may fail to capture the overall genetic and transcriptomic characteristics of GBM patients. The molecular comparison between GSCs and original tumors is limited. The selection of GFs and GF cocktails used to isolate different GSCs from single tumors was based on rational criteria, but these factors cannot fully recapitulate the complexity and individual variations of the microenvironmental cues that act on GBM cells. Additionally, the assessment of GSC tumorigenicity was performed in immunocompromised mice, which lack essential components of the GBM microenvironment. Therefore caution must be taken in interpreting the relationships between GSC molecular features and their *in vivo* properties.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human subjects
  - Neurosphere (NS) derivation
  - Conventional cell lines
  - Mice and experimental tumors
- METHOD DETAILS
  - DNA and RNA extraction and retrotranscription
  - Gene copy number evaluation
  - Gene sequencing
  - Droplet digital PCR (ddPCR)
  - NGS target panel analysis
  - WES analysis
  - FISH analysis
  - Bulk gene expression profiling
  - scRNA-seq and data processing
  - Quantitative real-time PCR (qPCR)
  - *In situ* RNA hybridization (RNAscope)
  - Western blotting
  - Flow-cytometric analysis
  - Pseudodifferentiation assay
  - Immunohistochemistry
  - Extreme limiting dilution assay
  - Cell proliferation assay
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112816.

ACKNOWLEDGMENTS

We thank D. Cantarella, S. Gilardi, S. Giove, B. Martinoglio, M. Patanè, R. Porporato, P. Bernabei, and N. Calandra for technical help and D. Gramaglia for secretarial assistance. This work was supported by the AIRC-Italian Association for Cancer Research under grants Special Program Molecular Clinical Oncology 5×1000 N.21052 (to C.B. and P.M.C), IG N.19933 (to C.B.), and N.23820 (to P.M.C.); Italian Ministry of Health RC 2022-2023 (to C.B. and S.P.); and Comitato per Albi98 (to C.B.). The graphical abstract and Figure S8 were created with Biorender.com.

AUTHOR CONTRIBUTIONS

Conceptualization, F.D.B., F.O., and C.B.; methodology, F.D.B. and F.O.; formal analysis, F.D.B., F.O., G.C., C.I., and A.C.; investigation, F.D.B., F.O., M. Prelli, E.C., R. Albano, G.R., J.E., A.D.A., M. Panero, C.D.A., L.C., M.C., F.P., A.B., and P.L.P.; resources, A.M., P.Z., R. Altieri, I.M., P.C., D.G., S.P., G.F., P.L.P.; writing–original draft, F.D.B., F.O., and C.B.; writing–review & editing, F.D.B., F.O., G.C., C.I., P.L.P., and C.B. Supervision and funding acquisition, P.M.C. and C.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 16, 2022
Revised: April 5, 2023
Accepted: June 30, 2023
Published: July 27, 2023

REFERENCES

1. Low, J.T., Ostrom, Q.T., Cioffi, G., Neff, C., Waite, K.A., Kruchko, C., and Barnholtz-Sloan, J.S. (2022). Primary brain and other central nervous system tumors in the United States (2014-2018): A summary of the CBTRUS statistical report for clinicians. Neurooncol. Pract. *9*, 165–182. https://doi.org/10.1093/nop/npac015.
2. Sottoriva, A., Spiteri, I., Piccirillo, S.G.M., Touloumis, A., Collins, V.P., Marioni, J.C., Curtis, C., Watts, C., and Tavaré, S. (2013). Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc. Natl. Acad. Sci. USA *110*, 4009–4014. https://doi.org/10.1073/pnas.1219747110.
3. Kim, J., Lee, I.H., Cho, H.J., Park, C.K., Jung, Y.S., Kim, Y., Nam, S.H., Kim, B.S., Johnson, M.D., Kong, D.S., et al. (2015). Spatiotemporal Evolution of the Primary Glioblastoma Genome. Cancer Cell *28*, 318–328. https://doi.org/10.1016/j.ccell.2015.07.013.
4. Nicholson, J.G., and Fine, H.A. (2021). Diffuse Glioma Heterogeneity and Its Therapeutic Implications. Cancer Discov. *11*, 575–590. https://doi.org/10.1158/2159-8290.CD-20-1474.
5. Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al. (2019). An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. Cell *178*, 835–849.e21. https://doi.org/10.1016/j.cell.2019.06.024.
6. Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401. https://doi.org/10.1126/science.1254257.
7. Snuderl, M., Fazlollahi, L., Le, L.P., Nitta, M., Zhelyazkova, B.H., Davidson, C.J., Akhavanfard, S., Cahill, D.P., Aldape, K.D., Betensky, R.A., et al. (2011). Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. Cancer Cell *20*, 810–817. https://doi.org/10.1016/j.ccr.2011.11.005.
8. Little, S.E., Popov, S., Jury, A., Bax, D.A., Doey, L., Al-Sarraj, S., Jurgensmeier, J.M., and Jones, C. (2012). Receptor tyrosine kinase genes amplified in glioblastoma exhibit a mutual exclusivity in variable proportions reflective of individual tumor heterogeneity. Cancer Res. *72*, 1614–1620. https://doi.org/10.1158/0008-5472.CAN-11-4069.
9. Szerlip, N.J., Pedraza, A., Chakravarty, D., Azim, M., McGuire, J., Fang, Y., Ozawa, T., Holland, E.C., Huse, J.T., Jhanwar, S., et al. (2012). Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. Proc. Natl. Acad. Sci. USA *109*, 3041–3046. https://doi.org/10.1073/pnas.1114033109.
10. Wang, J., Cazzato, E., Ladewig, E., Frattini, V., Rosenbloom, D.I.S., Zairis, S., Abate, F., Liu, Z., Elliott, O., Shin, Y.J., et al. (2016). Clonal evolution of glioblastoma under therapy. Nat. Genet. *48*, 768–776. https://doi.org/10.1038/ng.3590.
11. Orzan, F., De Bacco, F., Crisafulli, G., Pellegatta, S., Mussolin, B., Siravegna, G., D'Ambrosio, A., Comoglio, P.M., Finocchiaro, G., and Boccaccio, C. (2017). Genetic Evolution of Glioblastoma Stem-Like Cells From Primary to Recurrent Tumor. Stem Cell. *35*, 2218–2228. https://doi.org/10.1002/stem.2703.
12. Barthel, F.P., Johnson, K.C., Varn, F.S., Moskalik, A.D., Tanner, G., Kocakavuk, E., Anderson, K.J., Abiola, O., Aldape, K., Alfaro, K.D., et al. (2019). Longitudinal molecular trajectories of diffuse glioma in adults. Nature *576*, 112–120. https://doi.org/10.1038/s41586-019-1775-1.
13. Touat, M., Li, Y.Y., Boynton, A.N., Spurr, L.F., Iorgulescu, J.B., Bohrson, C.L., Cortes-Ciriano, I., Birzu, C., Geduldig, J.E., Pelton, K., et al. (2020). Mechanisms and therapeutic implications of hypermutation in gliomas. Nature *580*, 517–523. https://doi.org/10.1038/s41586-020-2209-9.
14. Körber, V., Yang, J., Barah, P., Wu, Y., Stichel, D., Gu, Z., Fletcher, M.N.C., Jones, D., Hentschel, B., Lamszus, K., et al. (2019). Evolutionary Trajectories of IDH^{wt} Glioblastomas Reveal a Common Path of Early Tumorigenesis Instigated Years ahed of Initial Diagnosis. Cancer Cell *35*, 692–704.e612. https://doi.org/10.1016/j.ccell.2019.02.007.
15. Mitchell, K., Troike, K., Silver, D.J., and Lathia, J.D. (2021). The evolution of the cancer stem cell state in glioblastoma: emerging insights into the next generation of functional interactions. Neuro Oncol. *23*, 199–213. https://doi.org/10.1093/neuonc/noaa259.
16. Gimple, R.C., Yang, K., Halbert, M.E., Agnihotri, S., and Rich, J.N. (2022). Brain cancer stem cells: resilience through adaptive plasticity and hierarchical heterogeneity. Nat. Rev. Cancer *22*, 497–514. https://doi.org/10.1038/s41568-022-00486-x.
17. Greaves, M. (2015). Evolutionary determinants of cancer. Cancer Discov. *5*, 806–820. https://doi.org/10.1158/2159-8290.CD-15-0439.
18. Lee, J.H., Lee, J.E., Kahng, J.Y., Kim, S.H., Park, J.S., Yoon, S.J., Um, J.Y., Kim, W.K., Lee, J.K., Park, J., et al. (2018). Human glioblastoma arises from subventricular zone cells with low-level driver mutations. Nature *560*, 243–247. https://doi.org/10.1038/s41586-018-0389-3.
19. Vescovi, A.L., Galli, R., and Reynolds, B.A. (2006). Brain tumour stem cells. Nat. Rev. Cancer *6*, 425–436. https://doi.org/10.1038/nrc1889.
20. Friedmann-Morvinski, D., Bushong, E.A., Ke, E., Soda, Y., Marumoto, T., Singer, O., Ellisman, M.H., and Verma, I.M. (2012). Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. Science *338*, 1080–1084. https://doi.org/10.1126/science.1226929.
21. Suvà, M.L., Rheinbay, E., Gillespie, S.M., Patel, A.P., Wakimoto, H., Rabkin, S.D., Riggi, N., Chi, A.S., Cahill, D.P., Nahed, B.V., et al. (2014). Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. Cell *157*, 580–594. https://doi.org/10.1016/j.cell.2014.02.030.
22. Lee, J., Kotliarova, S., Kotliarov, Y., Li, A., Su, Q., Donin, N.M., Pastorino, S., Purow, B.W., Christopher, N., Zhang, W., et al. (2006). Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. Cancer Cell *9*, 391–403. https://doi.org/10.1016/j.ccr.2006.03.030.
23. Galli, R., Binda, E., Orfanelli, U., Cipelletti, B., Gritti, A., De Vitis, S., Fiocco, R., Foroni, C., Dimeco, F., and Vescovi, A. (2004). Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. Cancer Res. *64*, 7011–7021. https://doi.org/10.1158/0008-5472.CAN-04-1364.
24. Brennan, C.W., Verhaak, R.G.W., McKenna, A., Campos, B., Noushmehr, H., Salama, S.R., Zheng, S., Chakravarty, D., Sanborn, J.Z., Berman, S.H., et al. (2013). The somatic genomic landscape of glioblastoma. Cell *155*, 462–477. https://doi.org/10.1016/j.cell.2013.09.034.
25. Orzan, F., Pagani, F., Cominelli, M., Triggiani, L., Calza, S., De Bacco, F., Medicina, D., Balzarini, P., Panciani, P.P., Liserre, R., et al. (2020). A simplified integrated molecular and immunohistochemistry-based algorithm allows high accuracy prediction of glioblastoma transcriptional subtypes. Lab. Invest. *100*, 1330–1344. https://doi.org/10.1038/s41374-020-0437-0.
26. Behnan, J., Stangeland, B., Langella, T., Finocchiaro, G., Murrell, W., and Brinchmann, J.E. (2016). Ultrasonic Surgical Aspirate is a Reliable Source For Culturing Glioblastoma Stem Cells. Sci. Rep. *6*, 32788. https://doi.org/10.1038/srep32788.
27. Vartholomatos, E., Vartholomatos, G., Alexiou, G.A., and Markopoulos, G.S. (2021). The Past, Present and Future of Flow Cytometry in Central

Nervous System Malignancies. Methods Protoc. *4*, 11. https://doi.org/10.3390/mps4010011.

28. Kim, H., Nguyen, N.P., Turner, K., Wu, S., Gujar, A.D., Luebeck, J., Liu, J., Deshpande, V., Rajkumar, U., Namburi, S., et al. (2020). Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. Nat. Genet. *52*, 891–897. https://doi.org/10.1038/s41588-020-0678-2.

29. Prager, B.C., Bhargava, S., Mahadev, V., Hubert, C.G., and Rich, J.N. (2020). Glioblastoma Stem Cells: Driving Resilience through Chaos. Trends Cancer *6*, 223–235. https://doi.org/10.1016/j.trecan.2020.01.009.

30. Mazzoleni, S., Politi, L.S., Pala, M., Cominelli, M., Franzin, A., Sergi Sergi, L., Falini, A., De Palma, M., Bulfone, A., Poliani, P.L., and Galli, R. (2010). Epidermal growth factor receptor expression identifies functionally and molecularly distinct tumor-initiating cells in human glioblastoma multiforme and is required for gliomagenesis. Cancer Res. *70*, 7500–7513. https://doi.org/10.1158/0008-5472.CAN-10-2353.

31. Li, Y., Li, A., Glas, M., Lal, B., Ying, M., Sang, Y., Xia, S., Trageser, D., Guerrero-Cázares, H., Eberhart, C.G., et al. (2011). c-Met signaling induces a reprogramming network and supports the glioblastoma stem-like phenotype. Proc. Natl. Acad. Sci. USA *108*, 9951–9956. https://doi.org/10.1073/pnas.1016912108.

32. Schulte, A., Günther, H.S., Martens, T., Zapf, S., Riethdorf, S., Wülfing, C., Stoupiec, M., Westphal, M., and Lamszus, K. (2012). Glioblastoma stem-like cell lines with either maintenance or loss of high-level EGFR amplification, generated via modulation of ligand concentration. Clin. Cancer Res. *18*, 1901–1913. https://doi.org/10.1158/1078-0432.CCR-11-3084.

33. De Bacco, F., Casanova, E., Medico, E., Pellegatta, S., Orzan, F., Albano, R., Luraghi, P., Reato, G., D'Ambrosio, A., Porrati, P., et al. (2012). The MET Oncogene Is a Functional Marker of a Glioblastoma Stem Cell Subtype. Cancer Res. *72*, 4537–4550. https://doi.org/10.1158/0008-5472.

34. Verginelli, F., Pisacane, A., Gambardella, G., D'Ambrosio, A., Candiello, E., Ferrio, M., Panero, M., Casorzo, L., Benvenuti, S., Cascardi, E., et al. (2021). Cancer of unknown primary stem-like cells model multi-organ metastasis and unveil liability to MEK inhibition. Nat. Commun. *12*, 2498. https://doi.org/10.1038/s41467-021-22643-w.

35. deCarvalho, A.C., Kim, H., Poisson, L.M., Winn, M.E., Mueller, C., Cherba, D., Koeman, J., Seth, S., Protopopov, A., Felicella, M., et al. (2018). Discordant inheritance of chromosomal and extrachromosomal DNA elements contributes to dynamic disease evolution in glioblastoma. Nat. Genet. *50*, 708–717. https://doi.org/10.1038/s41588-018-0105-0.

36. Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., Bernard, L., Viale, G., Pelicci, P.G., and Di Fiore, P.P. (2010). Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. Cell *140*, 62–73. https://doi.org/10.1016/j.cell.2009.12.007.

37. Yi, E., Chamorro González, R., Henssen, A.G., and Verhaak, R.G.W. (2022). Extrachromosomal DNA amplifications in cancer. Nat. Rev. Genet. *23*, 760–771. https://doi.org/10.1038/s41576-022-00521-5.

38. Wang, Q., Hu, B., Hu, X., Kim, H., Squatrito, M., Scarpace, L., deCarvalho, A.C., Lyu, S., Li, P., Li, Y., et al. (2017). Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. Cancer Cell *32*, 42–56.e6. https://doi.org/10.1016/j.ccell.2017.06.003.

39. Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., et al. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98–110. https://doi.org/10.1016/j.ccr.2009.12.020.

40. Lu, X., Maturi, N.P., Jarvius, M., Yildirim, I., Dang, Y., Zhao, L., Xie, Y., Tan, E.J., Xing, P., Larsson, R., et al. (2022). Cell-lineage controlled epigenetic regulation in glioblastoma stem cells determines functionally distinct subgroups and predicts patient survival. Nat. Commun. *13*, 2236. https://doi.org/10.1038/s41467-022-29912-2.

41. Riddick, G., and Fine, H.A. (2011). Integration and analysis of genome-scale data from gliomas. Nat. Rev. Neurol. *7*, 439–450. https://doi.org/10.1038/nrneurol.2011.100.

42. Suwala, A.K., Stichel, D., Schrimpf, D., Maas, S.L.N., Sill, M., Dohmen, H., Banan, R., Reinhardt, A., Sievers, P., Hinz, F., et al. (2021). Glioblastomas with primitive neuronal component harbor a distinct methylation and copy-number profile with inactivation of TP53, PTEN, and RB1. Acta Neuropathol. *142*, 179–189. https://doi.org/10.1007/s00401-021-02302-6.

43. Perry, A., Miller, C.R., Gujrati, M., Scheithauer, B.W., Zambrano, S.C., Jost, S.C., Raghavan, R., Qian, J., Cochran, E.J., Huse, J.T., et al. (2009). Malignant gliomas with primitive neuroectodermal tumor-like components: a clinicopathologic and genetic study of 53 cases. Brain Pathol. *19*, 81–90. https://doi.org/10.1111/j.1750-3639.2008.00167.x.

44. Louis, D.N., Perry, A., Wesseling, P., Brat, D.J., Cree, I.A., Figarella-Branger, D., Hawkins, C., Ng, H.K., Pfister, S.M., Reifenberger, G., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. *23*, 1231–1251. https://doi.org/10.1093/neuonc/noab106.

45. Wang, L.B., Karpova, A., Gritsenko, M.A., Kyle, J.E., Cao, S., Li, Y., Rykunov, D., Colaprico, A., Rothstein, J.H., Hong, R., et al. (2021). Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell *39*, 509–528.e20. https://doi.org/10.1016/j.ccell.2021.01.006.

46. Ming, G.L., and Song, H. (2011). Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron *70*, 687–702. https://doi.org/10.1016/j.neuron.2011.05.001.

47. Turner, K.M., Deshpande, V., Beyter, D., Koga, T., Rusert, J., Lee, C., Li, B., Arden, K., Ren, B., Nathanson, D.A., et al. (2017). Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature *543*, 122–125. https://doi.org/10.1038/nature21356.

48. Hoogstrate, Y., Draaisma, K., Ghisai, S.A., van Hijfte, L., Barin, N., de Heer, I., Coppieters, W., van den Bosch, T.P.P., Bolleboom, A., Gao, Z., et al. (2023). Transcriptome analysis reveals tumor microenvironment changes in glioblastoma. Cancer Cell *41*, 678–692.e7. https://doi.org/10.1016/j.ccell.2023.02.019.

49. Osuka, S., and Van Meir, E.G. (2017). Overcoming therapeutic resistance in glioblastoma: the way forward. J. Clin. Invest. *127*, 415–426. https://doi.org/10.1172/JCI89587.

50. Balanis, N.G., Sheu, K.M., Esedebe, F.N., Patel, S.J., Smith, B.A., Park, J.W., Alhani, S., Gomperts, B.N., Huang, J., Witte, O.N., and Graeber, T.G. (2019). Pan-cancer Convergence to a Small-Cell Neuroendocrine Phenotype that Shares Susceptibilities with Hematological Malignancies. Cancer Cell *36*, 17–34.e7. https://doi.org/10.1016/j.ccell.2019.06.005.

51. Chkheidze, R., Raisanen, J., Gagan, J., Richardson, T.E., Pinho, M.C., Raj, K., Achilleos, M., Slepicka, C., White, C.L., Evers, B.M., et al. (2021). Alterations in the RB Pathway With Inactivation of RB1 Characterize Glioblastomas With a Primitive Neuronal Component. J. Neuropathol. Exp. Neurol. *80*, 1092–1098. https://doi.org/10.1093/jnen/nlab109.

52. Prat, M., Crepaldi, T., Pennacchietti, S., Bussolino, F., and Comoglio, P.M. (1998). Agonistic monoclonal antibodies against the Met receptor dissect the biological responses to HGF. J. Cell Sci. *111*, 237–247. https://doi.org/10.1242/jcs.111.2.237.

53. Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. *2*, 401–404. https://doi.org/10.1158/2159-8290.CD-12-0095.

54. Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., et al. (2010). Towards a knowledge-based Human Protein Atlas. Nat. Biotechnol. *28*, 1248–1250. https://doi.org/10.1038/nbt1210-1248.

55. Bowman, R.L., Wang, Q., Carro, A., Verhaak, R.G.W., and Squatrito, M. (2017). GlioVis data portal for visualization and analysis of brain tumor expression datasets. Neuro Oncol. *19*, 139–141. https://doi.org/10.1093/neuonc/now247.

56. Baralis, E., Bertotti, A., Fiori, A., and Grand, A. (2012). LAS: a software platform to support oncological data management. J. Med. Syst. *36* (*Suppl 1*), S81–S90. https://doi.org/10.1007/s10916-012-9891-6.

57. Fu, L., and Medico, E. (2007). FLAME, a novel fuzzy clustering method for the analysis of DNA microarray data. BMC Bioinf. *8*, 3. https://doi.org/10.1186/1471-2105-8-3.

58. Steinhaus, R., Proft, S., Schuelke, M., Cooper, D.N., Schwarz, J.M., and Seelow, D. (2021). MutationTaster2021. Nucleic Acids Res. *49*, W446–W451. https://doi.org/10.1093/nar/gkab266.

59. Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J. Immunol. Methods *347*, 70–78. https://doi.org/10.1016/j.jim.2009.06.008.

60. Dobin, A., and Gingeras, T.R. (2015). Mapping RNA-seq Reads with STAR. Curr. Protoc. Bioinformatics *51*, 11–14. https://doi.org/10.1002/0471250953.bi1114s51.

61. Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., et al. (2021). GENCODE 2021. Nucleic Acids Res. *49*, D916–D923. https://doi.org/10.1093/nar/gkaa1087.

62. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

63. Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106. https://doi.org/10.1186/gb-2010-11-10-r106.

64. Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinf. *14*, 128. https://doi.org/10.1186/1471-2105-14-128.

65. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90–W97. https://doi.org/10.1093/nar/gkw377.

66. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. https://doi.org/10.1038/ncomms14049.

67. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

68. McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst. *8*, 329–337.e4. https://doi.org/10.1016/j.cels.2019.03.003.

69. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

70. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

71. De Bacco, F., D'Ambrosio, A., Casanova, E., Orzan, F., Neggia, R., Albano, R., Verginelli, F., Cominelli, M., Poliani, P.L., Luraghi, P., et al. (2016). MET inhibition overcomes radiation resistance of glioblastoma stem-like cells. EMBO Mol. Med. *8*, 550–568. https://doi.org/10.15252/emmm.201505890.

72. Corti, G., Bartolini, A., Crisafulli, G., Novara, L., Rospo, G., Montone, M., Negrino, C., Mussolin, B., Buscarino, M., Isella, C., et al. (2019). A Genomic Analysis Workflow for Colorectal Cancer Precision Oncology. Clin. Colorectal Cancer *18*, 91–101.e3. https://doi.org/10.1016/j.clcc.2019.02.008.

73. Crisafulli, G., Mussolin, B., Cassingena, A., Montone, M., Bartolini, A., Barault, L., Martinetti, A., Morano, F., Pietrantonio, F., Sartore-Bianchi, A., et al. (2019). Whole exome sequencing analysis of urine trans-renal tumour DNA in metastatic colorectal cancer patients. ESMO Open *4*, e000572. https://doi.org/10.1136/esmoopen-2019-000572.

74. Crisafulli, G., Sartore-Bianchi, A., Lazzari, L., Pietrantonio, F., Amatu, A., Macagno, M., Barault, L., Cassingena, A., Bartolini, A., Luraghi, P., et al. (2022). Temozolomide Treatment Alters Mismatch Repair and Boosts Mutational Burden in Tumor and Blood of Colorectal Cancer Patients. Cancer Discov. *12*, 1656–1675. https://doi.org/10.1158/2159-8290.CD-21-1434.

75. Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowicz, M.L., Kropiwnicki, E., Jagodnik, K.M., et al. (2021). Gene Set Knowledge Discovery with Enrichr. Curr. Protoc. *1*, e90. https://doi.org/10.1002/cpz1.90.

76. De Bacco, F., Orzan, F., Erriquez, J., Casanova, E., Barault, L., Albano, R., D'Ambrosio, A., Bigatto, V., Reato, G., Patanè, M., et al. (2021). ERBB3 overexpression due to miR-205 inactivation confers sensitivity to FGF, metabolic activation, and liability to ERBB3 targeting in glioblastoma. Cell Rep. *36*, 109455. https://doi.org/10.1016/j.celrep.2021.109455.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-EGFR antibodies (clone E30)	Dako Cytomation	cat#M7239; RRID:AB_2721108 (IHC 1:100, 1h)
anti-EGFR-PE (EGFR.1) antibodies	Immunological Science	cat#MAB-9680PE (FC 1:20)
anti-PDGFRA-PE (16A1) antibodies	Biolegend	cat#323506; RRID:AB_2268113 (FC 1:10)
anti-PDGFRA antibodies	Thermo Fisher Scientific	cat#TA804956; RRID:AB_2627546 (IHC 1:50, O/N)
anti-HGF R/MET-PE-Cy7 (95106) antibodies	R&D System	cat#MAB3582; RRID:AB_884334 (FC 1:20)
anti-MET C-12 antibodies	SantaCruz Biotechnology	cat#sc-10; RRID:AB_631940 (IHC 1:50, O/N)
anti-MET DL21 antibodies	Prat et al.[52]	N/A (WB 1:1500)
anti-Synaptophisin (D35E4) antibodies	Cell Signaling Technology	cat#5461; RRID:AB_10698743 (WB 1:1000)
anti- Synaptophisin (clone SP11) antibodies	Thermo Fisher Scientific	cat#MA5-14532; RRID:AB_10983675 (IHC 1:100, 1h)
anti-BTG2 antibodies	Atlas Antibodies	cat#HPA002355; RRID:AB_1078304 (WB 1:1000)
anti-β3-tubulin (TU-20) antibodies	Cell Signaling Technology	cat#4466; RRID:AB_1904176 (WB 1:1000)
Alexa Fluor® 488 anti-β3-tubulin antibodies	BD Biosciences	cat#560381; RRID:AB_1645344 (IF 1:50)
anti-NeuN (14H6L24) antibodies	Thermo Fisher Scientific	cat#702022; RRID:AB_2633050 (WB 1:250)
anti-ASCL1 (E6Y1B) antibodies	Cell Signaling Technology	cat#55467; RRID:AB_2936422 (WB 1:1000)
anti-Doublecortin (EPR19997) antibodies	Abcam	cat#ab207175; RRID:AB_2894710 (WB 1:500)
anti-GFAP (N-18) antibodies	Santa Cruz Biotechnology	cat#sc-6171; RRID:AB_641023 (WB 1:200, IF 1:50)
anti-GFAP (clone 6F2) antibodies	Dako Cytomation	cat#M0721 (IHC 1:150, 1h)
anti-Vimentin (VIM 3B4) antibodies	Millipore	cat#CBL202; RRID:AB_93387 (WB 1:500)
anti-YAP1 antibodies	Proteintech	cat#13584-1-AP; RRID:AB_2218915 (WB 1:2000)
anti-GalC (H300) antibodies	Santa Cruz Biotechnology	cat#sc-67352; RRID:AB_2108531 (IF 1:50)
anti-CD44 (8E2) antibodies	Cell Signaling Technology	cat#5640; RRID:AB_10547133 (WB 1:1000)
anti-CD44-FITC (MEM233) antibodies	Immunological Sciences	cat#MAB-1141F (FC 1:20)
anti-CD24-FITC (SN3) antibodies	Immunological Sciences	cat#MAB-1417F (FC 1:20)
anti-CD45 PE-CF594 (HI30) antibodies	BD Horizon	cat#562312; RRID:AB_11154590 (FC 1:20)
anti-CD56 PE-Cy7 (N901) antibodies	Beckman Coulter	cat#A21692; RRID:AB_2892144 (FC 1:20)
anti-Vinculin antibodies	Sigma-Aldrich	cat#V9131; RRID:AB_477629 (WB 1:200)
Chemicals, peptides, and recombinant proteins		
FGF2	PeproTech	cat#100-18B
EGF	Sigma-Aldrich	cat#E1257
PDGFBB	PeproTech	cat#100-14B
HGF	PeproTech	cat#100-39H
B27 PLUS supplement	Thermo Fisher Scientific	cat#17504044
Critical commercial assays		
All Prep DNA/RNA Mini kit	Qiagen	cat#80284
RNeasy Mini kit	Qiagen	cat#74004
Maxwell® RSC miRNA Tissue kit	Promega	cat#AS4500
ReliaPrep™ gDNA Tissue Miniprep System	Promega	cat#A2051
PowerPlex® 16 HS System	Promega	cat#DC2101
Platinum® Taq Hot-Start DNA Polymerase	Thermo Fisher Scientific	cat#14966001

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
2x Phanta Max Master Mix	Vazyme	cat#P515-01
BigDye Terminator v3.1 Cycle Sequencing kit	Thermo Fisher Scientific	cat#4337455
CleanSEQ Dye-Terminator Removal Kit	Beckman Coulter	cat#A29151
TaqMan™ Universal PCR Master Mix	Thermo Fisher Scientific	cat#4304437
ddPCR Supermix for Probes (No dUTP)	Bio-Rad	cat#1863023
RNA 6000 Nano kit	Agilent	cat#5067-1511
Qubit™ RNA BR Assay kit	Thermo Fisher Scientific	cat#Q10210
SureSelectXT HumanAllExon V5+UTRs	Agilent	cat#5190-6213
TruSeq Stranded mRNA Library Prep	Illumina	cat#20020595
Agilent High Sensitivity DNA Kit	Agilent	cat#5067-4626
Qubit™ dsDNA HS kit	Thermo Fisher Scientific	cat#Q32851
IDT for Illumina TruSeq RNA UN Indexes	Illumina	cat#20022371
MiSeq Reagent Kit v2	Illumina	cat#MS-102-2002
NVSEQ 6000 S1 Rgt Kit v1.5	Illumina	cat#20028318
MICROTUBE-15 AFA BEADS SCREW-CAP (25)	COVARIS	cat#COV520045
Agencourt AMPure XP - 60 mL	Beckman Coulter	cat#A63881
NXSEQ AMPFREE LOW DNA LIBRARY KIT 12RXN/48RXN	Lucigen	cat#14000-1/cat#14000-2
xGen™ Stubby Adapter and UDI Primer Pairs Index 1-96	IDT	cat#10005924/cat#10005921
XGEN HYBRIDIZATION AND WASH KIT	IDT	cat#1080577
xGen® Universal Blockers - TS Mix	IDT	cat#1075474
XGEN LIBRARY AMPLIFICATION PRIMER MIX	IDT	cat#1077675
xGen Custom Hyb Panel-Accel	IDT	N/A
KAPA HIFI HS RM (6.25ML)	Roche	cat#7958935001
NSQ 500/550 Mid Output KT v2.5 (300 CYS)	Illumina	cat#20024905
MiSeq® Reagent Kit v3 (600 cycle)	Illumina	cat#MS-102-3003
Chromium Next GEM Single Cell 3′ Kit v3.1	10× Genomics	cat#1000268
Chromium Next GEM Chip G Single Cell Kit	10× Genomics	cat#1000127
NovaSeq 6000 SP Reagent Kit (200 cycles)	Illumina	cat#20040326
RNAScope Duplex assay	Advanced Cell Diagnostics	cat#322500
Annexin V Apoptosis Detection Kit	Becton Dickinson	cat#556547
CellTiter-Glo®	Promega	cat#G7570
CN probe: CDKN2A	Thermo Fisher Scientific	cat#Hs03714372_cn
CN probe: CDKN2B	Thermo Fisher Scientific	cat#Hs02900430_cn
CN probe: PTEN	Thermo Fisher Scientific	cat#Hs02599450_cn
CN probe: EGFR	Thermo Fisher Scientific	cat#Hs.PT.58.27649789.g
CN probe: PDGFRA	Thermo Fisher Scientific	cat#Hs05935655_cn
CN probe: MET	Thermo Fisher Scientific	cat#Hs04993403_cn
CN probe: FGFR3	Thermo Fisher Scientific	cat#Hs00116878_cn
CN probe: HGF	Thermo Fisher Scientific	cat#Hs02789622_cn
CN probe: MYC	Thermo Fisher Scientific	cat#Hs00834648_cn
CN probe: NFKBIA	IDT	cat#Hs.PT.58.45537430.g
CN probe: MDM4	IDT	cat#Hs.PT.58.3793007.g
CN probe: CDK4	Thermo Fisher Scientific	cat#Hs01071103_cn
CN probe: PIK3CA	Thermo Fisher Scientific	cat#Hs00923687_cn
CN probe: RNAseP	Thermo Fisher Scientific	cat#4403328
CN probe: GREB	Thermo Fisher Scientific	cat#Hs01738470_cn
CN probe: APOA1	IDT	cat#Hs.PT.56A.40574746.g

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FISH probe: Vysis EGFR/CEP 7 FISH Probe Kit	Abbott Molecular	cat#1N35-20
FISH probe: ZytoLight SPEC MYC Dual Color BreakApart Probe	Zytovision	cat#Z-2090-50
ddPCR probe: TERT C228T_113	Bio-Rad	cat#dHsaEXD72405942
ddPCE probe: TERT C250T_113	Bio-Rad	cat#dHsaEXD46675715
gene expression probe: EGFR	Thermo Fisher Scientific	cat#Hs01076078_m1
gene expression probe: PDGFRA	Thermo Fisher Scientific	cat#Hs00183486_m1
gene expression probe: MET	Thermo Fisher Scientific	cat#Hs01565584_m1
gene expression probe: MYC	Thermo Fisher Scientific	cat#Hs00300643_m1
In situ RNA hybridization probe: Hs-CD44-C2	Advanced Cell Diagnostics	cat#311271-C2
In situ RNA hybridization probe: Hs-VIM	Advanced Cell Diagnostics	cat#310441
In situ RNA hybridization probe: Hs-VIM-C2	Advanced Cell Diagnostics	cat#310441-C2
In situ RNA hybridization probe: Hs-S100A11	Advanced Cell Diagnostics	cat#400861
In situ RNA hybridization probe: Hs-DCX-C2	Advanced Cell Diagnostics	cat#489551-C2
In situ RNA hybridization probe: Hs-CD24	Advanced Cell Diagnostics	cat#313021
In situ RNA hybridization probe: Hs-STMN4	Advanced Cell Diagnostics	cat#850151
Deposited data		
WES and NGS panel datasets	This paper	ENA: PRJEB55406, study ERP140298
RNAseq (bulk and single-cell) datasets	This paper	ENA: PRJEB55406, study ERP140298
Bulk RNAseq FeatureCounts (Figures 5 and 7B–7D)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
Single-cell RNAseq log transformed normalized counts (Figure 7G)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
STR analysis cohort 2 (GBMs and GSCs)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
IHC score in validation cohort (Figure 1G)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
Raw data for stem cell frequency (Figure 6A)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
Raw data for western blot (Figure 7F)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
Raw data for quantification of *in situ* RNA hybridization (Figures 7I and 7J)	This paper, Mendeley data	https://data.mendeley.com/datasets/n5tvhwpvg5/draft
Experimental models: Cell lines		
Human GBM tissue (cohort 2)	Città della Salute e Scienza	https://www.cittadellasalute.to.it/
Human GBM tissue (validation cohort)	Spedali Civili di Brescia	Orzan et al.[25]
A549 (lung cancer)	ATCC	cat#ATCC® CCL-185
A2780 (ovarian cancer)	ECACC	cat#93112519-1VL
SNB19 (GBM)	NCI-60 cancer panel	RRID:CVCL_0535
U87MG (GBM)	ATCC	cat#ATCC® HTB-14
U251 (GBM)	NCI-60 cancer panel	RRID:CVCL_0021
UACC257 (melanoma)	NCI-60 cancer panel	RRID:CVCL_1779
NHA (normal human astrocyte)	Lonza Bioscience	cat#CC-2565
SVG (immortalized astrocyte)	ATCC	cat#ATCC® CRL8621
Experimental models: Organisms/strains		
NOD.CB17-Prkdcscid/J	Charles River Laboratories	https://www.criver.com/products-services/find-model/jax-nod-scid-mice?region=27
Oligonucleotides		
Primers (see Table S5)	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
cBioPortal site for Cancer Genomics	Cerami et al.[53]	http://www.cbioportal.org/
The Human Protein Atlas	Uhlen et al.[54]	http://www.proteinatlas.org
Gliovis data portal	Bowman et al.[55]	http://gliovis.bioinfo.cnio.es/
GraphPad Prism 8.3.0	GraphPad software	www.graphpad.com
Laboratory Assistance Suite (LAS)	Baralis et al.[56]	http://las.ircc.it/las/laslogin/
Chromas Lite 2.01 software	Technelysium	http://www.technelysium.com.au/chromas_lite.html
COSMIC	Sanger Institute	http://cancer.sanger.ac.uk/cosmic
GEDAS software	Fu and Medico[57]	http://spurceforge.net/projects/gedas
MutationTaster2021	Steinhaus et al.[58]	https://www.genecascade.org/MutationTaster2021/
CytoVision® software	Olympus	https://www.selectscience.net/products/cytovision/
Summit 4.3 software	Beckman Coulter	PN998403C
ELDA: Extreme Limiting Dilution Analysis	Hu et al.[59]	http://bioinf.wehi.edu.au/software/elda/
R	R-Project for Stat computing	https://www.r-project.org/
bcl2fastq	Illumina	V2.20.0.422
STAR 2.5.465	Dobin et al.[60]	https://github.com/alexdobin/STAR
GENCODE 2021	Frankish et al.[61]	https://dspace.mit.edu/handle/1721.1/132262.2
Subread/featureCounts v.1.6.3	Liao et al.[62]	https://subread.sourceforge.net/
DESeq2 1.38.1	Anders et al.[63]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
enrichR	Chen et al.[64]; Kuleshov et al.[65]	https://maayanlab.cloud/Enrichr/
CellRanger 7.1.0	Zheng et al.[66]	https://10xgenomics.com
SEURAT v.4.3.0	Hao et al.[67]	https://satijalab.org/seurat/
Doublet Finder v.2.0.3	McGinnis et al.[68]	https://github.com/chris-mcginnis-ucsf/DoubletFinder
BWA-mem	Li and Durbin[69]	https://github.com/lh3/bwa
SAMtools 1.9	Danecek et al.[70]	http://www.htslib.org/
Image Lab software	Bio-Rad	cat#17006130
ImageJ software	NIH	https://ImageJ.nih.gov

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Carla Boccaccio (carla.boccaccio@ircc.it).

Materials availability

Neurospheres are available from the lead contact upon completed material transfer agreement after request by qualified academic investigators for non-commercial purposes.

Data and code availability

- Human Next Generation Sequencing data (DNA sequencing, and bulk and single-cell RNA sequencing) have been deposited at the European Nucleotide Archive (ENA) of European Bioinformatics Institute (EBI) under project accession number PRJEB55406, study ERP140298 (https://www.ebi.ac.uk/ena/browser/search) and are publicly available as of the date of publication. The accession number is listed in the key resources table. Bulk RNAseq FeatureCounts, single-cell RNAseq log transformed normalized counts, STR profiling and SNP ID of GSC families, IHC score in validation cohort, original western blots, limiting dilution assay data, and *in situ* RNA hybridization images have been deposited at Mendeley, and are publicly available as of the date of publication. STR profiling of GBM14, 17, 18 and 20 GSC families and SNP ID of GBM14 family are also available in Table S4.

- No original code is reported in this study.
- Any additional information required to reanalyze the data reported in this paper is available upon request from the lead contact.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human subjects

Glioblastoma samples were obtained from patients recruited at Fondazione IRCCS Istituto Neurologico C. Besta (Milan, Italy), or at Città della Salute e della Scienza (University of Torino, Italy), according to protocols approved by the respective institutional Ethical Committees. Informed written consent was gathered from all patients and studies were conducted according to the Declaration of Helsinki. All patient data and samples were de-identified before processing and key information (sex, age and diagnosis) is reported in Table S1 (cohort 1) and Table S2 (cohort 2). Briefly, patients were all adults (median age = 61.5, cohort 1, and = 67, cohort 2), and M:F ratio was 2.3 in cohort 1 (29:69) and 1.8 in cohort 2 (11:20). Sex and age were not considered as relevant factors in this study. Glioblastoma samples in the validation cohort were obtained from the Archive of Pathological Department of Spedali Civili of Brescia. Their use was approved by the Ethics Board of Spedali Civili of Brescia. Clinical characteristics of patients in this cohort were previously reported.[25] For the TCGA GBM cohort: (i) genetic data for correlation studies were obtained from the public cBioPortal website (Glioblastoma Multiforme TCGA Firehose Legacy, July 2022, https://www.cbioportal.org/)[53]; (ii) immunohistochemical evaluation of EGF, FGF2, HGF and PDGFB in public tissues were obtained from the Human Protein Atlas (http://www.proteinatlas.org).[54]

Neurosphere (NS) derivation

Conventional NSs (cohort 1) were derived starting from surgical samples (biopsies and, in a small percentage of cases, ultrasonic aspirate, UA, see Table S1 for details) of consecutive primary GBMs obtained at Fondazione IRCCS Istituto Neurologico C. Besta (Milan, Italy), as previously described.[33] NSs were propagated at clonal density in standard medium containing human FGF2 (20 ng/mL, PeproTech) and EGF (20 ng/mL, Sigma-Aldrich).

NS family members (cohort 2) were derived starting from UA of primary GBMs obtained at Città della Salute e della Scienza (University of Torino, Italy). Briefly, the fresh fluid obtained with the UA was collected in conical tubes, centrifuged 10 min at 270 g, RT and red blood cells were lysed by treatment with ACK for 15 min. The pellet was incubated with collagenase type 1 at 37°C for 20 min, mechanically dissociated by up and down pipetting with a sterile 18G needle in a 1 mL syringe and subjected to sequential 100 μm and 70 μm filtrations. Cellular resuspension (containing cancer cells, normal endothelial and peritumoral cells and blood white cells) was equally distributed in low-adherent 75 cm^2 flasks using 7 different media containing EGF, FGF2, HGF or PDGFB alone or combined at different concentration (2 or20 ng/mL). An additional medium devoid of any GF was also used. A simplified scheme of derivation procedure is provided in Figure S8. Part of the same cellular resuspension was collected for DNA analysis. Cell cultures were monitored (and cell culture media were refreshed) daily until NS appearance. Newly formed NSs were propagated at clonal density in derivation medium, considered stabilized after around 10 passages, and used as described.

Cohort 1 and 2 NSs were maintained in normoxic condition (20% O_2, 5% CO_2) at 37°C. NS identity and correspondence with the parental GBM has been routinely verified by short tandem repeated profiling (PowerPlex 16 HS System, Promega) or WES analysis of SNP ID (see below; Table S4).

Conventional cell lines

Human cell lines representative of epithelial tumors (A549, ATCC; A2780, ECACC), neuroectodermal tumors (SNB19, U251 and UACC257, NCI-60; U87MG, ATCC) and normal astrocytes (NHA, Lonza Bioscience; SVG, ATCC), used for CD56 specificity validation, were kept in culture according to manufacturer's instructions (see key resources table for details) and re-authenticated soon before experiments by short tandem repeated profiling (PowerPlex 16 HS System, Promega).

Mice and experimental tumors

Animal studies were performed according to ethical regulations and protocols approved by the Italian Ministry of Health (Authorization N. 223/2015-PR). Mice (6-8 week-old male NOD.CB17-Prkdcscid/J mice, Charles River Laboratories) were housed at a maximum of 6 per cage with a 12 h light/dark cycle at 22°C, and were monitored at a minimum of twice weekly for general performance status. For generation of experimental GBMs, GSC family members were previously engineered to express luciferase and GFP as previously described[71] and *in vitro* bioluminescence was verified. A minimum of 5 randomly assigned mice were used for each injected NS. Intracranial transplantation was conducted as described[71]; briefly, 2.5×10^5 dissociated cells were stereotactically injected into the mouse brain (0.7 mm anterior from bregma, 2 mm lateral from the midline and 2 mm below the pial surface). Mice were monitored by bioluminescence imaging (IVIS Lumina System, Caliper Life Sciences) and sacrificed at the appearance of evident suffering signs. Tumor onset was established at the appearance of bioluminescence signals (10^5photons/sec).

METHOD DETAILS

DNA and RNA extraction and retrotranscription

From NSs, nucleic acids were extracted using All Prep DNA/RNA Mini Kit (Qiagen) according to manufacturer's instructions with minor modifications for preserving miRNAs. Briefly, 1.5 volumes of Absolute ethanol were used in the RNA precipitation step and wash buffer RW1 was replaced by RWT buffer (Qiagen). In some experiments either miRNeasy Mini Kit (Qiagen) or RNeasy Mini Kit (Qiagen) were used. Purified mRNAs were reverse transcribed starting from 150 or 250 ng of total RNA and using High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). DNA from NSs and parental UAs was extracted using ReliaPrep gDNA Tissue Miniprep System (Promega); RNA was isolated using Maxwell RSC miRNA Tissue Kit (Promega). Unless specified otherwise, all procedures related to DNA extraction and RNA extraction, and retrotranscription were performed according to manufacturer's instructions.

Gene copy number evaluation

Gene copy number (CN) analysis was assessed by real-time PCR, using TaqMan Universal PCR Master MIX and the ABI PRISM 7900HT sequence detection system (Thermo Fisher Scientific). Primers and probes for TaqMan CN assays are reported in the key resources table and in Table S5. Relative gene CN data were calculated by normalizing against endogenous controls (*RNaseP* or *APOA1* and *GREB1*). Normal diploid human gDNA (either from 293T cell line, ATCC, or from PBMCs) was used as calibrator to obtain the ΔΔCt. The CN of each gene was calculated with the formula $2\times2^{-\Delta\Delta Ct}$. To discriminate between real *EGFR* amplification and chr7 polysomy, the calculated CN was normalized vs. CN of a usually not amplified reference gene mapped on chr7 (*HGF*). *EGFR* amplification is defined when CN is > 3 + *HGF* CN. For other genes, amplification is defined when CN is > 5; heterozygous deletion is defined when CN is < 1.5; CN gain is defined when CN is 3<CN < 5. In cultured cells, homozygous deletion corresponds to the absence of target gene PCR product in the presence of control gene PCR product. In tumor tissues, homozygous deletion is defined when CN < 1.

Gene sequencing

In NSs, *IDH1*, *TP53*, *PTEN* were amplified using Platinum Taq DNA Polymerase (Thermo Fisher Scientific) and specific primer pairs (Table S5). PCR conditions were as follows: 95°C for 3'; 3 × [95°C for 15″, 64°C for 30″, and 70°C for 1']; 3 × [95°C for 15″, 61°C for 30″, and 70°C for 1']; 3 × [95°C for 15″, 58°C for 30″, and 70°C for 1']; 37 × [95°C for 15″, 57°C for 30″, and 70°C for 1']; and 70°C for 5'. p*TERT* was amplified using 2× Phanta Max Master Mix (Vazyme) and specific primer pairs (Table S5). PCR conditions were as follows: 95°C for 3'; 40 × [95°C for 15″, 63.9°C for 15″, and 72°C for 30'']; and 72°C for 3'. PCR products were purified using illustra ExoProStar 1-Step (Merck) according to manufacturer's instructions. Cycle sequencing was performed using BigDye Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific). Sequencing products were purified using CleanSEQ Dye-Terminator Removal Kit (Beckman Coulter) and analyzed with a 3730 DNA Analyzer ABI capillary electrophoresis system (Thermo Fisher Scientific). Data were visualized by Chromas Lite 2.01 software (http://www.technelysium.com.au/chromas_lite.html) and compared with reference sequences from the Homo sapiens assembly GRCh37. Pathogenicity of all identified variants was assessed using MutationTaster2021[58] and the Catalog Of Somatic Mutations In Cancer (COSMIC, https://cancer.sanger.ac.uk/cosmic).

Droplet digital PCR (ddPCR)

To detect *TERT* promoter mutations (c.1-124C>T and c.1-146C>T) VAF, ddPCR was performed using probe–based assays with ddPCR Supermix for Probes (no dUTP) (Biorad), according to manufacturer's instructions. Droplets were generated using AutoDG Droplet Digital PCR System (Biorad) and analyzed on a QX200 Droplet Digital PCR System using QX Manager Software Standard Edition, Version 1.2 (Biorad). ddPCR assays are reported in key resources table.

NGS target panel analysis

Quantity and quality of fresh sample- or PBMC-derived gDNA was evaluated by means of Qubit dsDNA HS or BR Assay kit (Thermo Fisher Scientific) and 1% agarose gel electrophoresis run, respectively. Library preparation was performed starting from 400 ng of DNA, firstly fragmented by using the M220 Focused-ultrasonicator (Covaris), followed by a clean-up step with an optimized ratio volume of AMPure XP beads (Beckman Coulter). Subsequent end-repair step and dA-tailing reaction of blunt-ended DNA fragments was performed by means of NxSeq AmpFREE Low DNA Library Kit (Lucigen) with small adjustments to increase the efficiency of the reactions. Adaptor ligation step has been performed with the same Lucigen's kit, by using xGen Stubby Adapter (IDT) as adaptors. After clean-up step with AMPureXP beads (Beckman Coulter), samples have been amplified (KAPA HiFi HotStart ReadyMix PCR Kit, Roche) concomitantly introducing unique sample barcodes (xGen Stubby Adapter-UDI Primers, IDT). Before target enrichment, QC of post-PCR libraries were checked by means of Qubit dsDNA BR Assay kit (Thermo Fisher Scientific) and of 2100 Bioanalyzer with a High-Sensitivity DNA assay kit (Agilent Technologies). The target of interest for GBM-custom panel design has been defined starting from the identification of genes relevant for tumorigenesis, evolution and emergence of drug resistance in GBM, thus including all coding regions of 75 genes. In details, the GBM-custom panel covers 232,760 bases with 2,570 designed probes (CUSTOM IDT xGen Custom Hyb Panel-Accel, IDT). Equal amounts of post-PCR libraries (750 ng) were pooled, for a maximum of 8 samples per pool, and subjected to the GBM-panel target enrichment with xGen Hybridization and Wash Kit (IDT)

plus xGen Universal Blockers - TS Mix (IDT) following manufacturer instruction, except for the choice to perform over-night hybridization in order to increase the on-target capture. A further amplification of the libraries has been performed with KAPA HiFi HotStart ReadyMix PCR Kit (Roche) and xGen Library Amplification Primer Mix (IDT), thus reaching needed amount of final libraries. Final libraries were quantified by means of Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and their fragment distribution evaluated using High-Sensitivity DNA assay kit (Agilent Technologies). Equal molar amounts of DNA libraries were pooled and, based on the total number of samples ready to be analyzed, sequenced using Illumina MiSeq or NextSeq500 sequencer (Illumina). Data analysis was performed using a bioinformatic pipeline previously described.[72,73] A metanormal was built from fastQ files obtained by 10 PBMC samples sequenced using the same laboratory procedures. Alignments from metanormal and UA samples were compared to identify mutations/indels in tumor and metanormal sample. Somatic alterations were present only in tumor while germline ones were common to both samples. NGS artifacts were further filtered following the methods previously described.[73,74] Then only variants with 5% significance level obtained with a Fisher's exact test, supported by a minimum of 4 mutated reads in regions with 5× minimum depth and with allele frequency >2.5% were considered. Indels were called using Pindel tool in both alignments and only somatic indels with fractional abundance >10% were reported. Gene CN variations analysis was performed in the matched samples (Tumor vs. Metanormal) for each patient as previously reported.[73,74]

WES analysis

For GSC families and UA samples, library preparation was performed starting from 1 μg of total DNA using Agilent SureSelectXT HumanAllExon V5+UTRs and sequenced on a Nextseq500 (Illumina). FastQ files were generated by Illumina sequencer and analyzed according to a previously set-up bioinformatic pipeline.[73] Briefly, human reads were mapped to assembly of human reference genome version 19 using BWA-mem algorithm[69] with standard parameters. PCR duplicates were removed using the RMDUP command of SAMtools package[70] and, to delete sequencing artifacts, reads having more than 3 mismatches compared with reference genome were filtered out. Only the bases with Phred Score >30 were selected for the genetic analysis. Furthermore, mutations supported by reads' groups having strand bias (mutations supported by ratio between strands >0.9 or <0.1), or by mismatch only in head/tail of the read (the first and the last nucleotide of the read) were ignored. In order to evaluate the similarity of UA with derivative GSCs, hierarchical cluster analysis was performed using NGS data of each family sample (GBM14, GBM17, GBM18, GBM20). For this analysis, the ratio between the median gene depth and median depth of the whole exome was defined as Gene Copy Number (GCN). The VAFs were normalized using GCN, obtaining adjusted VAFs, and, only when the adjusted fractional abundance was higher than 1%, the mutations were used to build distance matrix and agglomerative hierarchical clustering. The package Cluster v 2.1.0 of the software package R v 3.6.3 (www.r-project.org) was used with unweighted pair group average method (UPGMA) for the hierarchical clustering analysis and for the genetic distance identification among the samples. Mutations were grouped in private, common and shared and were shown (with corresponding adjusted frequency) in the heatmap.

Sample assignment of GBM14 cell lines was performed using the single nucleotide polymorphism (SNP) identification through NGS from WES data. In detail, the list of SNPs present in dbSNP version 155 in coding sequences was extracted and used to identify the sample allelic profile (AP) and to build the SNP identifier (SNP_ID) of each sample. All SNP_ID were compared to establish the correct sample matching and results were reported in Table S4. Each allele was considered only if the fractional abundance was higher than 10% with a minimum depth of 10X in both samples. Samples were considered "matched" if the allelic homology percentage was higher than 95%.

FISH analysis

In order to evaluate *EGFR* and *MYC* amplification, FISH analysis was performed using Vysis *EGFR*/CEP 7 FISH Probe Kit (Abbott), including an orange fluorochrome probe specific for *EGFR* and a green fluorochrome probe specific for the corresponding alpha satellite centromeric region as control, and the ZytoLight *SPEC MYC* Dual Color Break Apart Probe (Zytovision), also able to detect gene rearrangements. FISH was performed on both bioptic samples (*MYC*) and interphase nuclei (*MYC*, *EGFR*) according to manufacturer's instructions. For bioptic samples, formalin-fixed paraffin-embedded (FFPE) sections were deparaffinized, air-dried, incubated in Pre-Treatment Solution at 98 °C for 10 min, followed by proteolytic digestion using Pepsin (Agilent). After air dehydration, 10 μL of probe mixture were applied to each sample. Slides and probes were co-denatured at 75 °C for 10 min and hybridized at 37 °C for 16 h in the dark (Top Brite). After washing with Stringent Wash Buffer (Agilent), chromatin was counterstained with DAPI 150 ng/mL (Zytovision). For analysis of interphase nuclei, cell suspensions were incubated with the probe for 10 min at 75°C for co-denaturation and placed in a humidified chamber at 37°C O/N for the hybridization step. After washing with ISH Stringent Wash Buffer (Agilent), chromatin was counterstained with DAPI 150 ng/mL (Zytovision). An average of 100 cells was analyzed using an Olympus BX61 microscope (Olympus Corporation) and CytoVision software (Leica Biosystems). *EGFR* CN was determined by calculating the ratio between the gene and the centromere CN, and considered amplified if the ratio gene CN/centromere CN was ≥ 2. *MYC* CN was defined as increased (gain) when the ratio gene CN/nuclei was $3 < n < 6$ and amplified if the ratio was ≥ 6.

Bulk gene expression profiling

RNA integrity was assessed using the 2100 Bioanalyzer with RNA 6000 Nano kit (Agilent) and quantified using Qubit RNA BR Assay kit (Thermo Fisher Scientific). Libraries were prepared starting from up to 400–800 ng of total RNA with TruSeq Stranded mRNA Library Preparation Kit (Illumina) according to the manufacturer's protocol. Library size was assessed using the 2100 Bioanalyzer with a High

Sensitivity DNA assay kit (Agilent) and quantified using the Qubit dsDNA HS kit. Libraries were normalized to 10nM in TrisHcl 10mM pH = 8.5, pooled and sequenced using Illumina MiSeq and NovaSeq 6000 Sequencer (Illumina) performing 150 reaction cycles in single end. FastQ files were generated using bcl2fastq (Illumina). Each generated fastQ file was aligned using STAR 2.5.4[60] and mapped to the human GRCh38 genome reference. The GENCODE release 27 was used as transcriptome reference annotation, and gene expression quantification was performed with featureCounts.[61,62] Differential expression analysis was carried out in R environment (v4.2.2) with DESeq2[63] using the following thresholds: p value < 0.005 and logFC >2. GO analysis was performed by enrichR (https://maayanlab.cloud/Enrichr/).[64,65,75] Clustering and data representation were performed with the GEDAS software.[57]

scRNA-seq and data processing

Dissociated NSs were stained with Annexin V and DAPI as described[76] and sorted by means of Fluorescence-Activated Cell Sorting (FACS, MoFlo Astrios EQ, Beckman Coulter) for being Annexin V^{neg}/$DAPI^{neg}$ as to analyze only perfectly viable cells. Sorted cells were processed with Chromium Next GEM Single Cell 3' Reagent Kits v3.1 Dual Index and Chromium Controller instrument (10× Genomics, Inc) following manufacturer's instruction. Briefly, single cell suspensions were brought to a concentration of 10^3 cells/μL, and 5 × 10^3 cells were processed to obtain 3 × 10^3 targeted cells. For cDNA amplification 12 cycles were used, and for the index PCR 14 amplification cycles were performed. Final libraries were quantified with Qubit dsDNA HS Assay Kit (ThermoFisher Scientific) and their fragment distribution evaluated with High-Sensitivity DNA assay kit (Agilent Technologies). Equal amounts of DNA libraries were pooled and sequenced using Illumina NovaSeq6000 sequencer (Illumina Inc.), thus generating sequencing data as required by manufacturer's instruction (read1 28bp, read2 90bp, double index 10bp). Sequencing data were aligned to the human reference genome (GRCh38) and processed using the CellRanger 7.1.0 pipeline (10x Genomics)[66] to obtain 1577 cells for member E, 1323 for member F and 1906 for member NO GF, with genes assigned to at least 3 cells. The raw gene expression matrix from the CellRanger pipeline was filtered and normalized using the "LogNormalize" method from the Seurat R package.[67] Filtering was performed according to the following criteria: 1) cells with <10% of mitochondrial gene expression in counts and ribosomal genes <15% (# of cells E = 1157, # of cells F = 1005, # of cells NO GF = 1170), and 2) cells displaying more than 1000 expressed genes,i.e., genes with at least 1 read (# of cells E = 764, # of cells F = 333, # of cells NO GF = 916). Doublet Finder was employed to remove cell doublets and pN-pK parameters were chosen as suggested[68] (final dataset composed of # of cells E = 755, # of cells F = 330, # of cells NO GF = 901). The gene expression of filtered cells matrices was normalized to the total UMI counts per cell and transformed to the natural log scale.

To evaluate expression of transcriptional signatures in GSC family members, we calculated metagenes as the average expression of the gene sets in each cell. Accordingly, Montecarlo randomization was performed on Neftel gene signatures to estimate the null distribution in each profile for each score. Cells with scores higher than the 99^{th} percentile of corresponding null distribution were assigned. The results are reported in Table S3. To assign a unique MES-like or NPC-like score to each cell, the highest score between MES-like 1 and 2 or NPC-like 1 and 2 was considered.

Quantitative real-time PCR (qPCR)

Real-time PCR for evaluation of gene mRNA expression was performed using primer and probe sets (Thermo Fisher Scientific) listed in the key resources table, with TaqMan Universal PCR Master Mix and an ABI PRISM 7900HT sequence detection system (Thermo Fisher Scientific). Expression levels were normalized against endogenous controls (β-actin and β2 microglobulin). When appropriate, control cells were used as calibrators. Expression levels were reported as 40-Ct or as fold vs. control cells, and are the mean ± SEM of two independent experiments in triplicate.

In situ RNA hybridization (RNAscope)

RNA chromogenic *in situ* hybridization was performed on 5 μm thick formalin-fixed paraffin-embedded (FFPE) samples using RNAscope Duplex Assay (Advanced Cell Diagnostics, Bio-Techne) according to manufacturer's instructions. Briefly, FFPE sections were deparaffinized, dehydratated and pretreated using three consecutive incubations as follow: (i) hydrogen peroxide, 10 min RT, (ii) target retrieval, 15 min at boiling temperature, and (iii) ACD HybEZ hybridization with protease, 30 min at 40°C. Pretreated samples were incubated 2 h at 40°C with the indicated probes, and signals were amplified and detected with respectively green (channel 1) or fast red (channel 2) substrate. Hematoxylin was used for staining cellular nuclei and guided discrimination between the GBM and the PNC component. Staining was evaluated with the use of an Axio Zeiss microscope. Quantification was performed on ImageJ (NIH) according to RNAscope Duplex Assay manufacturer's instruction (Advanced Cell Diagnostics, Bio-Techne). Briefly, 5 high resolution (40×) images/probe pair were evaluated measuring both the percentage of cells expressing the target and the average target expression level. The latter was evaluated according to the ACD scoring system (0: no staining; 1: 1–3 dots/cell; 2: 4–9 dots/cell not in cluster; 3: 10–15 dots/cell in clusters; 4: >15 dots/cell in clusters). Target genes and probed regions are listed in the key resources table. Sequences of target probes, pre-amplifier, amplifier and label probes are proprietary (Advanced Cell Diagnostics, Bio-Techne).

Western blotting

NS cells were lysed in boiling Laemmli Buffer as previously described[71] and clarified by sonication and centrifugation (13,000 rpm for at least 10 min). Total amount of proteins obtained was quantified using the BCA System (Pierce). 10 to 20 μg of proteins were loaded and separated by SDS-PAGE on a 4–15 or 4-20% pre-casted polyacrylamide gel (Biorad) and transferred on nitrocellulose (BioRad).

After membrane saturation, primary antibodies used were MET, Synaptophysin, BTG2, NeuN, ASCL1, Doublecortin, GFAP, Vimentin, YAP1, CD44 and beta-3-tubulin (see key resources table). Antibodies were visualized with appropriate horseradish peroxidase-conjugated secondary antibodies (Jackson Lab) and enhanced chemiluminescence system (Promega). Blot images were captured using the ChemiDoc Touch Imaging System (Biorad) with Image Lab software (Biorad). Vinculin was used as protein loading control as indicated.

Flow-cytometric analysis

For fresh samples analysis, a portion of single cells obtained after UA dissociation was immediately stained with CD45 and CD56 antibodies (CD45 PE-CF594, clone HI30, #562312, BD Horizon; CD56 PE-Cy7, clone N901-NKH-1, #A21692, Beckman Coulter, see key resources table) and analyzed by flow cytometry. The CD45/CD56 marker pair was chosen to discriminate GBM cells ($CD45^{neg}/CD56^{pos}$) from blood-derived cells ($CD45^{pos}/CD56^{neg}$: leukocytes; $CD45^{pos}/CD56^{pos}$: NK cells) based on previous evidence[27] and analysis of a panel of cell lines including epithelial tumors, GBMs, a melanoma and the two astrocyte lines (NHA and SVG), showing that CD56 is expressed by GBMs but not by astrocytes (Figure S9). In the gated $CD45^{neg}/CD56^{pos}$ cells, cell viability was measured with DAPI staining.

For analysis of cell surface RTKs, and CD24 and CD44 markers, established cell cultures were mechanically dissociated, resuspended at the optimal concentration of 2×10^6 cells/mL in 1× PBS + BSA 1% for FcBlock, and then incubated with the indicated antibodies (see key resources table) for 15 min RT on a seesaw rocker in the dark. Dead cells were excluded as $DAPI^{pos}$ (Roche). For each condition, fully unstained cells were processed as a negative control and, in multiparametric analysis, the Fluorescence Minus One controls were used in order to set the correct positive gates. Samples were acquired on a CyAn ADP (Beckman Coulter) equipped with 488 nM, 405 nM and 642 nM solid-state lasers. Collected data were analyzed with Summit 4.3 software (Beckman Coulter), applying the available compensation matrices, including the VisiComp scaling algorithm to avoid overcompensation errors.

Pseudodifferentiation assay

NSs were dissociated at single cell levels and 10^6 cells were seeded on a poly-lysinated multichambered slide in 1% FBS. 7 days after seeding, cells were fixed with PAF 4% for 20 min. Fixed cells were permeabilized with Triton 0.2% for 10 min, washed and saturated with PBS-TWIN BSA 5% for 1 h. Primary antibodies beta3-tubulin, GFAP and GalC (see key resources table), diluted 1:50 in PBA-BSA 0.5%, were incubated for 1 h at RT and then O/N at 4°C. AlexaFluor conjugated secondary antibodies (Thermo Fisher Scientific), diluted 1:1000 in PBA-BSA 0.5%, were added and slides were incubated for 1 h. DAPI was added for 30 s. Images were captured using LASV4.2 software on a LEICA SPEII confocal microscope and are representative of at least three independent stainings. GBM14 NO GF and P members were subjected to different pseudodifferentiation protocols (modifying % of FBS, timing and matrix). The percentage of beta3-tubulin, GFAP and GalC positive cells has been measured by counting positive cells/DAPI positive cells in at least 5 different images.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tissue sections (either from human GBM validation cohort or from experimental GBMs) underwent single or double immunohistochemical staining. Briefly, sections were de-waxed, rehydrated, and endogenous peroxidase activity blocked with 0.3% H_2O_2 in methanol for 20 min. Antigen retrieval was performed using a microwave oven or thermostatic bath in 1.0 mM EDTA buffer (pH 8.0) or 1.0 mM citrate buffer (pH 6.0). Sections were then washed in TBS (pH 7.4) and incubated for 1 h or O/N with the specific primary antibody diluted in TBS 1% bovine serum albumin. Signal was revealed using the DAKO Envision+System-HRP Labeled Polymer Anti-Mouse or Anti-Rabbit (Dako Cytomation) or NovoLink Polymer Detection System (NovocastraTM) followed by Diaminobenzydine (DAB, Dako Cytomation) as chromogen and hematoxylin as counterstain. For double immunostains, after completing the first immune reaction, the second primary antibody was applied and labeled using MACH 4TM Universal AP Polymer Kit (Biocare Medical); chromogen reaction was developed with Ferangi BlueTM Chromogen System (Biocare Medical), and nuclei were faintly counterstained with hematoxylin. Images were acquired with a Nikon DS-Ri2 camera (4908 × 3264 full-pixel) mounted on a Nikon Eclipse 50i microscope equipped with Nikon Plan lenses using NIS-Elements 4.3 imaging software (Nikon Corporation). Primary antibodies used were listed in the key resources table and are the following: EGFR, GFAP, Synaptophysin, MET and PDGFRA.

Extreme limiting dilution assay

For evaluation of stem cell frequency in GSC family members, limiting dilution assay was performed as previously described.[71] Briefly, cells were seeded at the concentration of 100, 50, 25, 10, 5 and 1 cell per well (10 replicates for each condition) in 96-well plates in the derivation medium. For p2 and p3 passage, NSs were collected, dissociated, and seeded at the concentrations of 25, 12.5, 6.25, 3.125, 1.56 and 0.78 cells per well. Data reported are derived from passage p2. Wells containing NSs were counted and data were analyzed with ELDA software[59] (http://bioinf.wehi.edu.au/software/elda/).

Cell proliferation assay

10^3 dissociated cells per well were plated in a 96-well plate at day −1 in a medium devoid of any GF. Where indicated, on day 0 cells were stimulated with the full GF medium (EFPH20). Viability was measured by CellTiter-Glo Luminescent Cell Viability Assay (Promega) using a GloMax 96 Microplate Luminometer (Promega) at day 0 (ctrl), and at the indicated time points after treatment. Data were reported as mean ± SEM of at least two independent experiments (n > 4).

QUANTIFICATION AND STATISTICAL ANALYSIS

Where indicated, data were expressed as mean ± standard error of the mean (SEM) of at least two independent experiments. The number of biological (nontechnical) replicates for each experiment is indicated in the figure legends. Statistical comparisons were performed using the parametric Student *t* test, nested *t* test, Fisher's exact test and Chi-square test as reported. Co-occurrence and mutual exclusivity were evaluated by non-parametric Spearman correlations. Survival curves were analyzed using the Kaplan-Meier method with groups compared by respective median survival; Mantel-Cox tests were used for calculating log rank p values. A p value < 0.05 was considered significant. All statistical tests were performed with Prism v8.0 software (GraphPad). Where indicated, ELDA software was used.