Stem Cell Reports
Report

ISSCR

OPEN ACCESS

MicroRNA Signatures and Molecular Subtypes of Glioblastoma: The Role of Extracellular Transfer

Jakub Godlewski,[1,*] Ruben Ferrer-Luna,[2] Arun K. Rooj,[1] Marco Mineo,[1] Franz Ricklefs,[1,3] Yuji S. Takeda,[1] M. Oskar Nowicki,[1] Elżbieta Salińska,[4] Ichiro Nakano,[5] Hakho Lee,[6] Ralph Weissleder,[7] Rameen Beroukhim,[2] E. Antonio Chiocca,[1] and Agnieszka Bronisz[1,*]

[1]Department of Neurosurgery, Harvey Cushing Neuro-oncology Laboratories, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
[2]Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Cancer Program, BROAD Institute of MIT and Harvard, Cambridge, MA 02142, USA
[3]Department of Neurosurgery, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany
[4]Department of Neurochemistry, Mossakowski Medical Research Centre, Polish Academy of Sciences 02-106 Warsaw, Poland
[5]Department of Neurosurgery and Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL 35243, USA
[6]Center for Systems Biology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA
[7]Department of Systems Biology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA
*Correspondence: jgodlewski@bwh.harvard.edu (J.G.), abronisz@bwh.harvard.edu (A.B.)
http://dx.doi.org/10.1016/j.stemcr.2017.04.024

SUMMARY

Despite the importance of molecular subtype classification of glioblastoma (GBM), the extent of extracellular vesicle (EV)-driven molecular and phenotypic reprogramming remains poorly understood. To reveal complex subpopulation dynamics within the heterogeneous intratumoral ecosystem, we characterized microRNA expression and secretion in phenotypically diverse subpopulations of patient-derived GBM stem-like cells (GSCs). As EVs and microRNAs convey information that rearranges the molecular landscape in a cell type-specific manner, we argue that intratumoral exchange of microRNA augments the heterogeneity of GSC that is reflected in highly heterogeneous profile of microRNA expression in GBM subtypes.

INTRODUCTION

Intercellular dialogue between tumor cells mediated by extracellular vesicles (EVs) is a powerful means of communication that facilitates exchange of active molecules (Bronisz et al., 2014; Skog et al., 2008). It is the linchpin of the molecular network, the invisible causeway on which the global cellular transcriptome hums. However, the culpability of EV communication for phenotypic and molecular diversity within heterogeneous tumors is not fully recognized.

Intratumoral heterogeneity and invasiveness are the key characteristics of glioblastoma (GBM), the most common and most aggressive primary brain malignancy in adults, with a median survival of 14.2 months (Johnson and O'Neill, 2012). The subpopulation of highly tumorigenic and therapy-resistant GBM stem-like cells (GSCs) (Schonberg et al., 2014) retains stem cell characteristics, including self-renewal and undifferentiated status, but also exhibits varying degrees of phenotypic and molecular polymorphism. Discovering whether the underlying cause of invasiveness is inherent or is a response to microenvironmental stimuli has important implications for better understanding of GBM pathobiology.

The signatures of protein-coding gene expression and somatic copy-number alterations have revealed the existence of several distinct subtypes among GBM patients, known as mesenchymal, proneural, neural, and classical according to The Cancer Genome Atlas (TCGA) database (Phillips et al., 2006). The classification is further complicated by the fact that individual tumors contain a spectrum of subtypes and hybrid cellular states (Patel et al., 2014) and that GSC subpopulations retain transcriptome heterogeneity (Mao et al., 2013). These findings indicate that tissue-based classification likely shows merely characteristics of the predominant cellular component. Importantly, tissue and GSC subtype classification was also demonstrated by signatures of non-protein-coding genes, such as long non-coding RNAs (Du et al., 2013; Mineo et al., 2016). Significantly, microRNAs have not been shown to predict GBM classification and prognosis by global signature to date, while being strongly implicated as functionally deregulated in GBM as individual molecules (Godlewski et al., 2015).

We aimed to highlight the magnitude of EV/microRNA-driven propagation of molecular and phenotypic diversity of GSCs. Using intracranial xenografts of patient-derived GSCs, we selected subpopulations of cells with distinct transcriptomes, displaying proliferative/nodular or migratory/invasive modes, which are associated with mesenchymal-like or proneural-like subtype, respectively. The highly heterogeneous expression profile of microRNAs in GBMs was separable into two unsupervised classes that partially overlapped with previously determined molecular subtypes, with both subclasses of GSCs displaying differential cellular and EV microRNA profiles. The analysis of microRNA/target networks provided evidence that

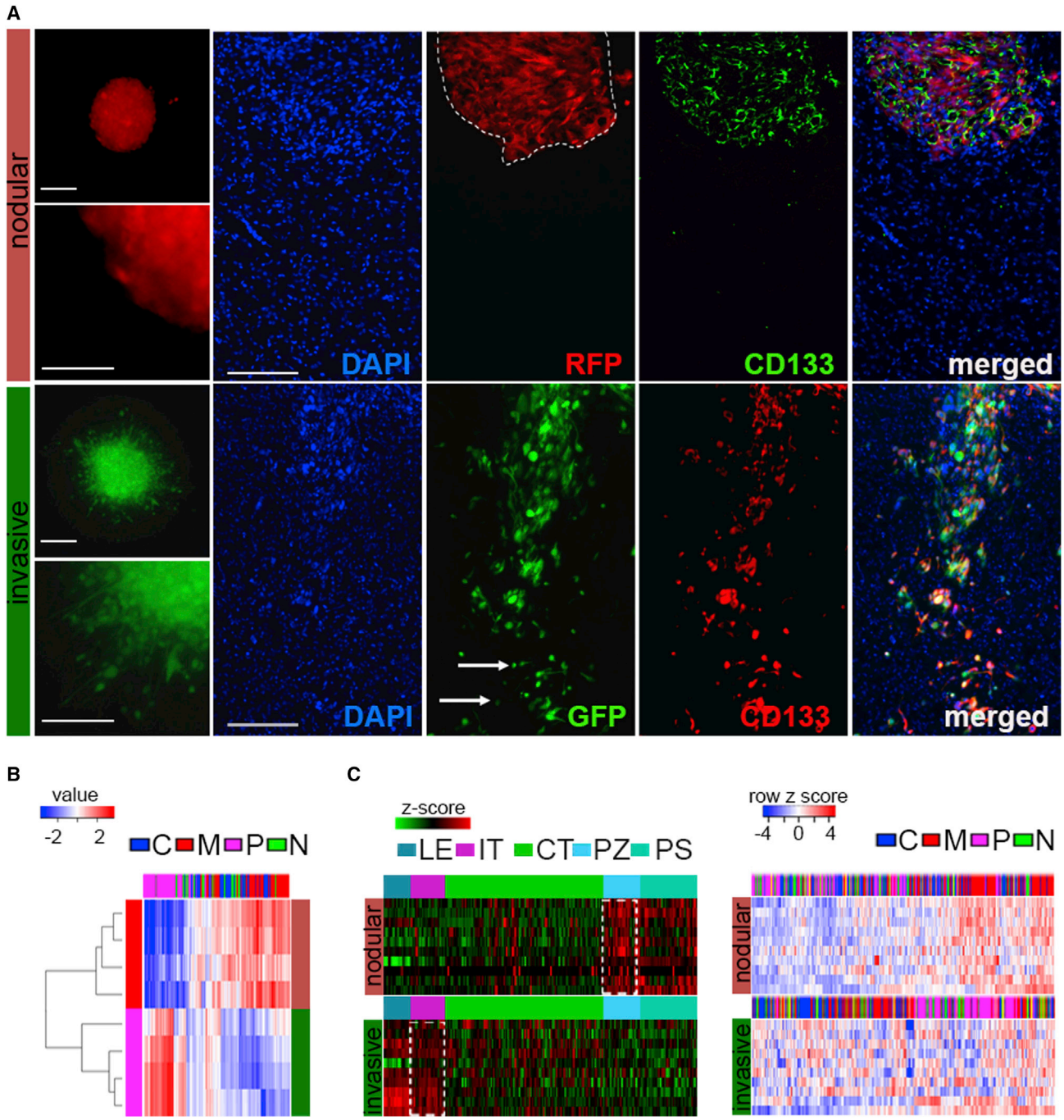


Figure 1. Intratumoral Architecture Is Imposed by Phenotype/Transcriptome-Diverse GSCs

(A) Patient-derived GSCs have distinct phenotypes in vitro and in vivo. Representative micrographs of GSC spheroids (left, n = 10 independent GSCs; scale bars, 100 μm and 50 μm) and GSC-derived intracranial tumors with CD133 immunostaining (right, n = 6 independent GSCs; scale bar, 150 μm). Nodular tumor and infiltrating tumor cells are indicated by the dashed line and arrows, respectively.

(B) Signature of genes with proproliferative or proinvasive function classifies GSC subpopulations. Gene expression (n = 8 independent GSCs, n = 4 per subclass) in selected categories of GSCs was queried with a gene signature retrieved from the TCGA GBM dataset, and identified by clustering as expression correlation analysis. C, classical; M, mesenchymal; P, proneural; N, neural.

(C) The phenotype-determining transcriptome overlap with tumor anatomic site-specific expression. The top ten genes in each category (n = 8 independent GSCs, n = 4 per subclass) (proproliferative or proinvasive) were queried with Ivy GAP database-based expression

(legend continued on next page)

EV/microRNAs are modifiers of both the molecular landscape and phenotype, acting via cell-dependent targeting that propagates GBM subtype heterogeneity.

RESULTS

Intratumoral Architecture Is Imposed by Phenotype/Transcriptome-Diverse GSCs

Recognizing molecular determinants which drive diverse GBM cell phenotypes would allow the identification of functional targets and provide much-needed insight into the phenotypic heterogeneity of GBM. Patient-derived GSCs revealed two major distinct subpopulations with nodular and invasive phenotypes by spheroid dispersal assay in vitro and intracranial xenografts in vivo (Figure 1A). The expression of GSC marker PROMININ 1 (CD133) showed significant upregulation in both subpopulations in vivo (Figure S1A), indicating that these cells retain a stem-like character in the brain microenvironment. The analysis of molecular and cellular function of genes deregulated in these two subpopulations showed proliferative and migratory modes of the transcriptome (Table S1 and Figure S1B). In fact, these genes' signatures, when queried with TCGA subtypes (Verhaak et al., 2010), clustered with mesenchymal or proneural subtypes (Figure 1B), indicating that these phenotypically diverse GSCs are characterized by distinct transcriptomic subtype classification. The phenotype-linked transcriptomics overlapped with tumor anatomic site, with mesenchymal-like/nodular signatures prevalent in perinecrotic zones and proneural-like/invasive signature in infiltrating areas of tumor (Figures 1C and S1C), suggesting that complex intratumoral architecture may arise from the co-existence of diverse GSCs within individual tumors.

GBM Subtypes Are Characterized by Highly Heterogeneous MicroRNA Profiles

Global analysis of microRNA expression showed that the transcriptional subtype diversity observed in nodular and invasive GSCs is reflected by microRNA signatures (Figures 2A and S2A). The comprehensive analysis of subtype-specific microRNAs and their mRNA targets showed that the downregulated microRNAs in each GSC subpopulation negatively correlated with expression of their targets in a tumor anatomic site-dependent manner (Figure 2B [left panels] and Table S2). However, targets of highly expressed microRNA in these two GSC subpopulations did not show an anatomic site-dependent pattern (Figure 2B, right panels). The analysis of microRNA expression in GBM TCGA subtypes identified by protein-coding gene signatures revealed a highly heterogeneous pattern with no apparent clustering (Figures 2C [left] and S2B). However, unsupervised analysis of the GBM microRNA expression profiles by non-negative matrix factorization (NMF) (Brunet et al., 2004) identified two major clusters (C1 and C2) (Figure 2C, right; Figures S2C and S2D), which largely aligned with the molecular classes previously determined by the gene expression analyses (Figure 2D). C1 class (nodular) was significantly enriched in mesenchymal GBM tumors, whereas C2 class (invasive) exhibited enrichment in proneural and neural molecular groups (Figure S2E). The distinct profile of microRNA expression identified in GSC subpopulations in vitro, taken together with the observed negative correlation of microRNAs and their targets' expression in vivo, suggested tumor microenvironment-dependent regulation of microRNA expression and targeting.

To validate whether the observed diversity of GSC cellular microRNAs was recapitulated by the microRNA composition of EVs released by these GSCs, we isolated, quantified, and characterized EVs. Interestingly, significant heterogeneity of EVs released by these two GSC subpopulations was observed regarding their size, shape, and expression of EV marker CD 63 molecule (CD63), but not the total number of particles or their RNA content (Figures S2F–S2J). Comprehensive analysis of cellular and EV microRNA in nodular and invasive GSCs showed that the microRNA signature separated GSCs and GSC EVs (Figures 2E, S2K, and S2L). The fact that cellular microRNA from two subpopulations of GSCs clustered together rather than with their own EV microRNA indicated that GSC EV microRNA profiles only partially mimic cellular microRNA expression, with sets of microRNAs enriched/depleted in EVs, and that subpopulation-specific EV microRNA signatures also exist (Figures S2M and S2N). Similar to GSC microRNA, GSC EV microRNA significantly separated into two classes which recapitulated the same classification as cellular microRNAs, despite the fact that different sets of microRNAs were identified in cells and EVs (Figure 2F). We thus hypothesized that in a setting as diverse as GBM, microenvironment-driven signaling may mediate dynamic transitions within tumor anatomic niches.

signature in different anatomic areas of GBM (left; LE, leading edge; IT, infiltrating tumor; CT, cellular tumor; PZ, perinecrotic zone; PS, pseudopalisading cells), or a gene signature retrieved from the TCGA GBM dataset and identified by clustering with subtype prediction (right; C, classical; M, mesenchymal; P, proneural; N, neural).
See also Figure S1.

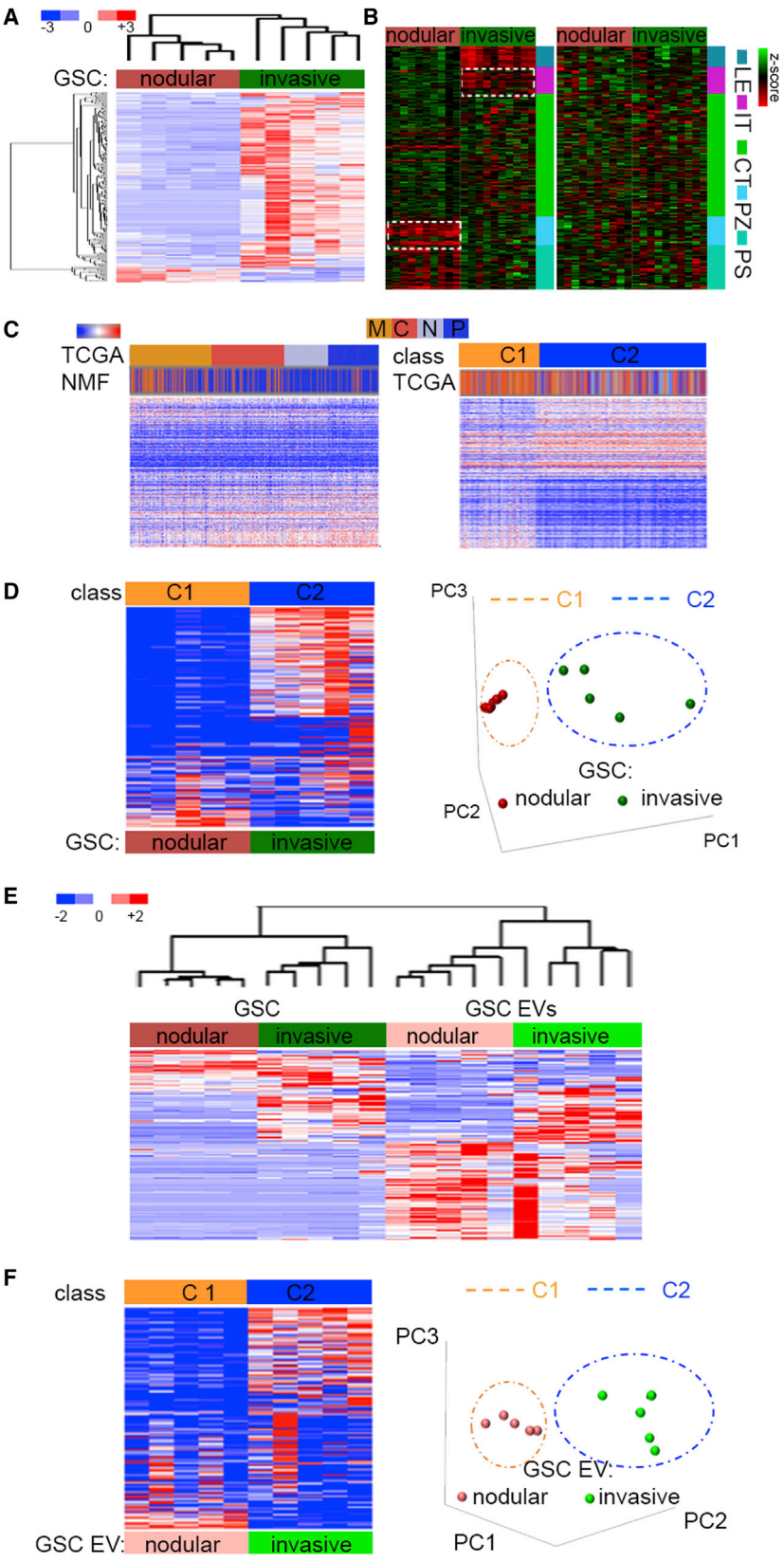


Figure 2. GBM Subtypes Are Characterized by Highly Heterogeneous MicroRNA Profiles

(A) MicroRNA profile distinguishes nodular and invasive GSCs. MicroRNA sets that vary coherently between GSCs (n = 10 independent GSCs, n = 5 per subclass) were identified by supervised clustering (fold >2, $p < 0.05$).

(B) MicroRNAs downregulated in GSC subpopulations (n = 10 independent GSCs, n = 5 per subclass) lack correlation with tumor anatomic site expression of their targets. IPA-based analysis of selected microRNA/mRNA target expression showed negative correlation (low microRNA/high target expression, left panels), and lack of correlation (high microRNA/low target expression, right panels). The top ten genes in each category were queried with Ivy GAP database-based expression signature in different areas of GBM (LE, leading edge; IT, infiltrating tumor; CT, cellular tumor; PZ, perinecrotic zone; PS, pseudopalisading cells). White dashed box indicates genes upregulated in PZ and LE zones.

(C) TCGA-classified GBM subtypes are characterized by highly heterogeneous profiles of microRNA expression. Hierarchical clustering of expression of 534 microRNAs in a core set of TCGA GBM samples (n = 173 patient samples) in supervised analysis (TCGA subtypes, top cluster) compared with unsupervised analysis (using NMF, bottom cluster) (left); and in supervised analysis (NMF-based classification into two classes C1 and C2, top cluster) versus unsupervised analysis (TCGA subtypes, bottom cluster) (right; C, classical; M, mesenchymal; P, proneural; N, neural; 200 microRNAs, false discovery rate [FDR] < 0.05).

(D) GSC microRNA expression data reveals two profiles in TCGA-classified GBM subtypes. Hierarchical clustering (left) and principal component analysis (right) of microRNAs using 396 predictive microRNAs expressed in GSC subpopulations (n = 10 independent GSCs, n = 5 per subclass) and ordered based on gene subtype predictions using the core set of TCGA GBM samples (n = 173 patient samples) (NMF1 versus NMF2; 200 microRNAs, FDR < 0.05).

(E) GSCs and GSC EV microRNA profiles separate cells and EVs. Hierarchical clustering of expression of 692 microRNAs in GSCs (n = 10 independent GSCs, n = 5 per subclass) and GSC EVs (n = 10 EVs from independent GSCs, n = 5 per subclass) in unsupervised analysis.

(F) GSC EV microRNA expression data reveal two profiles in TCGA-classified GBM subtypes. Hierarchical clustering (left) and principal component analysis (right) of EV microRNAs using the

(legend continued on next page)

GSC EVs Support a Subpopulation-Specific Invasive Phenotype

The treatment of invasive GSCs with EVs derived from nodular GSCs enhanced their migration in vitro, suggesting that EV exchange may contribute to cellular phenotype (Figure S3A and Movie S1). Thus, a heterogeneous sphere migration assay was performed to mimic co-existence of these subtypes within the tumor. Significantly, invasive cells migrated farther away from the spheroid core in the presence of other cells, either other invasive cells (Figure S3B) or, even more profoundly, in the presence of nodular cells (Figures 3A and S3C; Movie S2). Importantly, nodular cells remained within the spheroid regardless of co-culture variants. This phenotype was also recapitulated in vivo in a heterogeneous GSC intracranial xenograft model, showing subpopulation-specific invasive and nodular behavior (Figure 3B, left). As a proof of concept, we selected GBM-high microRNA-31 (miR-31), which was enriched in donor nodular GSCs and their EVs (Figures S2M and S2N). Frequent exchange of EVs and EV microRNAs and their spread away from the cell of origin suggested both the existence of molecular hybrids and propagation of heterogeneity across the tumor (Figures 3B [right], 3C, and S3D). Together, these data suggested that EV exchange did not cause a phenotypic switch but rather facilitated formation of interdependent tumor organization. The analysis of relation between expression of microRNAs from selected classes (C1 and C2) and genes expressed differentially in tumor anatomic niches (perinecrotic/core zone versus infiltration/invasive zone) revealed a positive association in GBM tissue (Figures S3E and S3F). Thus, EV-mediated transfer of bioactive molecules leads to increased heterogeneity, allowing a more robust response to microenvironmental challenges and leading to increased cell survival (Ricklefs et al., 2016).

Transfer of EV-Encapsulated MicroRNAs Propagates GBM Heterogeneity

EVs are complex structures with cargo composed of multiple classes of molecules; thus, their uptake may result in EV-dependent alterations much broader than those caused by microRNA alone. It is likely impossible to "tease out" microRNA from other active molecules present in EVs. Transfer of EVs between GSC subpopulations resulted in shifts of clustering with significant global deregulation of microRNAs in each GSC subpopulation treated with EVs derived from other subpopulations (Figure 4A). The most enriched microRNAs in EV-treated cells (Figure S4A) showed a diverse pattern of expression in GBM TCGA subtypes (Figure 4B). However, EV microRNAs from invasive GSCs were associated with a significantly worse outcome in mesenchymal tumors. Conversely, EV microRNAs from nodular GSCs were associated with a significantly worse outcome in proneural but not mesenchymal tumor (Figures 4C and S4B). Nodular GSC- and EV-specific miR-31 (Figures S2M and S2N) was enriched in recipient invasive GSC upon exposure to nodular GSC EVs (Figures 4D [right], S4C, and S4D), or after sorting from heterogeneous spheroids (Figure 4D, left), recapitulating exchange observed in vivo (Figure 3C). Interestingly, putative miR-31 targets were differentially expressed in both GSC subpopulations, indicating that different sets of targets exist in each subgroup of GSCs (Figure S4E). In fact, invasive GSC-specific miR-31 targets (Wong et al., 2015) were significantly downregulated upon treatment with EVs from nodular GSCs, but deregulation of miR-31 by transfection with an miR-31 mimic or inhibitor in donor nodular GSCs had no significant effect on their uniformly low expression (Figure 4E). The analysis of transcriptome of invasive GSCs treated with EVs derived from nodular GSC expressing either control or miR-31 inhibitor showed potent gene rearrangement with significant impact on miR-31 targets (Figure S4F). However, these targets did not show a tumor anatomic niche-specific pattern (Figure S4G), indicated that both cellular and EV microRNAs have a cell-specific function, targeting effectors existing exclusively in particular GSC subpopulations.

DISCUSSION

Our data indicated that phenotype-linked transcriptomics of GSCs overlapped with tumor anatomic site, with mesenchymal-like/nodular signatures prevalent in perinecrotic zones and proneural-like/invasive signature in infiltrating areas of tumor, suggesting that these GSCs both shape and adapt to microenvironmental conditions, and that complex intratumoral architecture likely arises from the co-existence of diverse GSCs within individual tumors (Patel et al., 2014).

The distinct microRNA profile identified in GSC subpopulations in vitro, and the negative correlation of microRNAs and their targets' expression in vivo, suggested tumor microenvironment-dependent regulation of microRNA expression and targeting. The concordance between gene and microRNA unsupervised expression

curated list of 298 predictive microRNAs secreted in EVs released by distinct GSC subpopulations (n = 10 independent GSCs, n = 5 per subclass) and ordered based on gene subtype predictions using the core set of 173 TCGA GBM samples (n = 173 patient samples). FDR < 0.05.
See also Figure S2.

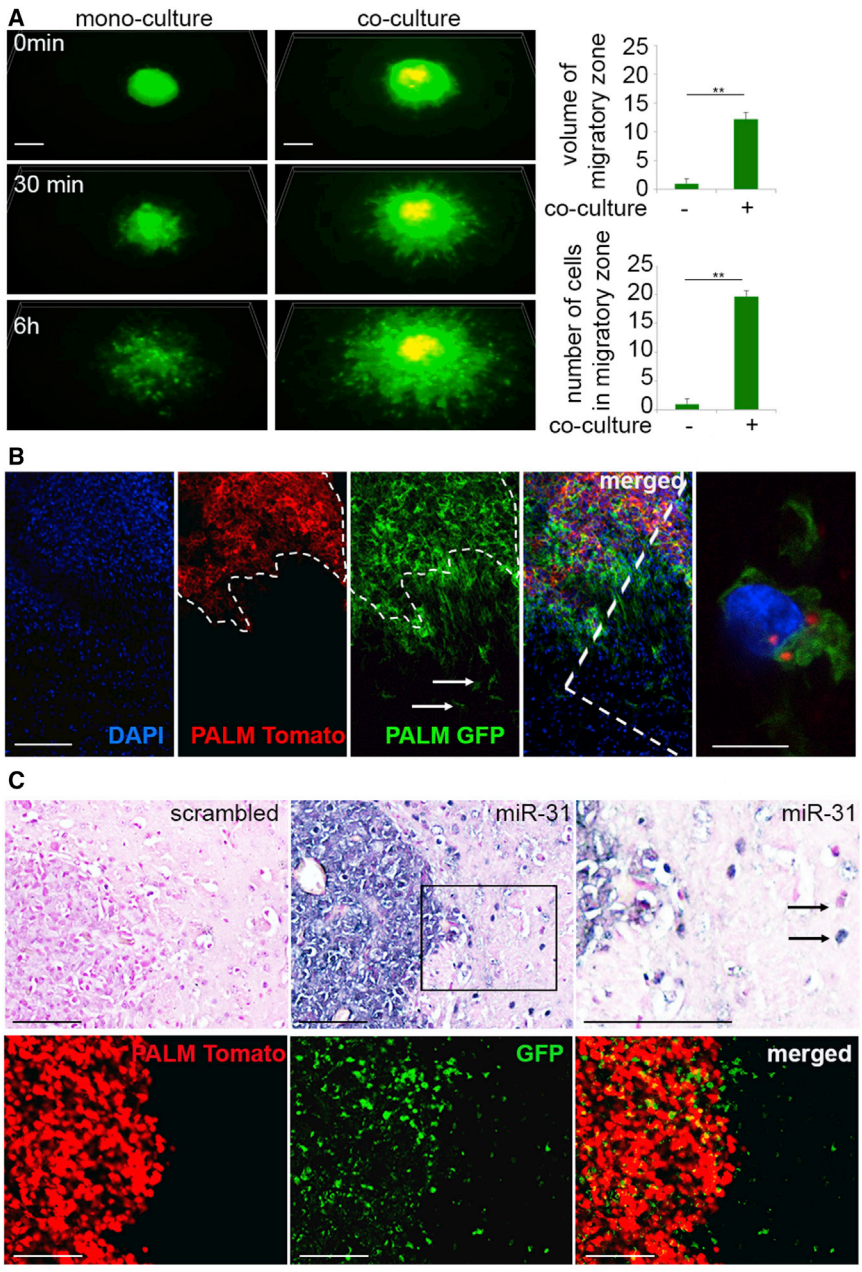


Figure 3. GSC EVs Support Subpopulation-Specific Invasive Phenotype

(A) Heterogeneity of GSC spheroids promotes migration of invasive GSCs. Representative micrographs of GSC spheroids in monoculture (n = 3 independent experiments per three independent GSCs) (left) and heterogeneous co-culture (n = 3 independent experiments per three independent GSCs) (middle) in 3D time-lapse frames are shown. GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs. Scale bar, 100 μm. Relative quantification of migratory zone volume (top right) and number of cells migrated out of spheroid core (bottom right) in mono- versus co-culture. **p < 0.01.

(B) Nodular and invasive phenotype of GSCs is recapitulated in vivo. Representative micrographs of co-implanted heterogeneous tumors (n = 6 independent experiments) are shown. GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs. Nodular tumor burden and infiltrating tumor cells are indicated by dashed line and arrows, respectively. Intratumoral EV transfer between cells is shown on high-power magnification inset. Scale bars, 150 μm and 10 μm.

(C) GSC EV is transferred intratumorally. Representative micrographs of co-implanted heterogeneous tumors (n = 3 independent experiments) are shown. Scrambled or microRNA ISH (nodular specific miR-31) (top) and GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs (bottom) from consecutive sections. Positive and negative microRNA detection is indicated by arrows. Scale bars, 100 μm.

See also Figure S3.

analyses suggested that distinction between mesenchymal and proneural molecular subclasses might be at least partially driven by the microRNA expression signatures.

Our data strongly implicated that EV-mediated transfer of bioactive molecules leads to increased heterogeneity not due to passive transfer but via cell-specific targeting, as both cellular and EV microRNAs have a cell-specific function, targeting effectors existing exclusively in particular GSC subpopulations. EV-microRNA transfer between different subpopulations of tumor cells should be thus recognized as an important aspect of tumor intricacy that may propagate heterogeneity of GBM; thus, EV-microRNA secretion and uptake may be an additional trait of cellular adaptation into different anatomic niches. Recent evidence suggests that the microRNA repertoire in EVs only partially mirrors that of cellular microRNA and, in fact, its specific pattern may be surprisingly different from that of secreting cells (Koppers-Lalic et al., 2014; Skog et al., 2008). Our data strongly support the existence of an active mechanism of microRNA loading or, rather, the co-existence of diverse mechanisms, as global,

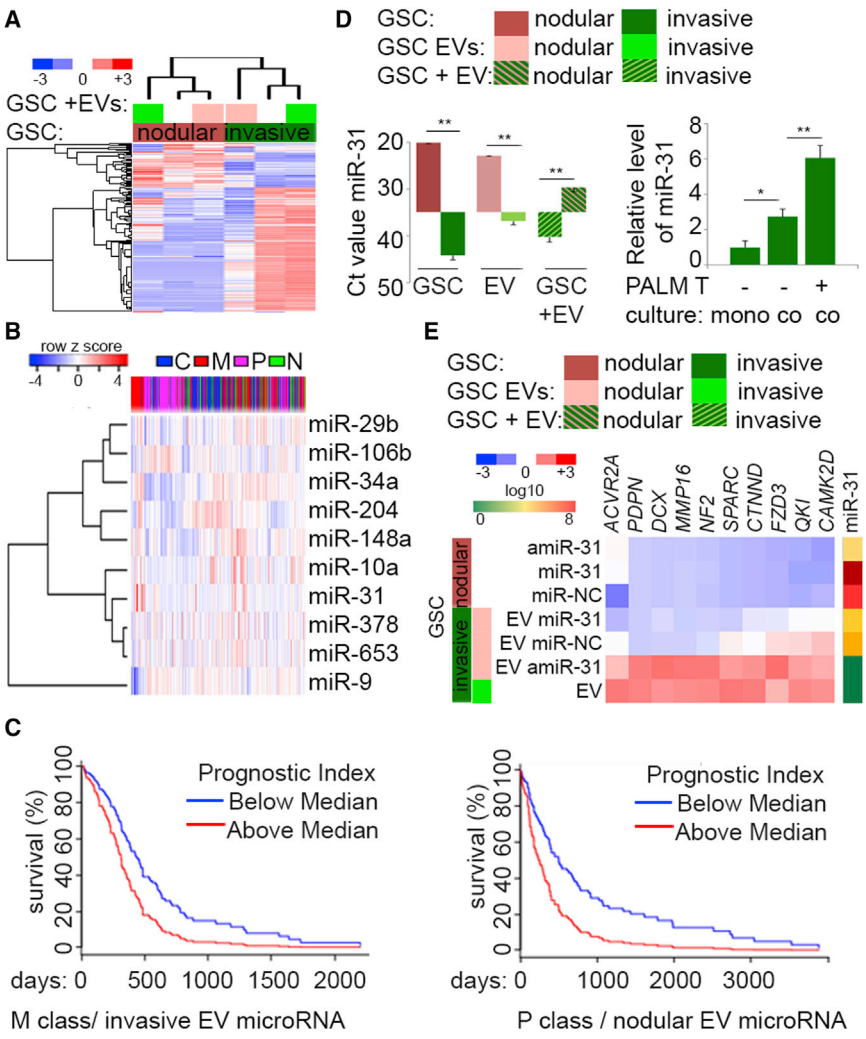


Figure 4. Transfer of EV-Encapsulated MicroRNAs Propagates GBM Heterogeneity

(A) Exchange of EV between distinct GSCs shifts subpopulation-specific microRNA signatures. Unsupervised hierarchical clustering of expression of 307 microRNAs in non-treated GSCs (n = 4 independent GSCs, n = 2 per subclass) and EV-treated GSCs (n = 4 independent EVs, n = 2 per subclass) is shown.

(B) MicroRNAs upregulated upon EV uptake are diversely expressed in GBM. MicroRNA sets that are coherently upregulated in nodular and invasive GSCs (n = 4 independent GSCs, n = 2 per subclass) upon treatment with EVs ($p < 0.05$, fold >2) were queried with TCGA-classified GBM dataset and identified by clustering with subtype prediction. C, classical; M, mesenchymal; P, proneural; N, neural.

(C) Survival analysis in mesenchymal (left) and proneural (right) GBM subtypes based on the impact of the prognostic index of multiple microRNAs (miR-148a, miR-204, miR-34a, miR-106b, and miR-9 [left], and miR-31, miR-653, miR-378a, miR-29b, and miR-10a [right]) based on retrospective data extrapolated from the TCGA. For mesenchymal GBM (n = 125 patient samples), log-rank p = 0.004, Prognostic Index hazard ratio = 1.83, p = 0.004. For proneural GBM (n = 112 patient samples) log-rank p = 0.001, Prognostic Index hazard ratio = 2.09, p = 0.001.

(D) MiR-31 is EV-transferred between GSC subpopulations. Left: qPCR analysis of miR-31 in donor nodular GSCs (n = 3 independent GSCs), their EVs (n = 3 independent GSC EVs), and recipient invasive GSCs (n = 3 independent GSCs). Right: monoculture spheroid of GFP-tagged invasive GSC (mono-) and co-culture spheroids of GFP-tagged invasive GSCs and PALM-Tomato (PALM T) nodular GSCs were sorted for GFP-positive (co-culture negative) or double-positive (GFP and Tomato [co-culture positive]). Data (n = 3 independent experiments) are shown as the mean raw Ct value ± SD, $^{**}p < 0.01$ (left); and as mean ± SD, $^{*}p < 0.05$, $^{**}p < 0.01$ (right).

(E) EV-miR-31 targets subclass GSC-specific genes. Nodular GSCs (n = 3 independent GSCs) were transfected with control (NC), microRNA mimic (miR-31), and microRNA inhibitor (amiR-31) (top three rows), and invasive GSC were treated with EVs derived from such nodular GSCs or their own EVs (n = 3 independent GSCs). qPCR analysis of selected targets and miR-31 is shown as hierarchical clustering and $\log_{10}$ assessed based on the value of expression, respectively.

See also Figure S4.

non-random distribution of microRNA was detected in subclasses of GSC EVs. The complexity of solid tumors, including GBM, and their distinct pathophysiology relies on anatomic niches that transmit and receive signals through cellular and acellular mediators (Jones and Wagers, 2008). These components are highly reliant on each other and undergo constant architectural, phenotypic, and transcriptomic rearrangements depending on fluctuating microenvironmental contexts as the disease progresses. The brain tumor "ecosystem" is composed of distinct phenotypic and transcriptomic cell components, and our analysis of cellular and EV microRNA load discovered additional aspects of intratumoral diversity. EVs/microRNA as transcriptome and signaling communication tandem modulators arrange both molecular and phenotypic traits. We thus argue that observed highly heterogeneous profiles of microRNA expression in GBM and the co-existence of diverse subtypes and hybrid-stage cells within individual tumors is propagated by intratumoral exchange of microRNA.

EXPERIMENTAL PROCEDURES

Human Specimens and Primary Cells

Tumor samples were obtained as approved by The Harvard Medical School (HMS) Institutional Review Board. Surgery was performed by E.A.C. and I.N. GSCs were obtained by dissociation of tumor samples and cultivated in stem cell-enriching conditions. The unique identity of cultured patient-derived cells (n = 12) was confirmed by short tandem repeat analysis (Kim et al., 2016).

Purification of EVs

The conditioned media were collected, and EVs were isolated by differential centrifugation and analyzed using a NanoSight.

MicroRNA/mRNA Expression Analysis

Nanostring microRNA technology was used to search for unique microRNA signatures in GSC and GSC EVs. Whole human genome oligo microarray was performed by Arraystar. The SBI's Exo-NGS service was used to build the Illumina NGS libraries followed sequencing using a 1 × 50-bp single-end Illumina HiSeq NGS and Maverix Analytic Platform.

In Vitro Assays

Nodular GSCs were labeled with either RFP or GFP (using the lentiviral pCDH vector), or PALM-Tomato (using the lentiviral CSCW2 vector [Lai et al., 2015]); and invasive GSCs were labeled with GFP or PALM-GFP.

For EV transfer, 3×10^5 GSCs/mL were maintained overnight in unsupplemented medium, followed by 24 hr of treatment with EVs.

For 3D spheroid dispersal assay, GSCs were dissociated to single cells using Accutase (Life Technologies), and plated at 200 (noninvasive GSCs) or 1,000 (invasive GSCs) cells/well in a 96-well plate for 48 hr. Spheroids were then transferred into collagen with unsupplemented medium or medium supplemented with EVs and analyzed after 0–6 hr or in time-lapse.

For co-culture assays, single-cell suspensions of co-culture (at ratio 1:4) were cultured for 48 hr before assay. Monocultured invasive GSC spheroids served as controls for co-cultured invasive/nodular spheroids followed by sorting for populations of pure PALM-Tomato (nodular), GFP (invasive), and double-positive (invasive with nodular PALM-Tomato EVs).

In Vivo Studies

Female athymic mice were purchased from Envigo. Mice were housed in the HMS animal facility in accordance with NIH regulations. Protocols were approved by the HMS Institutional Animal Care and Use Committee. Intracranial tumor injection was performed as described by Ricklefs et al. (2016). For GSC implantation and co-implantation experiments, either 1×10^3 nodular GSCs or 5×10^5 invasive GSCs, or both combined were used.

Data Analysis

Functional bioinformatics analyses were performed using Qiagen's Ingenuity Pathway Analysis (IPA; www.qiagen.com/ingenuity). Experimental and clinical data were analyzed using the GBM-BioDP (Celiku et al., 2014). Clinical data were downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov/). Gene expression in the various anatomical regions of glioblastoma was analyzed using the Ivy Glioblastoma Atlas Project (http://glioblastoma.alleninstitute.org/).

Level 3 microRNA expression data (unc.edu_GBM.H-miRNA_8 × 15K.Level_3.1.8.0) from 479 glioblastomas were obtained from TCGA (Brennan et al., 2013). An unpaired, two-tailed t test was used to compare two groups. One-way ANOVA, followed by Bonferroni's test, was conducted to test for significance among multiple groups. Fisher's exact test was used to identify associations between unsupervised microRNA classes and previously determined GBM subtypes. $p < 0.05$ was considered significant.

ACCESSION NUMBERS

The accession number for the gene microarray reported in this paper is GEO: GSE89501.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, two tables, and two movies and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2017.04.024.

AUTHOR CONTRIBUTIONS

All authors agree to be accountable for all aspects of the work.R.F.-L. analyzed TCGA data. Y.S.T., A.K.R., and F.R. performed in vitro and in vivo assays. M.M. characterized EV cargo. M.O.N. performed microscope analysis. E.S., H.L., R.W., and R.B. assisted with writing the manuscript and interpretation of data. I.N. acquired patients' specimens. E.A.C. acquired patients' specimens and assisted with writing the manuscript. J.G. and A.B. conceived and designed the overall work, analyzed and interpreted data, and wrote the manuscript.

ACKNOWLEDGMENTS

This research was supported by NCI P01 CA069246, NCI P01 CA069246-20 (both to E.A.C.), and NCI 1R01 CA176203-01A1 (to J.G.).

Received: December 12, 2016
Revised: April 20, 2017
Accepted: April 21, 2017
Published: May 18, 2017

REFERENCES

Brennan, C.W., Verhaak, R.G., 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.

Bronisz, A., Wang, Y., Nowicki, M.O., Peruzzi, P., Ansari, K.I., Ogawa, D., Balaj, L., De Rienzo, G., Mineo, M., Nakano, I., et al. (2014). Extracellular vesicles modulate the glioblastoma

microenvironment via a tumor suppression signaling network directed by miR-1. Cancer Res. *74*, 738–750.

Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2004). Metagenes and molecular pattern discovery using matrix factorization. Proc. Natl. Acad. Sci. USA *101*, 4164–4169.

Celiku, O., Johnson, S., Zhao, S., Camphausen, K., and Shankavaram, U. (2014). Visualizing molecular profiles of glioblastoma with GBM-BioDP. PLoS One *9*, e101239.

Du, Z., Fei, T., Verhaak, R.G., Su, Z., Zhang, Y., Brown, M., Chen, Y., and Liu, X.S. (2013). Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. Nat. Struct. Mol. Biol. *20*, 908–913.

Godlewski, J., Krichevsky, A.M., Johnson, M.D., Chiocca, E.A., and Bronisz, A. (2015). Belonging to a network—microRNAs, extracellular vesicles, and the glioblastoma microenvironment. Neuro Oncol. *17*, 652–662.

Johnson, D.R., and O'Neill, B.P. (2012). Glioblastoma survival in the United States before and during the temozolomide era. J. Neuro Oncol. *107*, 359–364.

Jones, D.L., and Wagers, A.J. (2008). No place like home: anatomy and function of the stem cell niche. Nat. Rev. Mol. Cell Biol. *9*, 11–21.

Kim, S.H., Ezhilarasan, R., Phillips, E., Gallego-Perez, D., Sparks, A., Taylor, D., Ladner, K., Furuta, T., Sabit, H., Chhipa, R., et al. (2016). Serine/threonine kinase MLK4 determines mesenchymal identity in glioma stem cells in an NF-kappaB-dependent manner. Cancer Cell *29*, 201–213.

Koppers-Lalic, D., Hackenberg, M., Bijnsdorp, I.V., van Eijndhoven, M.A., Sadek, P., Sie, D., Zini, N., Middeldorp, J.M., Ylstra, B., et al. (2014). Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. Cell Rep. *8*, 1649–1658.

Lai, C.P., Kim, E.Y., Badr, C.E., Weissleder, R., Mempel, T.R., Tannous, B.A., and Breakefield, X.O. (2015). Visualization and tracking of tumour extracellular vesicle delivery and RNA translation using multiplexed reporters. Nat. Commun. *6*, 7029.

Mao, P., Joshi, K., Li, J., Kim, S.H., Li, P., Santana-Santos, L., Luthra, S., Chandran, U.R., Benos, P.V., Smith, L., et al. (2013). Mesenchymal glioma stem cells are maintained by activated glycolytic metabolism involving aldehyde dehydrogenase 1A3. Proc. Natl. Acad. Sci. USA *110*, 8644–8649.

Mineo, M., Ricklefs, F., Rooj, A.K., Lyons, S.M., Ivanov, P., Ansari, K.I., Nakano, I., Chiocca, E.A., Godlewski, J., and Bronisz, A. (2016). The long non-coding RNA HIF1A-AS2 facilitates the maintenance of mesenchymal glioblastoma stem-like cells in hypoxic niches. Cell Rep. *15*, 2500–2509.

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.

Phillips, H.S., Kharbanda, S., Chen, R., Forrest, W.F., Soriano, R.H., Wu, T.D., Misra, A., Nigro, J.M., Colman, H., Soroceanu, L., et al. (2006). Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell *9*, 157–173.

Ricklefs, F., Mineo, M., Rooj, A.K., Nakano, I., Charest, A., Weissleder, R., Breakefield, X.O., Chiocca, E.A., Godlewski, J., and Bronisz, A. (2016). Extracellular vesicles from high-grade glioma exchange diverse pro-oncogenic signals that maintain intratumoral heterogeneity. Cancer Res. *76*, 2876–2881.

Schonberg, D.L., Lubelski, D., Miller, T.E., and Rich, J.N. (2014). Brain tumor stem cells: molecular characteristics and their impact on therapy. Mol. Aspects Med. *39*, 82–101.

Skog, J., Wurdinger, T., van Rijn, S., Meijer, D.H., Gainche, L., Sena-Esteves, M., Curry, W.T., Jr., Carter, B.S., Krichevsky, A.M., and Breakefield, X.O. (2008). Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nat. Cell Biol. *10*, 1470–1476.

Verhaak, R.G., 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.

Wong, H.K., Fatimy, R.E., Onodera, C., Wei, Z., Yi, M., Mohan, A., Gowrisankaran, S., Karmali, P., Marcusson, E., Wakimoto, H., et al. (2015). The cancer genome atlas analysis predicts MicroRNA for targeting cancer growth and vascularization in glioblastoma. Mol. Ther. *23*, 1234–1247.

Stem Cell Reports, Volume *8*

Supplemental Information

MicroRNA Signatures and Molecular Subtypes of Glioblastoma: The Role of Extracellular Transfer

Jakub Godlewski, Ruben Ferrer-Luna, Arun K. Rooj, Marco Mineo, Franz Ricklefs, Yuji S. Takeda, M. Oskar Nowicki, Elżbieta Salińska, Ichiro Nakano, Hakho Lee, Ralph Weissleder, Rameen Beroukhim, E. Antonio Chiocca, and Agnieszka Bronisz

Figure S1

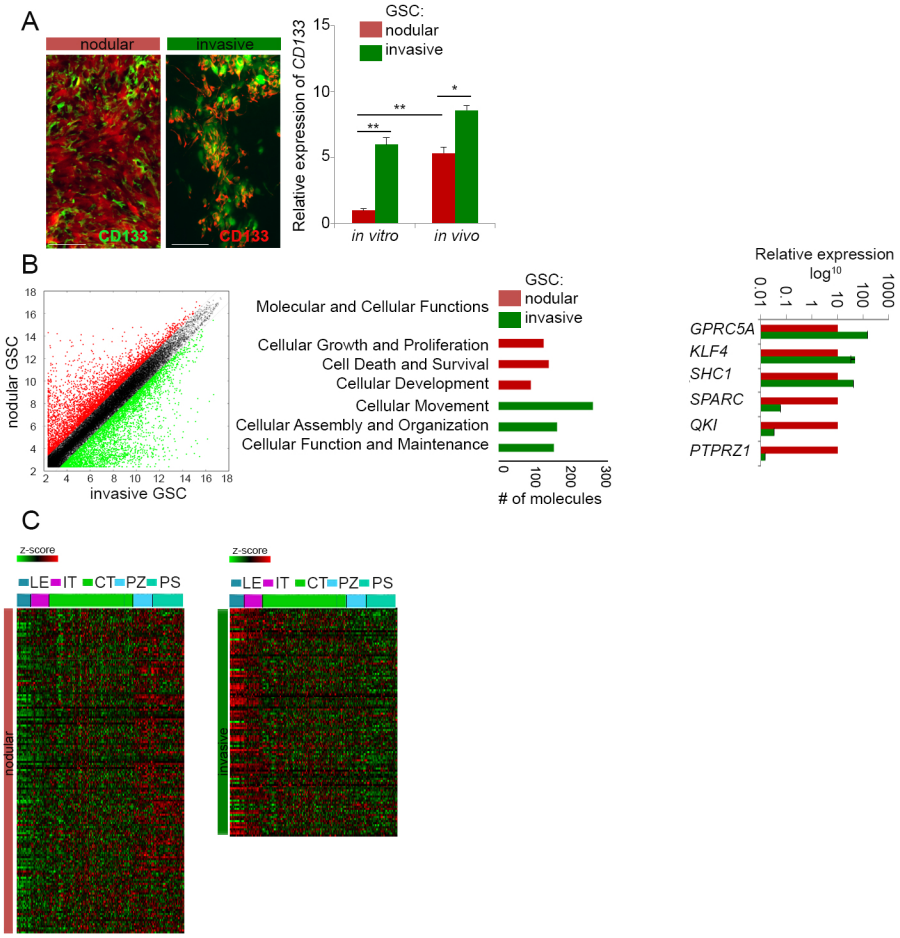
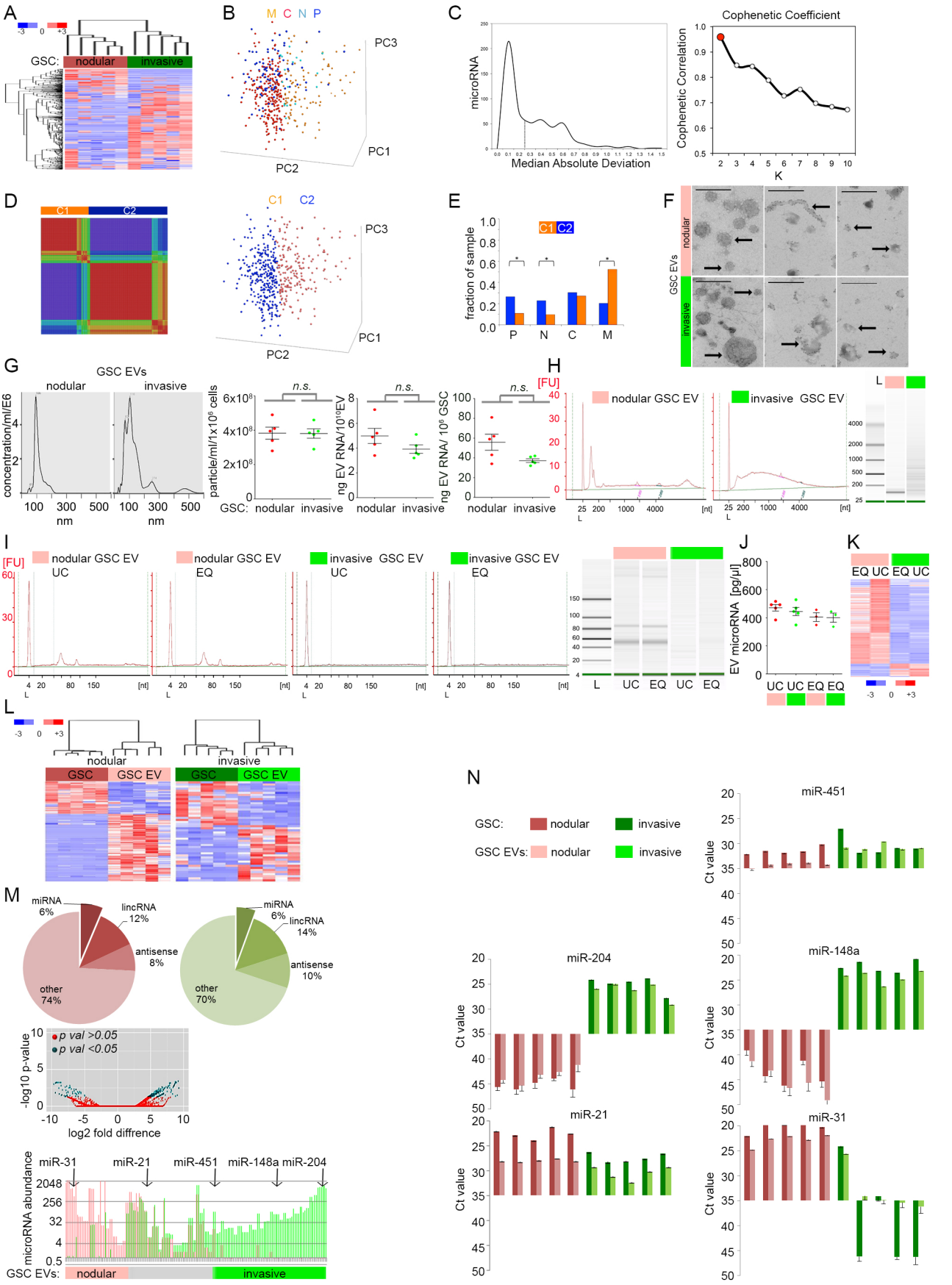

A
nodular
invasive
CD133
CD133
GSC:
nodular
invasive
Relative expression of CD133
15
10
5
0
**
**
*
in vitro
in vivo
B
nodular GSC
invasive GSC
18
16
14
12
10
8
6
4
2
2 4 6 8 10 12 14 16 18
GSC:
nodular
invasive
Molecular and Cellular Functions
Cellular Growth and Proliferation
Cell Death and Survival
Cellular Development
Cellular Movement
Cellular Assembly and Organization
Cellular Function and Maintenance
0
100
200
300
# of molecules
Relative expression
log10
0.01
0.1
1
10
100
1000
GPRC5A
KLF4
SHC1
SPARC
QKI
PTPRZ1
C
z-score
LE IT CT PZ PS
nodular
z-score
LE IT CT PZ PS
invasive

Figure S2


A
GSC:
nodular
invasive
B
M C N P
PC3
PC1
PC2
C
microRNA
Median Absolute Deviation
Cophenetic Coefficient
Cophenetic Correlation
K
D
C1
C2
E
fraction of sample
P
N
C
M
F
GSC EVs
nodular
invasive
G
GSC EVs
concentration/ml/E6
particle/ml/1x10^6 cells
ng EV RNA/10^10EV
ng EV RNA/ 10^6 GSC
n.s.
H
[FU]
nodular GSC EV
invasive GSC EV
I
UC
EQ
J
EV microRNA [pg/µl]
K
L
GSC
GSC EV
M
miRNA 6%
lincRNA 12%
antisense 8%
other 74%
lincRNA 14%
antisense 10%
other 70%
p val >0.05
p val <0.05
-log10 p-value
log2 fold diffrence
microRNA abundance
miR-31
miR-21
miR-451
miR-148a
miR-204
N
GSC:
GSC EVs:
Ct value

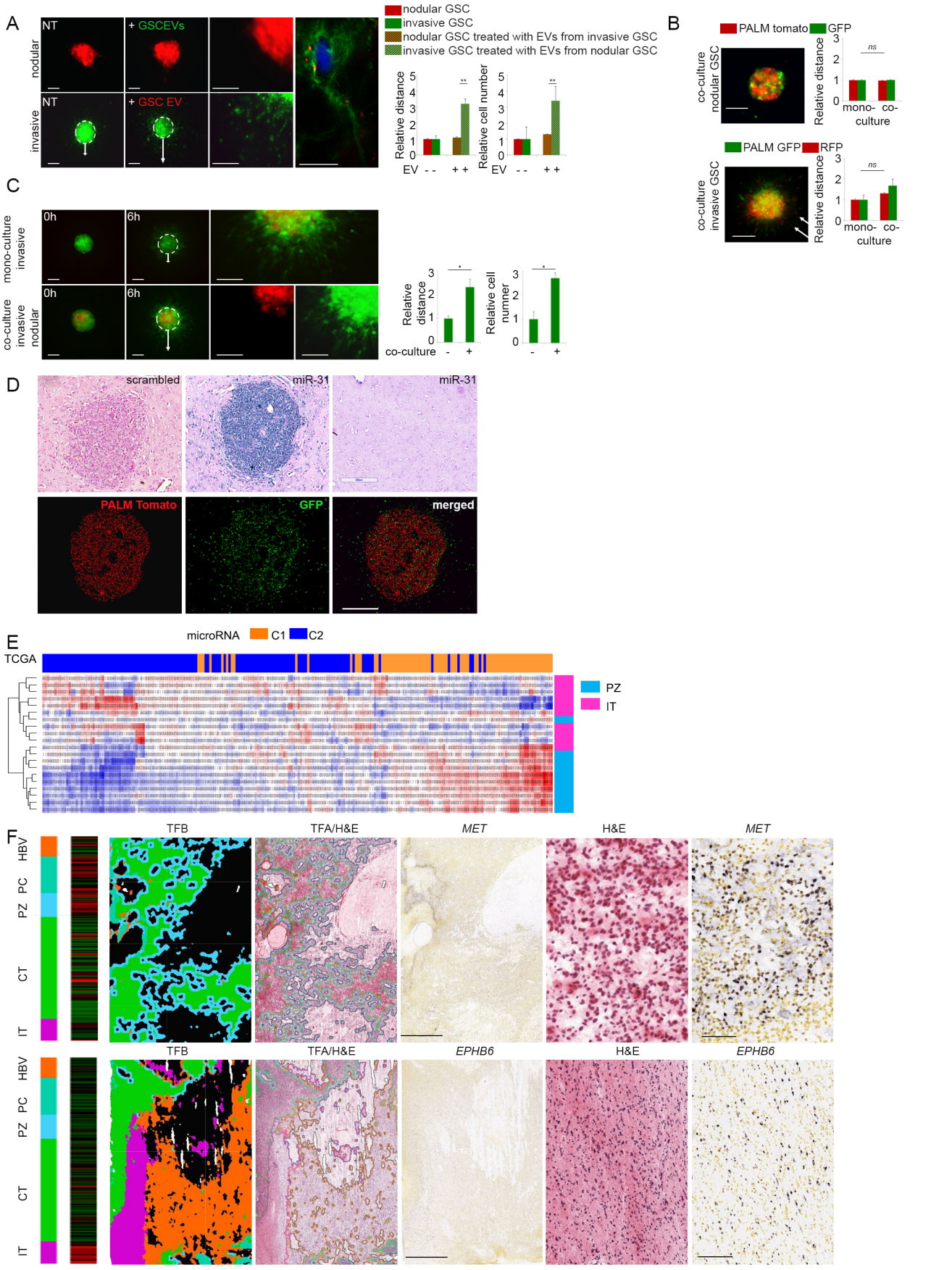
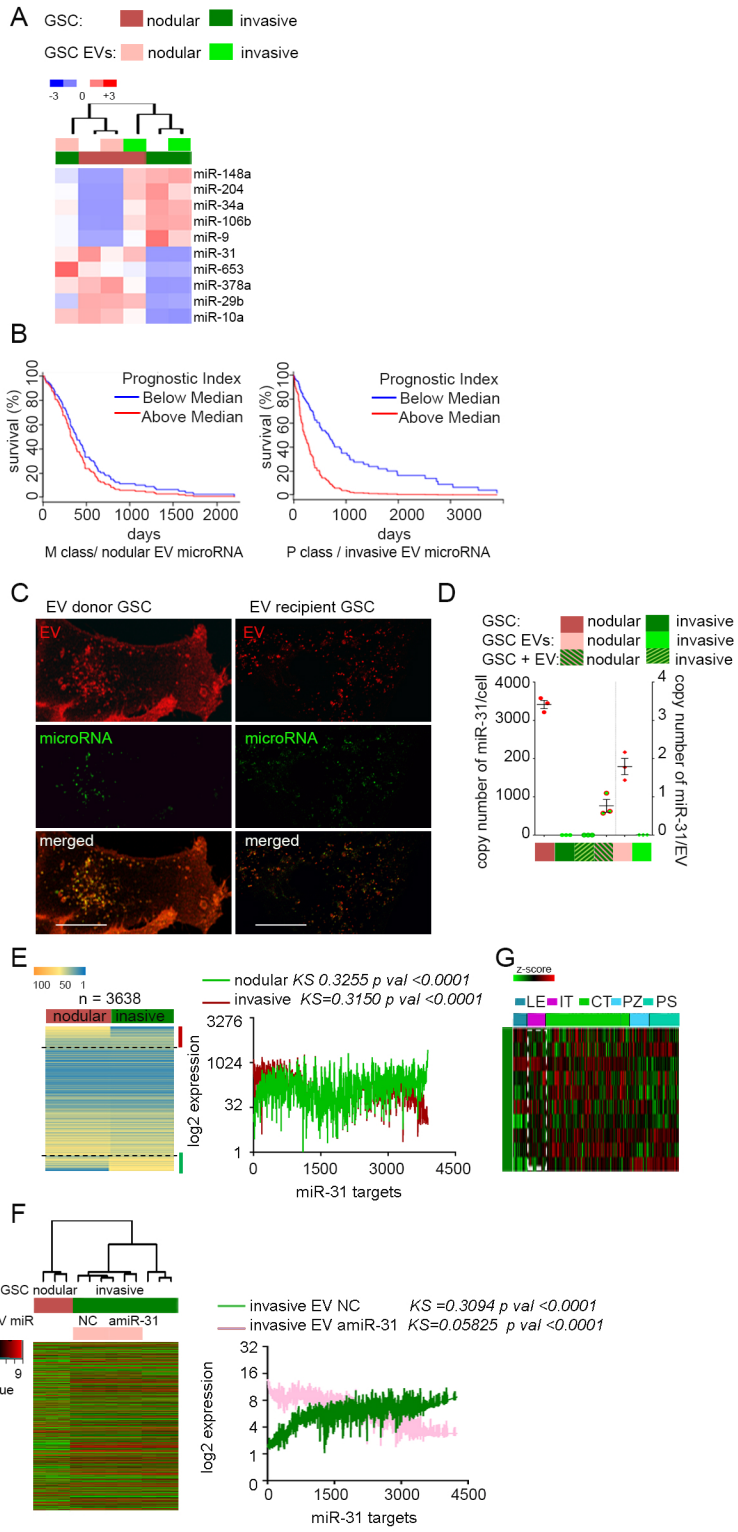

A
NT
+ GSCEVs
NT
+ GSC EV
nodular
invasive
nodular GSC
invasive GSC
nodular GSC treated with EVs from invasive GSC
invasive GSC treated with EVs from nodular GSC
Relative distance
Relative cell number
EV - -
+ +
B
PALM tomato
GFP
co-culture nodular GSC
Relative distance
ns
mono-
co-culture
PALM GFP
RFP
co-culture invasive GSC
C
0h
6h
mono-culture invasive
co-culture invasive nodular
Relative distance
Relative cell numner
co-culture
D
scrambled
miR-31
miR-31
PALM Tomato
GFP
merged
E
microRNA
C1
C2
TCGA
PZ
IT
F
TFB
TFA/H&E
MET
H&E
MET
HBV
PC
PZ
CT
IT
TFB
TFA/H&E
EPHB6
H&E
EPHB6

Figure S4


A
GSC: nodular invasive
GSC EVs: nodular invasive
-3 0 +3
miR-148a
miR-204
miR-34a
miR-106b
miR-9
miR-31
miR-653
miR-378a
miR-29b
miR-10a
B
Prognostic Index
Below Median
Above Median
survival (%)
days
M class/ nodular EV microRNA
P class / invasive EV microRNA
C
EV donor GSC
EV recipient GSC
EV
microRNA
merged
D
GSC:
GSC EVs:
GSC + EV:
nodular
invasive
copy number of miR-31/cell
copy number of miR-31/EV
E
n = 3638
nodular
invasive
nodular KS 0.3255 p val <0.0001
invasive KS=0.3150 p val <0.0001
log2 expression
miR-31 targets
G
z-score
LE IT CT PZ PS
F
GSC nodular invasive
EV miR NC amiR-31
count
value
invasive EV NC KS =0.3094 p val <0.0001
invasive EV amiR-31 KS=0.05825 p val <0.0001
log2 expression
miR-31 targets

SUPPLEMENTAL FIGURES and DATA LEGEND

Figure S1 Supplemental to Figure 1

A. Representative micrographs of GSC-originated intracranial tumors with CD133 immunostaining (left, scale bar: 50 μm), and qPCR analysis of CD133 expression (right, n = 6 independent GSCs). Data shown as mean ± SD, * $p < 0.05$, ** $p < 0.01$.

B. Gene expression analysis in nodular vs. invasive GSCs (n=8 independent GSCs, 4 per subclass) by scatter plot (left); molecular and cellular function of genes significantly deregulated ($p< 0.05$, fold>2) by IPA analysis (middle); qPCR analysis of the top three genes for each GSC subclass (right). Data shown as mean ± SD, all data $p < 0.01$.

C. The positively correlated genes in GSCs (n=8 independent GSCs, 4 per subclass) in each category (pro-proliferative n=160 genes , or pro-invasive n=124 genes) were queried with Ivy GAP database-based expression signatures and identified by clustering with different anatomic areas of GBM (left; LE - leading edge; IT - infiltrating tumor; CT - cellular tumor; PZ - perinecrotic zone; PS - pseudopalisading cells).

Figure S2. Supplemental to Figure 2

A. MicroRNA expression in GSCs was identified by unsupervised clustering (n = 10 independent GSCs, 5 per subclass).

B. Unsupervised principal component analysis (PCA) based on expression of 221 microRNAs in TCGA GBM (n=479 patient samples) samples labeled by molecular classes (C - classical; M - mesenchymal; P - proneural; N – neural). The main principal components projected in tridimensional space bore about 90% of variation.

C. Median absolute deviation (MAD) distribution (left) for 534 microRNAs in 479 TCGA GBM samples (n=479 patient samples), note that 221 miRNA scored above established threshold (0.25).

Cophenetic coefficients for 10 hierarchical clustered matrices (right) obtained from miRNA (221) displaying enough variation. Note that K=2 suggests most two robust classes.

D. Selected model of consensus negative matrix factorization (K=2) showing a robust two-class partition (left). Principal component analysis (right) based on expression of 221 microRNAs in TCGA

GBM (n=479 patient samples) labeled for unsupervised classes (C1 and C2) obtained after perform unsupervised NMF analysis.

E. Distribution of GBMs (n=479 patient samples) and associations between previously determined molecular subtypes (C - classical; M - mesenchymal; P - proneural; N – neural) and unsupervised microRNA classes (C1 and C2) (Significant associations * $p<0.001$, Fisher exact test)

F. GSCs secrete heterogeneous EVs. Representative micrographs of GSC EVs (n=6 independent GSC EVs). Arrows indicate diverse size (top), shape (middle) and differential expression of CD63 (bottom). Scale bar: 500 nm.

G. Analysis of frequency and size of GSC-derived EVs (n=10 independent GSCs, 5 per subclass) by NanoSight (left). Data quantification (right) is shown as mean ± SD.

H. Total RNA derived from GSC EVs (n=10 independent GSC EVs, 5 per subclass) was profiled using a total RNA platform Agilent Bioanalyzer. Representative RNA profiles are shown.

I. Total RNA derived from GSC EVs by ultracentrifugation (UC) (n=8 independent GSC EVs, 4 per subclass) or ExoQuick (EQ) (n=6 independent GSC EVs, 3 per subclass) was profiled using a small RNA platform Agilent Bioanalyzer. Representative RNA profiles are shown.

J. Analysis of EV microRNA content in small RNA fraction using a small RNA platform Agilent Bioanalyzer in GSC-derived EVs isolated by ultracentrifugation (UC) (n=8 independent GSC EVs, 4 per subclass) or ExoQuick (EQ) (n=6 independent GSC EVs, 3 per subclass). Data is shown as mean ± SD.

K. Analysis of microRNA signature in nodular and invasive GSC EVs isolated by ultracentrifugation (UC) (n=4 independent GSC EVs, 2 per subclass) or ExoQuick (EQ) (n=4 independent GSC EVs, 2 per subclass). Data is shown as a heat map.

L. Hierarchical clustering of 119 microRNAs in nodular (left) (n=5 independent GSC) and 49 microRNAs in invasive (right) GSCs (n=5 independent GSC EVs) and their EVs is shown in supervised analysis ($p < 0.05$, fold >2).

M. Deep sequencing analysis of GSC EV microRNAs correlates with NS data and distinguishes between EVs derived from different subpopulation of GSCs (n=2 independent GSC EVs). Non-coding RNA fraction in GSC EV (top). MicroRNA changes between EVs derived from nodular vs. invasive GSC (middle) and microRNA abundance (bottom) are shown. Selected microRNAs indicated by arrows.

N. Expression of selected microRNAs in GSCs and GSC EVs. Value based on qPCR in GSC and GSC EV is shown (n=5 independent GSCs or EVs per subclass). Data shown as mean ± SD.

Figure S3. Supplemental to Figure 3

A. Representative micrographs of GSC spheroids (n=6 independent GSCs, 3 per subclass) (GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs) co-cultured with EVs; intraspherical EV uptake is shown on high power magnification insets. Scale bars: 100 μm, 50 μm, 10 μm (left). Relative quantification of average distance of cells migrating out of the spheroid core and the number of cells that migrated out of the spheroid core (right). Data shown as mean ± SD * $p< 0.05$.

B. Representative micrographs of nodular (top) or invasive (bottom) GSC spheroids (co-culture of PALM-GFP and PALM-TOMATO cells from one subclass) (n=3 independent experiments), scale bars: 100 μm. Relative quantification of migratory zone (for GSCs monoculture vs. co-culture) that migrated out of the spheroid core (right). Data shown as mean ± SD $p>0.05$.

C. Representative micrographs of GSC spheroids (n=6 independent GSCs, 3 per variant) (GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs) in mono- and heterogeneous co-culture, scale bars: 100 μm and 50 μm (right). Relative quantification of migratory zone volume and number of cells that migrated out of the sphere core (right). Data shown as mean ± SD * $p < 0.05$.

D. Representative micrographs of co-implanted heterogeneous tumors (n=3 independent experiments) are shown. Scrambled or microRNA ISH (nodular-specific miR-31) (top) and GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs (bottom) from consecutive sections. Tissue staining 5 mm away from tumor is shown on top left. Scale bars: 200 μm.

E. The C1-C2 microRNA signature separates infiltrating and perinecrotic zones. Expression correlation between microRNAs and genes from different tumor anatomic sites. Heatmap of the correlation between the microRNA and differentially expressed genes in each category (pro-proliferative n=10 genes, or pro-invasive n=10 genes) (see Figure 1C) and its association with infiltrating tumor (IT) and perinecrotic (PZ) zones of the tumor and C1-C2 class of microRNA. Genes (left cluster) and microRNAs (top cluster) were retrieved from TCGA. Red color represents positive correlation, whereas blue color represents negative correlation ($p<0.05$).

F. Anatomic Structures for Selected Genes. Left: The selected, representative genes from each category (pro-proliferative (*MET*, top) and pro-invasive (*EPHB6*, bottom) (r=-0.09)) were queried with Ivy GAP database-based expression signature (n=42 patients samples) and identified by clustering with different anatomic areas of GBM (LE - leading edge; IT - infiltrating tumor; CT - cellular tumor;

PZ - perinecrotic zone; PS - pseudopalisading cells). Right: tumor feature boundary (TFB), tumor feature annotation (TFA) with H&E, ISH (scale: 1600 microns) magnification of H&E and ISH (200) microns.

Figure S4. Supplemental to Figure 4

A. MicroRNA sets that were coherently upregulated in invasive GSCs upon treatment with EVs derived from nodular GSCs were identified by supervised clustering (n=2 per subclass, fold >2, $p < 0.05$).

B. Survival analysis in mesenchymal (left) and proneural (right) GBM subtypes based on the impact of the prognostic index of multiple microRNAs (miR-31, miR-653, miR-378a, miR-29b, and miR-10a (left), and miR-148a, miR-204, miR-34a, miR-106b, and miR-9 (right) based on retrospective data extrapolated from the TCGA. For mesenchymal GBM (n=125 patient samples): Log rank p= 0.192; for proneural GBM (n=112 patient samples): Log rank p= not determined (0.000).

C. EV-microRNAs are transferred between GSC subpopulations. Representative micrographs (n=6 independent experiments) of EV donor PALM-Tomato GSCs expressing FAM-microRNA (left) and EV recipient GSCs (right). Scale bars: 10 μm.

D. EV miR-31 is efficiently transferred into recipient cells. Copy number of miR-31 in donor (nodular GSCs) and recipient (invasive GSCs treated with nodular derived EV or their own EVs) and EVs (n=3 independent GSCs or EVs) was quantified using an NS spike control.

E. Putative miR-31 targets are differentially expressed in GSC subpopulations. Expression of 3638 transcripts with 3′ UTR elements harboring matches to position 2–8 of the seed region of the miR-31 in nodular vs. invasive GSCs was assessed based on the value of expression (dashed line show cut off: fold difference >2, $p < 0.05$) (left) and correlation analysis by linear regression plot with KS test ($p > 0.0001$) (right) based on microarray data (n=8 independent GSCs, 4 per subclass).

F. Putative miR-31 targets are deregulated upon EV treatment. Hierarchical clustering and heat map (left) was performed based on "All Targets Value" in nodular and invasive GSCs and invasive GSCs treated with either EVs derived from nodular GSCs expressing either negative control (NC) or miR-31 inhibitor (amiR-31). The result of hierarchical clustering shows distinguishable gene expression profiling (n=12 independent GSCs, 3 per variant). Expression of detected 3392 transcripts with 3′ UTR elements harboring matches to position 2–8 of the seed region of the miR-31 in invasive GSCs treated with EVs derived from nodular GSCs expressing either negative control (NC) or miR-31

inhibitor (amiR-31) was assessed based on correlation analysis by linear regression plot with KS test (p>0.0001) (right) based on microarray data, (n=6 independent GSCs, 3 per variant).

G. Top 10 putative targets of miR-31 de-regulated upon EV treatment in invasive GSCs were queried with Ivy GAP database-based expression signature (n=42 patient sample) in different areas of GBM.

Table S1 Supplemental dataset to Figure 1B, C and Figure S1B

Curated list of genes from relevant biological classes: "cellular growth and proliferation" (p= 1.93E-14 - 3.00E-04) and "cellular movement" (p= 3.16E-31 - 2.22E-06). The log ratio difference between invasive and nodular GSCs.

Table S2. Supplemental dataset to Figure 2B

List of microRNA/mRNA targets with positive correlation in nodular GSCs and in invasive GSCs.

Movie S1. Supplemental to Figure S3A

Migration of invasive GSCs is supported by EVs. Representative time-lapse movie of GSC cultured in un-supplemented medium that were non-treated (left) or EV-treated (right).

Movie S2. Supplemental to Figure 3A and Figure 3SB

Heterogeneous GSC sphere promote migration of invasive GSCs. Representative time-lapse movie of GSC in monoculture (left) and heterogeneous co-culture spheroids (right) in supplemented medium. PALM-GFP-labeled invasive GSCs and PALM-Tomato-labeled nodular GSCs.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Detailed methods include the following:

- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS

- *Human specimens and primary cells*

Tumor samples were obtained as approved by Harvard Medical School (HMS) and Ohio State University (OSU) IRBs. Surgery was performed by E.A.C. and I.N. Primary human GSCs (G2, G6, G12, G33, G34, G44, G62, G88, G146, G157, G91, G816) were obtained by dissociation of tumor samples and cultivation in stem cell-enriching conditions (Bronisz et al., 2014; Mao et al., 2013) as neurospheres using Neurobasal (Gibco) supplemented with 1% glutamine (Gibco), 2% B27 (Gibco) and 20 ng/mL EGF and FGF–2 (PeproTech). The unique identity of cultured patient-derived cells was confirmed by short tandem repeat analysis - by University of Arizona Genetics Core, Arizona Research Laboratories, Division of Biotechnology, http://uagc.arl.arizona.edu/ (Kim et al., 2016). The tumorigenicity of GSCs used in this study was confirmed by intracranial xenografts (Mao et al., 2013; Miyazaki et al., 2012; Nakano et al., 2011). of Arizona Genetics Core, Arizona Research Laboratories, Division of Biotechnology, http://uagc.arl.arizona.edu/

- *Constructs*

Palmitoylation sequences of growth cone-associated protein (GAP43) were genetically fused to the NH_2 terminus of GFP (PalmGFP) and tdTomato (PalmtdTomato) and cloned into lentiviral constructs as described (Lai et al., 2015). These vectors in combination with lentiviral packaging constructs were used to establish stable lines (Godlewski et al., 2010a). GSCs stably transduced with PALM-Tomato (nodular GSC) and PALM-GFP (invasive GSC) were sorted using fluorescence activated cell sorting by the Harvard Medical School (HMS) core facility, http://immdiv.hms.harvard.edu/FlowCore/Flowcore.html

- *Purification of EVs*

The conditioned media were collected and EVs were isolated by differential centrifugation and analyzed using a NanoSight as described (Bronisz et al., 2014). Briefly, media were centrifuged at 500*g* for 10 min to eliminate cells and at 16,500*g* for 20 min, followed by filtration through a 0.22 μm filter (Millipore). EVs were pelleted by ultracentrifugation at 120,000*g* for 70 min. For treatment experiments and RNA analysis EV pellets were washed once in PBS. As control for ultracentrifugation method, the ExoQuick purification (System Bioscience) was used according to manufacturer's protocol. EV numbers were determined using a Nanosight LM10 nanoparticle characterization system (NanoSight). All nanoparticle tracking analyses (NTAs) were done with identical experiment settings in

triplicates. Particles were measured for 60s and for optimal results concentrations were adjusted to obtain ~50 particles in the field of view (Shao et al., 2012).

- *Microscopy*

The Transmission Electron microscope Tecnai G² Spirit BioTWIN with AMT 2k CCD camera was used to analyze EVs stained with immunogold-labeled anti-CD63 antibody in Electron Microscopy Facility at Harvard Medical School http://electron-microscopy.hms.harvard.edu/. Ultrathin frozen sections and immunogold-labeled anti-CD63 antibody were prepared in Cellular Neuroscience Core Laboratory http://harvard.eagle-i.net/i/0000012b-00c7-c357-db6e-7a3f80000000.

All fluorescent and light microscopy based assays were monitored using a Zeiss LSM510 and Nikon Eclipse Ti.

MicroRNA EV uptake was monitored using donor cells stably labeled with PALM tomato and transiently transfected with FAM-labeled microRNA (Ambion). 24 h after transfection conditioned media from donor cells ($1*10^6$/1ml medium) were harvested by centrifugation at 200g for 10 minutes, supernatant filtrated and added to recipient cells in ratio 1:1 (1ml of media from donor cell per $1*10^6$ recipient cells) without purification or concentration. Recipient cells grew in glass microscopic slides.

EV uptake in neurospheres was visualized by embedding either PALM-Tomato nodular GSC and/or PALM-GFP invasive GSC spheres as mono- or co-culture spheroids in collagen (Advanced Biomatrix) and visualized by confocal microscopy. To visualize EV transfer within a spheroid, they were fixed by PFA/sucrose and cryo-sectioned (Ricklefs et al., 2016).

For *in vivo* EV uptake, brain sections were processed as described (Bronisz et al., 2014; Mao et al., 2013; Miyazaki et al., 2012; Nakano et al., 2011). Briefly, brain tissue was placed in 4% paraformaldehyde for 24 hours and then in 30% sucrose for 48 hours. Sections of 20 μm were evaluated for CD133 (1:100, Amersham) signal or green/red fluorescence. For quantification of staining/fluorescence, 3 sections per tumor were analyzed.

- *In situ hybridization*

In situ hybridization (ISH) for miR-31 was performed by Phylogeny Inc. (http://phylogenybioscience.com) as previously described (Bronisz et al., 2012; Godlewski et al., 2010b; Godlewski et al., 2008; Nuovo, 2008) in the formalin-fixed, paraffin-embedded brain tissue by using the MiRCURY LNA microRNA ISH Optimization Kit and locked nucleic acid (LNA)-modified, 5′ end digoxygenin (DIG)-conjugated probes (Exiqon, Vedbaek, Denmark). Briefly, paraffin sections (6-μm

thick) were de-paraffinized and treated with 15 µg/mL proteinase-K (Ventana) at 37°C for 4 minutes. After dehydration, slides were incubated with 1x and 5x probes cocktail (25pmol/100 ml or 125pmol per 100ml) of microRNA negative control LNA 5' DIG labeled probe (5'-DIG-GTGTAACACGTCTATACGCCCA-3') or miR-31 LNA 5'DIG labeled probe (5'-DIG-AGCTATGCCAGCATCTTGCCT -3'-DIG) at 50°C for 8 hours. This was followed by stringency washes with 5× standard saline citrate, 1× saline sodium citrate, and 0.2× saline sodium citrate buffers at 50°C; DIG blocking reagent (Roche, Mannheim, Germany) in maleic acid buffer containing 2% sheep serum at room temperature for 15 minutes; and alkaline phosphatase–conjugated anti-DIG (diluted 1:500 in blocking reagent; Roche) at room temperature for 2 hours. Enzymatic development was performed by incubating the slides with 4-nitro-blue tetrazolium and 5-brom-4-chloro-3'-indolylphosphate substrate (Roche) at 30°C for 2 hours to allow formation of dark-blue 4-nitro-blue tetrazolium formazan precipitate, followed by nuclear fast red counterstain (Vector Laboratories, Burlingame, CA) at room temperature for 10 minutes. Slides were then dismantled in water, dehydrated in alcohol solutions, and mounted with mounting medium (Vector Laboratories). Scrambled probe and U6 small nuclear RNA–specific probe were used as system control.

- *In vivo* studies

For all *in vivo* studies we used 6 experimental groups. Female athymic nude mice were purchased from Envigo. For all studies mice were housed at Harvard Medical School (HMS) animal facility in accordance with all NIH regulations. For intracranial tumor injection, cells were analyzed for viability using the Muse Count & Viability Reagent on the Muse Cell Analyzer (Millipore) following the manufactures' instructions to normalize number of viable cell prior to the transplantation of $1x10^3$ viable nodular GSCs (n=6) or $5x10^5$ viable invasive GSC (n=6), (stably transfected with PALM tomato and PALM GFP respectively) and stereotactically injected (2mm right lateral, 0.5mm frontal to the bregma and 4mm deep) into the brain of 6-8 week old mouse. For invasion studies, $1x10^3$ nodular GSCs and $5x10^5$ invasive GSC were admixed (n=6) and injected simultaneously. Animals were scarified 10 days after injection.

- *In vitro* assays

For all *in vitro* assays we used between 3 and 10 experimental variants per group. For EV treatment we used 3 experimental variants per group with EVs collected from one donor cells added to recipient cells in pairs: EVs from G33, G34, G88 (nodular GSCs) were added in to G6, G44, G146 (invasive

GSCs) and vice versa. For control treatment EVs from each GSC were collected and added to the same cells.

For spheroid formation assay GSC neurospheres were dissociated to single cell suspension using Accutase (Life Technologies) according to manufacturer's protocol, and plated at 200 (nodular GSCs transduced with PALM tomato, or cellular GFP) or 1000 (invasive GSCs transduced with PALM GFP, or cellular RFP) cells/well for mono-culture sphere assay or admixed (at 1:4 ratio or 1:1 ratio) for co-culture sphere (from different subclass and same subclass respectively) assay and cultured for 48h in neurosphere medium. Using this ratio the sphere is composed of roughly the same number of cells from each subclass after 48 h (Ricklefs et al., 2016).

For 3D spheroid migration assay (Del Duca et al., 2004) GSC neurospheres were transferred into 300 µl neutralized collagen (Advanced Biomatrix) I solution in an 8-well chamber slide (Nalgene, Rochester, NY, USA). Collagen was neutralized to pH 7.5 using 1N NaOH and supplemented with plain neurobasal. After polymerization at 37°C, the collagen was overlaid with 300 µl of neurobasal plain medium with or EVs from donor cells (at concentration 250 EVs per cell) and analyzed during 6 h in time-lapse.

For EV treatment recipient mono-cultured GSC neurospheres were maintained overnight in un-supplemented medium followed by 24h treatment with EVs derived from donor cells at concentration 250 EV particles per recipient cell.

For co-culture GSC neurosphere EV microRNA transfer assay PALM tomato labeled nodular GSCs were admixed with regular cytosolic GFP-labeled invasive GSCs (not PALM GFP to avoid the transfer of GFP-labeled EVs from invasive GSCs into nodular GSCs). As a control invasive GSCs were cultured in mono-culture neurosphere. After 48 spheroids were dissociated and sorted using flow cytometry. Cell were sorted based on: tomato (donor cells not included in analysis), GFP, and double positive GFP/tomato cells. RNA from sorted cells was used for qPCR analysis of microRNAs.

To validate the effect of EV miR-31 on mRNA targets expressed in invasive GSCs, EVs were collected from 3 nodular GSC (donor cells) and added to invasive GSCs (recipient cells) in triplicates in concentration 250 EV per cell. After 24h recipient cells were washed with PBS, spun down (200xg) and then processed for RNA isolation.

To test whether the same targets were also deregulated in nodular GSCs, we performed oligo transfection in triplicates using miR-31 mimic, miR-31 inhibitor or negative control oligo (Ambion)

using Lipofectamine 2000 (Invitrogen). 48 h after transfection cells were washed with PBS spun down (200xg) and then processed for RNA isolation.

- *RNA analysis*

RNA isolation

Total cellular RNA and EV RNA were extracted using Trizol (Invitrogen) and an SBI exoRNA isolation kit (System Bioscience) respectively, and treated with RNase-free DNase (Qiagen). RNA was quantified by Nanodrop and analyzed using a Bioanalyzer 2100 (Bronisz et al., 2014; Godlewski et al., 2008). The total RNA analysis was performed to verify level of 28S and 18S rRNA peaks in EV RNA and to estimate RNA concentration by comparing peak areas of a ladder with RNA fragments of known concentration and peak areas of the unknown samples. Small RNA analysis was performed to analyze microRNA content in EV microRNA. Gel-like images are provided for visualization of fragment sizing and distribution, as well as for a visual representation of the RNA ladder.

Absolute quantification of hsa-microRNA-31

Synthetic microRNA spike-in control was serially diluted 5 times for concentrations in 128 pmol to 0.125 pmol range on Nanostring platform. To calculate copy number of miR-31 algorithm was used where 152.23 pm in 1ug RNA equals 9.17×10^{13} of 20nt molecule. Each sample (cellular RNA and EV RNA) was quantified with respect to their own universal reference. The analysis of Nanostring data is described in following paragraph.

Quantitative RT-PCR

Reverse transcription was performed using miScript Reverse Transcription Kit (Bio-Rad). miScript SYBR Green PCR Kit (Qiagen) and miScript Primer Assays for six different microRNAs were used to amplify mature forms of miR-31, miR-21, miR-148 and miR-451 on Thermal Cycler (Bio-Rad). U6 small nuclear 2 (U6b) was used as the internal control. For quantitation of mRNA targets levels, reverse transcription was performed using the SuperscriptTM II Reverse Transcriptase Kit (Invitrogen). Then, the cDNAs were amplified using the QuantiFast SYBR Green PCR Kit (Qiagen). 18S rRNA was used as the internal control. MicroRNA and mRNA expression levels were normalized to their corresponding internal control genes, and relative change was calculated using ΔΔCt method or presented as raw Ct value.

Primers (Invitrogen) sequence:

ACVR2A	For	5′-GCCAGCATCCATCTCTTGAAGAC-3′
	Rev	5′-GATAACCTGGCTTCTGCGTCGT-3′
NF2	For	5′-TGAACGCACGAGGGATGAGTTG-3′
	Rev	5′-GCCTTTTCAGCCAACAGGTCAG-3′
MMP16	For	5′-GATTCAGCCATTTGGTGGGAGG-3′
	Rev	5′-CCCTTTCCAGACTGTGATTGGC-3′
DCX	For	5′-TATGCGCCGAAGCAAGTCTCCA-3′
	Rev	5′-CATCCAAGGACAGAGGCAGGTA-3′
PDPN	For	5′-GTGACCCTGGTTGGAATCATAG-3′
	Rev	5′-TCAGCTCTTTAGGGCGAGTA-3′
CTNND2	For	5′-TTTCCCTCGGTCCAGTCTAA-3′
	Rev	5′-GATGCCTCCTTGTCTCCTTATC-3′
CAMK2D	For	5′-ACACGGTGACTCCTGAAGCCAA-3′
	Rev	5′-GTCTCCTGTCTGTGCATCATGG-3′
FZD3	For	5′-GGCTCTCATAGTTGGCATTCCC-3′
	Rev	5′-TGGAGTACCTGTCGGCTCTCAT-3′
QKI	For	5′-GCAGAGTACGGAAAGACATGT-3′
	Rev	5′-ATAGGTCCCACAGCATCAGG-3′
SPARC	For	5′TGGACTCTGAGCTGACCGAATT-3′
	Rev	5′-ATGGATCTTCTTCACCCGCAG-3′

- *MicroRNA and mRNA expression analysis*

Nanostring

Nanostring microRNA technology was used to search for unique microRNA signatures in GSC and GSC-derived EVs. Sample preparation was performed according to the manufacturer's instructions. For data analysis positive and negative corrections, as well as a sample content normalization to the raw data was applied as per manufacturer guidelines. The first stage of data analysis was a correction to positive controls. Each experiment contained synthetic spike-in controls in the early stage preparatory mix that allow for the correction of sample-to-sample variation due to assay-specific factors such as differences in amount of input material or reagents. The positive correction was calculated by:

c×(ms); where c is count data for a gene in a given sample, m is the mean of the sum of the positive controls across all samples, and s is the sum of all of the positive controls for that given sample. The positive correction was applied to the data with t-test or one-way ANOVA.

Next, the negative correction subtracted background noise from the positively-corrected data using counts of sequence tags known to be absent from the assay, using the maximum of the negative controls which is a one-tailed Student's t-test using the negative controls as one group against all samples of a given gene as the other group, and using this to determine the significance of a given gene with a p-value cutoff by: *{c−m0forp-value<pand(c−m)≥0forp-value<pand(c−m)<0orforp-value≥p*; where c is all count data for a given gene, m is the mean of the count data for that gene across all samples, and p is the p-value cutoff used. If the p-value was at or above the cutoff (default 0.05), the counts for such gene were regarded as not significant and are set to 0 for all samples. If the p-value is below the cutoff, the mean for the count data for such gene were subtracted from the counts for the gene and count values that fell below zero were set to zero. Next, geometric mean of the top 100 microRNAs in all samples was used to effectively normalize data relative to total microRNA present.

Microarray

Whole Human Genome Oligo Microarray was performed by Arraystar (Rockville, MD). Briefly, total RNA from each sample was amplified and transcribed into fluorescent cDNA using the manufacturer's Agilent's Quick Amp Labeling protocol, version 5.7 (Agilent Technologies, Santa Clara, CA). The labeled cRNAs were hybridized to the Whole Human Genome Oligo Microarray (4 × 44K; Agilent Technologies). After having washed the slides, the arrays were scanned by the Agilent Scanner G2505B (Agilent Technologies). Agilent Feature Extraction software version 11.0.1.1 (Agilent Technologies) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX software version 11.5.1 (Agilent Technologies). Raw data of detected genes were normalized for further analysis. Differentially expressed genes were identified through fold change filtering. The data were deposited into Gene Expression Omnibus (http://www-ncbi-nlm-nih-gov.ezp-prod1.hul.harvard.edu/geo; accession number GSE89501).

- QUANTIFICATION, BIOINFORMATIC AND STATISTICAL ANALYSIS

For all experiments (unless specified otherwise) an unpaired, two-tailed t-test was used to compare 2 groups. One-way ANOVA, followed by Bonferroni's test, was conducted to test for significance among multiple groups. $P < 0.05$ was considered significant. Data are expressed as mean ± SD. Statistical analyses were performed using GraphPad Prism software. All significantly deregulated microRNAs were visualized as a heatmap and analyzed by principal component analysis using dChip software with the Statistical R package.

- *Functional bioinformatics analyses.*

Functional bioinformatics analyses were performed using QIAGEN's Ingenuity® Pathway Analysis (IPA®) using casual network (Kramer et al., 2014). Analysis was performed using gene expression microarray and Nanostring data from 6 nodular and 6 invasive GSCs. Lists of mRNAs and microRNAs differentially expressed between GSCs were analyzed based on the IPA library of canonical pathways (content date 2012-05-08). Such approach do not take individual expression levels into account but instead assume that transcriptionally altered genes have been determined using a suitable cutoff applied to the measured expression change which allowed to curate list where each gene in the dataset, can be either up- or down-regulated. The significance of the association between each list and a canonical pathway was measured by Fisher's exact test. As a result, a *P*-value was obtained, determining the probability that the association between the genes in experimental data set and a canonical pathway can be explained by chance alone. To identify biological functions that were most significant to experimental data sets, the functional analysis was performed using canonical pathway analysis between GSC subclasses. Right-tailed Fisher's exact test was used to calculate a significant *P*-value for each functional category as referenced in Ingenuity® Knowledge Base. The obtained *P*-value was further adjusted using the Benjamini–Hochberg correction. We focused the analysis of IPA categories with an adjusted $P \leq 0.01$ directly related to cellular functions, providing dynamical information about enriched biological functions. The goal of Downstream Effects Analysis was to identify those biological processes and functions that were likely to be caused by up- and down-regulated genes and predicted whether those processes are increased or decreased (these data were incorporated into Supplemental Tables S1 and S2). For microRNA-targets analysis we used IPA's microRNA Target Filter. It contains content from TarBase, TargetScan, miRecords, and micro-RNA related findings from the published literature. It also includes filtering tools that sort microRNA targets

and query them with microarray data to examine microRNA-mRNA pairings (these data were incorporated into Supplemental Tables S3-S6).

- Analysis of protein coding genes profiles of GSC subclass

The collection of the data from TCGA (Cancer Genome Atlas Research, 2008) was compliant with all applicable laws, regulations, and policies for the protection of human subjects, and necessary ethical approvals were obtained. Experimental and clinical data were analyzed using the GBM-BioDP (URL: http://gbm-biodp.nci.nih.gov) (Celiku et al., 2014). Gene expression data included data from three platforms HT_HG-U133A (488 patient samples×12042 features), HuEx-1_0-st-v2 (437 patient samples×18631 features), AgilentG4502A_07_1/2 (101+396 patient sample×17813 features). The data from these three platforms were aggregated (Verhaak et al., 2010). GSCs microarray data were queried for cluster analysis with clinical data of 564 patients. The experimental data were already pre-processed as a part of the TCGA data.

- Classification of GSC in to molecular subtype

The GSC genes expression profiles were used to predict the molecular subtype in supervised models using predictive analysis (PAM) (Tibshirani et al., 2002). This strategy enables analysis of uploaded expression data, and their classification into GBM subtypes. Using the classification of 201 samples, and the corresponding gene expressions for genes common in Verhaak 840 panel we entered GSC data to predict GSC class identity to display the correlation between GSC gene expression and GBM subclass.

- Displaying heatmap clustering mRNA and microRNA expression data for selected genes

The samples (columns of the heatmap) were annotated in three ways: first, according to cluster membership (the optimal number of clusters was determined using NbClust); second, according to TCGA (Verhaak et al., 2010) subtype; and third, by inspecting the status of a prognostic index (which was computed by weight averaging the gene expressions with the regression coefficients of a multi-gene Cox proportional hazards model). The gene names were annotated with their respective Hazard Ratios in a multi-gene Cox proportional hazards model.

- Performing multi-microRNA survival analysis

We performed two types of survival analyses: first, we stratified the samples according to the heatmap cluster membership, where the optimal number of clusters is picked out algorithmically. We used a Kaplan-Meier model to analyze the differences in survival. We also analyzed the effect of the

gene expressions on survival by constructing a multi-gene prognostic index. A Cox proportional hazards model was constructed with each gene as a covariate. A prognostic index was computed for each sample by weight averaging of each gene's expression by its regression coefficient in the Cox model. Samples were then stratified by the prognostic index (the default option is stratification down the median). The prognostic index status is depicted in the third annotation bar on the heatmap. We constructed a model for the full cohort of samples, as well as for samples stratified according to their GBM subtype. In some situations the Cox-model could not be computed as the model is non-convergent; in those instances plots were not informative and p-value was not calculated.

- Displaying heatmap clustering of gene and microRNA expression data correlation

To access the correlation between expression of microRNA from selected classes (C1 and C2) and genes expressed in tumor anatomic niches (with perinecrotic/core zone vs. infiltration/invasive zone) we first analyzed the expression correlation between genes from proliferative and invasive modes(Figure 1C) and microRNA in GBM using GBM-BioDP searches. Next we associated microRNA with C1-C2 classes that resulted in multiple hits of genes and microRNAs, which were displayed as a heatmap of the correlation between the gene and microRNA expressions. The associations between C1-C2 classes and infiltrative or perinecrotic zones were shown as clusters. Each cell of the heatmap represents how the expression of the gene (row) and the microRNA (column) are correlated, and is annotated with the correlation value.

- The analysis of correlation between gene expression and anatomic region of glioblastoma

Gene expression in the various anatomical regions of glioblastoma tumors was analyzed using the Ivy Glioblastoma Atlas Project (http://glioblastoma.alleninstitute.org/). Gene expression in five major anatomic structures of glioblastoma: Leading Edge (LE), Infiltrating Tumor (IT), Cellular Tumor (CT), Microvascular Proliferation (MVP), and Pseudopalisading Cells around Necrosis (PAN) were quantified by RNA sequencing in the Anatomic Structures RNA-Seq Study. The Ivy GAP cohort is comprised of 42 tumors. Samples from the anatomic structures were collected by laser microdissection and validated by ISH. Curated list of genes expressed in GSC based on IPA analysis (Supplemental Table 1-6) was queried with IVY GAP dataset and expression was visualized as z-score heatmap.

- Analysis of miR-31 targets correlation in GSC subclass

The TargetScan was used to select transcripts with 3′ UTR elements harboring matches to position 2–8 of the seed region of the miR-31 (n=3066). The list of transcript was queried with the normalized

mRNA array data in 3 nodular and 3 invasive GSC, followed by the analysis based on value of expression (with fold difference between subclass >2 and p <0.05) and expression correlation analysis with a nonparametric Kolmogorov-Smirnov (KS) statistic test.

- Analysis of microRNA profiles in GSC subclasses

For unsupervised analysis, Level 3 microRNA expression data (unc.edu_GBM.H-miRNA_8x15K.Level_3.1.8.0) from 479 Glioblastomas were obtained from The Cancer Genome Atlas data portal (Brennan et al., 2013). Processed data were filtered in order to reduce noise and prioritize microRNA with expression values showing maximal variation across analyzed samples. Variation filter based on "median absolute deviation (MAD)" was applied. MicroRNA with a MAD score smaller than 0.25 across all samples were removed (Supplemental Fig. S2C left), rendering a total of 221 microRNAs. Then an unsupervised expression analysis was performed using NMF algorithm (Brunet et al., 2004) (Supplemental Fig. S2D) and projected with an independent Principal Component Analysis (PCA) (Supplemental Fig. S2D, right). In NMF model and class (K) selection, the cophenetic coefficient corresponding to hierarchical clustering of consensus matrices for K=2-10 indicated that optimum number of classes recognized was a robust two-class partition K=2 (Supplemental Fig. S2C, right). A Fisher exact test was used to identify associations between unsupervised microRNA classes and previously determined GBMs subtypes considering $p \leq 0.05$ as statistically significant.

- DATA AND SOFTWARE AVAILABILITY

- The datasets were deposited into Gene Expression Omnibus (http://www-ncbi-nlm-nih-gov.ezp-prod1.hul.harvard.edu/geo; accession number GSE89501).
- Softwares used in the study:

  Ivy GAP http://glioblastoma.alleninstitute.org/static/home

  GBM biodiscovery portal: https://gbm-biodp.nci.nih.gov

  dchip with R package: http://www.dchip.org/

  IPA: https://apps.ingenuity.com/

SUPPLEMENTAL REFERENCES

Brennan, C. W., Verhaak, R. G., 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.

Bronisz, A., Godlewski, J., Wallace, J. A., Merchant, A. S., Nowicki, M. O., Mathsyaraja, H., Srinivasan, R., Trimboli, A. J., Martin, C. K., Li, F.*, et al.* (2012). Reprogramming of the tumour microenvironment by stromal PTEN-regulated miR-320. Nature cell biology *14*, 159-167.

Bronisz, A., Wang, Y., Nowicki, M. O., Peruzzi, P., Ansari, K. I., Ogawa, D., Balaj, L., De Rienzo, G., Mineo, M., Nakano, I.*, et al.* (2014). Extracellular vesicles modulate the glioblastoma microenvironment via a tumor suppression signaling network directed by miR-1. Cancer research *74*, 738-750.

Brunet, J. P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. Proceedings of the National Academy of Sciences of the United States of America *101*, 4164-4169.

Celiku, O., Johnson, S., Zhao, S., Camphausen, K., and Shankavaram, U. (2014). Visualizing molecular profiles of glioblastoma with GBM-BioDP. PloS one *9*, e101239.

Del Duca, D., Werbowetski, T., and Del Maestro, R. F. (2004). Spheroid preparation from hanging drops: characterization of a model of brain tumor invasion. Journal of neuro-oncology *67*, 295-303.

Godlewski, J., Nowicki, M. O., Bronisz, A., Nuovo, G., Palatini, J., De Lay, M., Van Brocklyn, J., Ostrowski, M. C., Chiocca, E. A., and Lawler, S. E. (2010a). MicroRNA-451 Regulates LKB1/AMPK Signaling and Allows Adaptation to Metabolic Stress in Glioma Cells. Molecular Cell *37*, 620-632.

Godlewski, J., Nowicki, M. O., Bronisz, A., Nuovo, G., Palatini, J., De Lay, M., Van Brocklyn, J., Ostrowski, M. C., Chiocca, E. A., and Lawler, S. E. (2010b). MicroRNA-451 regulates LKB1/AMPK signaling and allows adaptation to metabolic stress in glioma cells. Molecular cell *37*, 620-632.

Godlewski, J., Nowicki, M. O., Bronisz, A., Williams, S., Otsuki, A., Nuovo, G., Raychaudhury, A., Newton, H. B., Chiocca, E. A., and Lawler, S. (2008). Targeting of the Bmi-1 oncogene/stem cell renewal factor by microRNA-128 inhibits glioma proliferation and self-renewal. Cancer research *68*, 9125-9130.

Kim, S. H., Ezhilarasan, R., Phillips, E., Gallego-Perez, D., Sparks, A., Taylor, D., Ladner, K., Furuta, T., Sabit, H., Chhipa, R.*, et al.* (2016). Serine/Threonine Kinase MLK4 Determines Mesenchymal Identity in Glioma Stem Cells in an NF-kappaB-dependent Manner. Cancer cell *29*, 201-213.

Kramer, A., Green, J., Pollard, J., Jr., and Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics *30*, 523-530.

Lai, C. P., Kim, E. Y., Badr, C. E., Weissleder, R., Mempel, T. R., Tannous, B. A., and Breakefield, X. O. (2015). Visualization and tracking of tumour extracellular vesicle delivery and RNA translation using multiplexed reporters. Nature communications *6*, 7029.

Mao, P., Joshi, K., Li, J., Kim, S. H., Li, P., Santana-Santos, L., Luthra, S., Chandran, U. R., Benos, P. V., Smith, L.*, et al.* (2013). Mesenchymal glioma stem cells are maintained by activated glycolytic metabolism involving aldehyde dehydrogenase 1A3. Proceedings of the National Academy of Sciences of the United States of America *110*, 8644-8649.

Miyazaki, T., Pan, Y., Joshi, K., Purohit, D., Hu, B., Demir, H., Mazumder, S., Okabe, S., Yamori, T., Viapiano, M.*, et al.* (2012). Telomestatin impairs glioma stem cell survival and growth through the disruption of telomeric G-quadruplex and inhibition of the proto-oncogene, c-Myb. Clinical cancer research : an official journal of the American Association for Cancer Research *18*, 1268-1280.

Nakano, I., Joshi, K., Visnyei, K., Hu, B., Watanabe, M., Lam, D., Wexler, E., Saigusa, K., Nakamura, Y., Laks, D. R.*, et al.* (2011). Siomycin A targets brain tumor stem cells partially through a MELK-mediated pathway. Neuro-oncology *13*, 622-634.

Nuovo, G. J. (2008). In situ detection of precursor and mature microRNAs in paraffin embedded, formalin fixed tissues and cell preparations. Methods *44*, 39-46.

Ricklefs, F., Mineo, M., Rooj, A. K., Nakano, I., Charest, A., Weissleder, R., Breakefield, X. O., Chiocca, E. A., Godlewski, J., and Bronisz, A. (2016). Extracellular Vesicles from High-Grade Glioma Exchange Diverse Pro-oncogenic Signals That Maintain Intratumoral Heterogeneity. Cancer research *76*, 2876-2881.

Shao, H., Chung, J., Balaj, L., Charest, A., Bigner, D. D., Carter, B. S., Hochberg, F. H., Breakefield, X. O., Weissleder, R., and Lee, H. (2012). Protein typing of circulating microvesicles allows real-time monitoring of glioblastoma therapy. Nature medicine *18*, 1835-1840.

Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proceedings of the National Academy of Sciences of the United States of America *99*, 6567-6572.

Verhaak, R. G., 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.