Spatiotemporal analysis of glioma heterogeneity reveals Col1A1 as an actionable target to disrupt tumor mesenchymal differentiation, invasion and malignancy

Andrea Comba[1,2,3], Syed M. Faisal[1,2,3], Patrick J. Dunn[1,2,3], Anna E. Argento[1,2], Todd C. Hollon[1], Wajd N. Al-Holou[1], Maria Luisa Varela[1,2,3], Daniel B. Zamler[1,2,3], Gunnar L Quass[7], Pierre F. Apostolides[7,8], Clifford Abel II[1,2,3],Christine E. Brown[9], Phillip E. Kish[1,5], Alon Kahana[5], Celina G. Kleer[3,6], Sebastien Motsch[4], Maria G Castro[1,2,3], Pedro R. Lowenstein[1,2,3,10], *

[1]Dept. of Neurosurgery, University of Michigan Medical School, Ann Arbor, 48109, MI, USA
[2]Dept. of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, 48109 MI, USA
[3]Rogel Cancer Center, University of Michigan Medical School, Ann Arbor, 48109, MI, USA
[4]School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, USA
[5]Ophthalmology & Visual Science, University of Michigan Medical School, Ann Arbor, 48109 MI, USA
[6]Dept. of Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA
[7]Kresge Hearing Research Institute, Department of Otolaryngology-Head & Neck Surgery, University of Michigan Medical School, Ann Arbor, MI 48109, USA
[8]Dept. of Molecular & Integrative Physiology, University of Michigan Medical School, Ann Arbor, MI 48109, USA.
[9]Departments of Hematology & Hematopoietic Cell Transplantation and Immuno-Oncology, City of Hope, Duarte, CA, USA.
[10] Dept. of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109,
*Correspondence to: pedrol@umich.edu

ABSTRACT

Intra-tumoral heterogeneity and diffuse infiltration are hallmarks of glioblastoma that challenge treatment efficacy. However, the mechanisms that set up both tumor heterogeneity and invasion remain poorly understood. Herein, we present a comprehensive spatiotemporal study that aligns distinctive intra-tumoral histopathological structures, oncostreams, with dynamic properties and a unique, actionable, spatial transcriptomic signature. Oncostreams are dynamic multicellular fascicles of spindle-like and aligned cells with mesenchymal properties, detected using *ex vivo* explants and *in vivo* intravital imaging. Their density correlates with tumor aggressiveness in genetically engineered mouse glioma models, and high-grade human gliomas. Oncostreams facilitate the intra-tumoral distribution of tumoral and non-tumoral cells, and the invasion of the normal brain. These fascicles are defined by a specific molecular signature that regulates their organization and function. Oncostreams structure and function depend on overexpression of COL1A1. COL1A1 is a central gene in the dynamic organization of glioma mesenchymal transformation, and a powerful regulator of glioma malignant behavior. Inhibition of COL1A1 eliminated oncostreams, reprogramed the malignant histopathological phenotype, reduced expression of the mesenchymal associated genes, induced changes in the tumor microenvironment and prolonged animal survival. Oncostreams represent a novel pathological marker of potential value for diagnosis, prognosis, and treatment.

INTRODUCTION

High grade gliomas (HGG) are the most prevalent and malignant brain tumors. They grow rapidly, invade surrounding normal brain, and recur within 12 months. Median survival is 18-20 months, in spite of current standard of care[1,2]. Despite some notable successful outcomes from the large cancer sequencing programs, which identified driver genes in a number of cancers, effective therapeutically actionable breakthroughs have not yet been identified in HGG[3-7].

HGG are highly heterogeneous at the histological, cellular, and molecular level. Heterogeneity of HGG is illustrated in addition by characteristic pathological structures such as pseudopalisades, microvascular proliferation, and areas of hypoxia and necrosis[2,8]. The molecular characterization of glioma heterogeneity identified three main molecular signatures: proneural, mesenchymal, and classical[4,9] . However, later studies demonstrated that all three transcriptomic signatures are expressed within individual tumors[5,10,11]. Rather than outright glioma subtypes, the consensus proposes that individual tumors are enriched in particular molecular subtypes. Thus, studies have correlated histological features with genetic alterations and transcriptional expression patterns. For example, highly aggressive histological features such as hypoxic, perinecrotic and microvascular proliferative zones have been associated with the mesenchymal molecular signature and worse prognosis[8]. However, the molecular classification has only minor clinical impact. Thus, alternative classification schemes using a pathway-based classification are currently being considered[12]. How these new classifications will deal with tumor heterogeneity remains to be explored. Moreover, different microenvironmental, metabolic, and therapeutic factors drive transitions of the GBM transcriptomic signature, particularly transitions to mesenchymal states. It is important to note that glioblastoma plasticity explains the high degree of tumor heterogeneity and prompts the selection of new clones at recurrence or therapy resistance[13-16]. It has been established that intra-tumoral heterogeneity is represented by four main cellular states, the progenitor, astrocyte, oligodendrocyte, and mesenchymal like-state, which represent tumor plasticity and are affected by the tumor microenvironment[15].

Tumoral mesenchymal transformation is a hallmark of gliomas[13,17,18]. A mesenchymal phenotype is defined by cells with spindle-like, fibroblast-like morphology associated with alterations in their dynamic cellular organization leading to an increase in cell migration and invasion[19,20]. The mesenchymal phenotype is controlled by particular transcription factors and downstream genes related to the extracellular matrix (ECM), cell adhesion, migration, and tumor angiogenesis[18,21,22].

However, the cellular and molecular mechanisms that regulate mesenchymal transformation in gliomas, especially concerning the mesenchymal features of invasive cells, has remained elusive. Integrating morphological features, spatially resolved transcriptomics, and cellular dynamics resulting from mesenchymal transformation, growth, and invasion are thus of great relevance to our understanding of glioma progression[13,20].

Cell migration is essential to continued cancer growth and invasion. Morphological and biochemical changes that occur during mesenchymal transformation allow glioma cells to move throughout the tumor microenvironment and invade the adjacent normal brain. Tumor cells also migrate along blood vessels, white matter tracks, and the subpial surface. Within the tumor microenvironment, aligned extracellular matrix fibers help guide the movement of highly motile mesenchymal-like cancer cells [23-25].

Our study reveals that malignant gliomas, both high grade human gliomas and mouse glioma models, display regular distinctive anatomical multicellular fascicles of aligned and elongated, spindle-like cell. We suggest they are areas of mesenchymal transformation. For the sake of simplicity in their description throughout the manuscript, we have named these areas 'oncostreams'.

Using time lapse laser scanning confocal imaging *ex vivo*, and multiphoton microscopy *in vivo* of high grade glioma explants we demonstrated that oncostreams are organized collective dynamic structures; they are present at the tumor core and at areas of tumor invasion of the normal brain. Collective motion is a form of collective behavior where individual units' (cells) movement is regulated by local intercellular interactions (i.e., attraction/repulsion) resulting in large scale coordinated cellular migration [26-30]. Collective motion plays an essential role in embryogenesis and wound healing[27,31-33]. Emergent organized collective motion patterns could help explain so far poorly understood tumoral behaviors such as invasion, metastasis, and especially recurrence[27,31].

Studies of tumor motility have concentrated on the behavior of glioma cells at the tumor invasive border[27,34,35]. Potential motility at the glioma core has not been studied in much detail so far. This study challenges the conventional belief that cells in the central core are non-motile and indicate that the glioma core displays collective migratory patterns. This would suggest that the capacity of gliomas to invade and grow, results from phenomena occurring at the tumor invasive border, and from the overall capacity of gliomas to organize collective motion throughout the tumor mass, from the tumor core to the tumor invasive border.

To study the molecular mechanisms underlying oncostream organization and function we used laser capture microdissection (LCM) followed by RNA-sequencing and bioinformatics analysis. We discovered that oncostreams are defined by a mesenchymal transformation signature enriched in extracellular matrix related proteins, and which suggest that Collagen1A1 (COL1A1) is a key determinant of oncostream organization. Inhibition of COL1A1 within glioma cells lead to oncostream loss and reshaping of the highly aggressive phenotype of HGG. These data indicate that COL1A1 is likely to constitute the tumor microenvironment scaffold, and to serve to organize areas of collective motion in gliomas.

COL1A1 has been shown previously to be a major component of the extracellular matrix in different cancers, including glioma, and has been reported to promote tumor growth and invasion[36,37]. Alternatively, some data suggest that collagen fibers could be passive barriers to resist tumor cell infiltration or provide biophysical and biochemical support for cell migration. Some studies reported that density of COL1A1 inversely correlates with glioma patient's prognosis. However, other studies showed that either increased or decreased deposition of collagen could be associated with increased tumor malignancy[38-40]. Therefore, it is important to further determine the role of COL1A1 in glioma invasion and continued growth.

This study provides a comprehensive study of the histological, morphological, and dynamic properties of glioma tumors. In addition, we uncover a novel characterization of the molecular mechanisms that define intra-tumoral mesenchymal transformation in gliomas and discuss their therapeutic implications. Oncostreams are anatomically and molecularly distinctive, regulate glioma growth and invasion, display collective motion, and are regulated by the extracellular matrix, specially by COL1A1. Inhibiting COL1A1 within glioma cells is a potential therapeutic strategy to mitigate glioma mesenchymal transformation, intra-tumoral heterogeneity, and thus, reduce deadly glioma invasion and continued growth.

RESULTS

Intra-tumoral multicellular fascicles of elongated and aligned cells in gliomas: oncostreams

High grade gliomas (HGG) are characterized by anatomical, cellular and molecular heterogeneity which determines, in part, tumor aggressiveness and reduces treatment efficacy[5,7,11]. Histopathological analysis of mouse and human gliomas revealed the presence of frequent distinct multicellular fascicles of elongated (spindle-like) and aligned cells (≈5-30 cells wide) distributed throughout the tumors. These structures resemble areas of mesenchymal transformation which we describe as "oncostreams" (**Fig. 1A-B**).

To study the presence and morphological characterization of oncostreams, we examined histological sections from various mouse glioma models as well as human glioma specimens **(Fig. 1 A-B).** We determined the existence of oncostreams in genetically engineered mouse models (GEMM) of glioma including NPA (Nras, shP53, shATRx) and NPD (Nras, shP53, PDGFβ) and other implantable models (GL26) (**Fig. 1A and Supplementary Fig. 1A-B**). Moreover, human glioma samples from primary resections and a xenograft glioma model, SJGBM2, established the presence of these multicellular structures in human tissue (**Fig. 1B and Supplementary Fig. 1C**). Morphological analysis determined that cells within histological areas corresponding to oncostreams have an aspect ratio of 2.63±0.19 (elongated or spindle-like cells) compared to the surrounding tissue where cells have an aspect ratio of 1.37±0.12 (round cells), both in mouse and human gliomas as shown in **Fig. 1C and Supplementary Fig. 1D.** We also determined that elongated cells within oncostreams are nematically aligned with each other, whereas outside of oncostreams, cell orientations are not aligned (**Fig. 1D and Supplementary Fig. 1E).**

To gain insight into the cellular features of oncostreams we asked if they are homogeneous or heterogeneous multicellular structures. We observed that in GEMM of gliomas, oncostreams are formed by GFP+ tumor cells, and are enriched in other tumor microenvironment cells such as ACTA2+ mesenchymal cells, Iba1+ and CD68+ tumor associated microglia/macrophages cells, Nestin+ cells and GFAP+ glial derived cells (**Fig. 1, E-G, and Supplementary Fig. 2 A-D**). The quantification of mesenchymal cells (ACTA2+), and tumor associated microglia/macrophages (TAM) cells (CD68+ and Iba1+) showed a significant enrichment of these populations within oncostreams compared to the surrounding areas (**Fig. 1E-G**). Moreover, non-tumoral cells within oncostreams were positively aligned along the main axes of oncostreams, and with tumor cells in mouse gliomas

(**Fig. 1H**). This suggests that oncostreams are mesenchymal-like structures which interact with TAM and mesenchymal cells.

To test if oncostreams form along existing brain structures, we evaluated their co-localization with white matter tracts. Although, occasional positive immune-reactivity (Neurofilament-L) was present within some areas of the tumors, oncostream fascicles were not preferentially organized along brain axonal pathways **(Supplementary Fig. 1F)**. These data indicate that oncostreams are fascicles of spindle-like aligned cells within glioma tumors, which contain tumor and non-tumor cells.

Oncostream density positively correlates with tumor aggressiveness and poor prognosis in mouse and human gliomas

Oncostreams are unique histological features that contribute to intra-tumoral heterogeneity suggesting a potential role in glioma progression and malignancy. To understand whether the presence of oncostreams correlates with tumor aggressiveness and clinical outcomes, we generated genetically engineered tumors of different malignant behaviors using the Sleeping Beauty Transposon system. These models reproduce the malignant histopathological features of gliomas as demonstrated in previous studies[41-44]. We induced tumors harboring two different genotypes: (1) Activation of RTK/RAS/PI3K pathway, in combination with p53 and ATRX downregulation (**NPA**), and, (2) RTK/RAS/PI3K activation, p53 downregulation, ATRX downregulation, and mutant IDH1-R132 expression (**NPAI**) **(Fig. 2A)**. IDH1-wild-type tumors (**NPA**) display a highly malignant phenotype and worse survival prognosis (Mediam survival (MS): 70 days), compared with tumors harboring the IDH1-R132R mutation, **NPAI,** (MS: 213 days) **(Fig. 2B)**. This outcome reproduces human disease, as patients with IDH1-mutant tumors also have prolonged median survival[1,44,45]. Tumor histopathological analysis showed a positive correlation between the density of oncostreams and tumor malignancy **(Fig. 2C- D)**. NPA (IDH1-WT) tumors exhibited larger areas of oncostreams within a highly infiltrative and heterogeneous glioma characterized by abundant necrosis, microvascular proliferation, pseudopalisades and cellular heterogeneity as described before [43,44]. Conversely, NPAI (IDH1-Mut) tumors display a very low density of oncostreams and a homogenous histology mainly comprised of round cells, low amounts of necrosis, no microvascular proliferation, absence of pseudopalisades and less invasive borders **(Fig. 2C and Supplementary Fig. 5).**

Further, to objectively identify and quantify tumor areas covered by oncostreams, we trained a fully convolutional neural network (fCNN) (**Supplementary Fig. 3 and 4A**). Our deep learning analysis

found that oncostreams occupied 15.28 ± 6.10% of the area in NPA tumors compared with only 1.18 ± 0.81 % in NPAI tumors (**Fig. 2C and D, and Fig. Supplementary 5A and B**). Cellular alignment analysis validated the presence or absence of oncostreams (**Fig. 2E**).

To determine whether oncostreams are linked to glioma aggressiveness in human patients, we evaluated a large cohort of TCGA glioma diagnostic tissue slides from the Genomic Data Commons Portal of the National Cancer Institute. We visually examined 100 TCGA-glioblastoma multiforme tissue sections (WHO Grade IV) and 120 TCGA-low grade glioma tissues (WHO Grade II and III) using the portal's slide image viewer (**Supplementary Table 1**). Oncostreams were present in 47% of TCGA-GBM grade IV tumors tissue, in 8.6 % of TCGA-LGG grade III, and were absent from TCGA-LGG grade II (**Fig. 3A-C and Supplementary Table 2**), consistent with tumor aggressiveness (http://gliovis.bioinfo.cnio.es)[46]. We then determined the presence of oncostreams across known molecular subtypes of HHG (Grade IV)[4]. We found oncostream fascicles in 59.4% of Mesenchymal (MES), 53.6% of Classical (CL) subtypes and only 26.7% of Proneural (PN) (**Fig. Supplementary 6A**). Finally, we evaluated oncostreams presence related to IDH status and 1p 19q co-deletion in LGG (Grade III). Oncostreams were present in 16.6% of IDH-WT subtype, 5% of IDHmut-non-codel and absent from IDHmut-codel subtype (**Fig. Supplementary 6B)**. These analyses suggest that oncostream presence is higher in Mesenchymal and Classical subtypes and correlates with IDH-WT status, and thus with a poor prognosis.

To validate the histological identification, we examined H&E images using our deep learning algorithm (**Fig. Supplementary 4B**). We observed a strong concordance (>84%) between machine learning and the manual histological identification of oncostreams (**Table Supplementary 3)**. Oncostream presence and their segmentation by deep learning is illustrated in **Fig. 3C and Supplementary Fig. 7 and 8**. Additionally, alignment analysis of glioma cells confirmed the existence of fascicles of elongated, mesenchymal-like cells in human gliomas (**Fig. 3D**). Thus, our deep learning algorithm validates our histological identification of oncostreams and confirms that the density of oncostream fascicles positively correlates with glioma aggressiveness.

The analysis of cellular heterogeneity showed that non-tumoral cells such as Iba1+ macrophages/microglia and GFAP+ glial derived cells were positively aligned within oncostreams tumoral cells (SOX+) in human HGG (**Fig. 3E**). Conversely, we detected that low grade gliomas (LGG) exhibited homogenous round cells, GFAP+ and Iba1+ cells throughout the tumor with no defined orientation or alignment (**Fig. 3F**).

Oncostreams are defined by a distinctive spatial transcriptome signature

To determine whether oncostreams fascicles are characterized by a specific gene expression profile, we performed a spatially-resolved transcriptomic analysis using laser capture microdissection (LCM) coupled to RNA sequencing (RNA-Seq). Oncostreams were dissected according to their morphological characteristics defined above. Surrounding areas of homogenous rounded cells were selected as non-oncostreams areas (control) (**Fig. 4A**). RNA-Seq analysis detected a set of 43 differentially expressed (DE) genes; 16 genes were upregulated and 27 downregulated within oncostreams (**Fig. 4 B-C and Table Supplementary 4**).

Functional enrichment analysis of DE genes, performed using the I-PathwayGuide platform (Advaita Corporation, MI, USA), showed that False Discovery Rate (FDR) corrected gene ontology (GOs) terms were associated with migration and extracellular matrix biological process. GOs such as "positive regulation of motility" "positive regulation of cell migration", "collagen catabolic processes" and "extracellular matrix organization" were the most over-represented biological processes (**Fig. 4D and Table Supplementary 5**). The upregulated DE genes within the relevant GOs include: COL1A1, MMP9, MMP10, ACTA2, ADAMTS2, CDH5, CYR61, PLP1 and those downregulated were ENPP2, AKAP12, BDKRB1 (**Fig. 4E and Fig. Supplementary 9**). Significant DE genes shared by related GOs are shown in **Supplementary Fig. 9**. These data indicate that oncostreams can be identified by a specific gene expression set and suggest a distinct role for oncostreams as intra-tumoral mesenchymal-like migratory assemblies within glioma tumors.

COL1A1 contributes to oncostream organization in high-grade gliomas

Histopathologically, oncostreams are spindle-like multicellular fascicles with a defined DE gene expression signature enriched in mesenchymal genes. The GO ontology analyses suggests a central role of collagen catabolic process and extracellular matrix organization in oncostreams function. To understand the molecular mechanisms that regulate oncostream organization and function, we identified critical genes using network analysis. Network interactions revealed that COL1A1 is a hub gene, one of the most highly connected nodes, representing a potential regulator of the network's signaling pathways and biological functions (**Fig 5A and Fig. Supplementary 10A**). We found that the most relevant COL1A1 related pathways include: Focal Adhesion, Extracellular Matrix

Organization and Integrin Signaling pathways (**Fig. Supplementary 10B-C and Table Supplementary 7 and 8**).

To analyze the role of COL1A1 in oncostream organization, we analyzed COL1A1 expression by immunofluorescence analysis. The COL1A1 gene encodes for the alpha-1 chain of type I collagen fibers. We observed that collagen fibers were aligned within oncostreams and overexpressed in more aggressive NPA (IDH1-WT) gliomas compared with NPAI (IDH1-Mut) tumors. COL1A1 expression was significantly lower and only found surrounding blood vessels in NPAI (IDH-Mut) tumors (**Fig. 5B-C**). Correspondingly, human GBM glioma tumors (IDH1-WT) with high oncostream densities showed prominent alignment of collagen fibers along these fascicles and higher COL1A1 expression compared to LGG (IDH1-Mut) (**Fig. 5D and E**).

Moreover, TCGA-glioma data indicate that COL1A1 has differentially higher expression in GBM histological Grade IV. LGG IDH-WT tumors display higher expression of COL1A1 than IDH1-Mutant. Within the GBM molecular subtype classification[4,9], the Mesenchymal group shows higher expression of COL1A1 than the Proneural and Classical groups (**Fig. Supplementary 11A**); the COL1A1 gene is clearly associated with the mesenchymal subtype. Analysis of patient survival related to COL1A1 expression showed that mesenchymal GBM subtype displayed a significantly shorter survival (MS: 10.4 months) for COL1A1 high tumors, compared to COL1A1 low (MS: 17.9 months) tumors. Classical and Proneural subtypes did not show survival differences associated to COL1A1 expression (**Fig. Supplementary 11B**). Thus, oncostreams represent intra-tumoral mesenchymal-like structures organized along collagen fibers.

COL1A1 depletion leads to oncostream loss, tumor microenvironment (TME) remodeling and increases in median survival

To evaluate the functional role of COL1A1 in oncostream formation we generated a COL1A1-deficient genetically engineered mouse glioma model. We generated COL1A1 wildtype, and COL1A1 knock-down tumors with different genetic backgrounds (**Fig Supplementary 12A-C**). COL1A1 downregulation increased median survival (MS) (**Fig. 5F and G**). The knockdown of COL1A1 in NPA tumors (NPAshCOL1A1) increased survival to MS: 123 days, compared to NPA control tumors (MS: 68 days) (**Fig. 5F**). Similarly, COL1A1 knockdown in NPD tumors harboring PDGFβ ligand upregulation (NPDshCOL1A1), also exhibited an increased median survival (MS: 98 days) compared to the NPD controls (MS: 74 days) (**Fig. 5G**).

To further analyze the effects of COL1A1 downregulation, we evaluated the histopathological features of glioma tumors, quantified oncostream density using deep learning analysis and evaluated COL1A1 expression within glioma tissues (**Fig. 5 H-I**). We observed that NPA tumors with COL1A1 downregulation showed a significant reduction of COL1A1 immunoreactivity within tumors; it was only maintained in small areas surrounding blood vessels (**Fig. 5J-K**). COL1A1 inhibition led to oncostream loss and reprogramming of the histopathological tumoral characteristics as evidenced by homogenous round cell morphology, resembling low grade tumors (**Fig. 5J-K**). Downregulation of COL1A1 in NPD tumors appeared less effective, with large areas of remaining COL1A1 (**Fig. %H-I**). Nonetheless, COL1A1 was downregulated within tumor cells and oncostream dismantling was significant compared to NPD control. Some oncostream areas remained associated with blood vessels which displayed significant amounts of COL1A1 (**Fig. 5J-K**).

We analyzed the effect of COL1A1 depletion on the intrinsic properties of tumoral cells. In vitro studies showed that COL1A1-knockdown cells exhibited a significantly decreased cell proliferation and cell migration compared to controls (**Supplementary Fig. 13 A-D**). Also, we observed that intracranial implantation of COL1A1-knockdown cells resulted in decreased tumor growth and progression when compared to controls (**Fig. Supplementary 13E**). *In vivo*, genetically engineered COL1A1 knockdown tumors displayed decreased cell proliferation (PCNA+ cells) (**Fig. 6A-B),** increased apoptosis via activation of Cleaved-Caspase 3, and downregulation of the anti-apoptotic protein Survivin **(Supplementary Fig. 14A-C**).

Furthermore, to determine whether COL1A1 downregulation within glioma cells modifies the glioma TME we analyzed changes in tumor associated macrophages (TAM), endothelial cells and mesenchymal cells. We found that COL1A1 knockdown tumors exhibited a decreased recruitment of CD68+ TAM (**Fig. 6C-D**), impaired CD31+ endothelial vascular proliferation (**Fig. 6E-F**) and diminished ACTA2+ perivascular mesenchymal cells (**Fig. 6G-H**). Moreover, inhibition of COL1A1 within glioma cells led to downregulation of fibronectin expression, a mesenchymal associated extracellular matrix protein (**Fig. Supplementary 14E-F**) that is associated with a more aggressive phenotype.

These preclinical animal models knocked down the expression of COL1A1 from the earliest stages of tumor development. Further, to evaluate the effects of the pharmacological degradation of deposited collagen fibers in highly malignant tumors we analyzed explants of brain tumor sections treated with collagenase. We observed that collagenase treatment decreased reticular fibers

(general collagen staining), reduced COL1A1 expression and disassemble fibers' alignment along tumoral cells and caused oncostreams depletion in a dose dependent manner (**Fig. Supplementary 15A-D**). These data indicate that oncostream organization and functions are regulated by COL1A1. COL1A1 knockdown within glioma cells decreased oncostream formation, reprogramed glioma mesenchymal transformation and remodeled the glioma TME, thus increasing animal survival. COL1A1 inhibition represents a novel approach for future translational development.

Oncostreams' mesenchymal patterns reveal intra-tumoral collective motion in gliomas

GO analysis indicates that biological processes such as positive regulation of motility/migration are enriched within oncostream fascicles. Overexpression of extracellular matrix (ECM)-associated proteins suggest a potential role of COL1A1 fibers in regulating oncostreams' motility. To study if oncostreams represent migratory structures within glioma tumors, we established a physiologically viable explant brain tumor slice model containing a high density of oncostreams (**Fig. 7A**). The movement of glioma cells expressing green fluorescent protein (GFP), within the thickness of each explant, was visualized using time-lapse confocal imagining and tracked using Fiji's plug-in Track-Mate **(Fig. 7A-C).**

Migration analyses show complex glioma cell dynamics throughout the tumor core. The glioma tumor core displays groups of cells (within particular zones) with similar nematic orientation and displaying complex movement patterns (**Fig. 7D and Fig. Supplementary 16A**) and, which represent collective motion[27,29-31]. Angle velocity distribution indicated the existence of three patterns of collective motility shown schematically in **Fig. 7D and F**: in 'Zone A' cells don't have a preferred direction, in 'Zone B' cells move in opposite directions (~ 135° and 315°), and in 'Zone C' all cells move with a predominant preferred direction (~ 45°) (**Fig. 7D and F).** We named these patterns 's*warm*' (Zone A), '*stream*' (Zone B), or '*flock*' (Zone C) (**Fig. 7G).** They were classified by likelihood analysis: the distribution of the angle velocity is constant in a *swarm* (all angle velocity are equally probable), bi-modal in a *stream* (cells are moving in equal but opposite directions), and uni-modal in a *flock* (cells move in one direction) (**Fig. 7H)**. These patterns were observed in all tumor slices examined (**Fig. Supplementary 18, 19 and 20**). Average cell speeds differed among the three patterns (**Fig. 7E, and Supplementary 18, 19 and 20**). In the tumor core, swarms moved faster and without orientation, followed by directionally moving flocks and streams (**Fig. Supplementary 27**). To determine which of these collective motion patterns match oncostream histological features, we

analyzed H&E sections corresponding to imaged organotypic slices (**Fig. Supplementary 16B).** Cells within histological areas corresponding to *streams* and *flocks* have an aspect ratio of 2.2 and 2.7, respectively, (spindle-like cells), while those within areas corresponding to *swarms* have an aspect ratio of 1.2 (round cells) **(Fig. Supplementary 16C-D)**. Moreover, elongated cells within *streams* and *flocks* are nematically aligned with each other, whereas round cells within *swarms* are not **(Fig. Supplementary 16E).** As predicted by our *in silico* model, [47] these results suggest that cell shape, or eccentricity, is driving feature in the organization of collective motion patterns (**Fig. Supplementary 16F**). Therefore, taking into account cell shape and alignment, we define oncostreams as the histological expression of collective motion patterns (*streams* and *flocks*). Notice that only the dynamic analysis of collective motion can differentiate between *streams* and *flocks*. At the histological level both appear as oncostreams.

In collective motion of flocks, interactions among individual cells are sufficient to propagate order throughout a large population of starlings[48]. To define if oncostream migration patterns recall organized collective motion behavior, we analyzed the organization of the cells by performing local pair-wise correlation analysis (relative position and pair directional correlation) by tumor zones (**Fig. Supplementary 17A-C**). These analyses indicate the spatial correlation of location and alignment between individual cells. We observed that within *swarms* cells are more separated, as neighbors are located at 20-40 µm. S*treams* and *flocks* have higher cell density, and the nearest neighbors are closer, at 20-30 µm (**Fig. Supplementary 17E and S18, S19, S20**). Pair-wise directional correlation with nearby neighbors showed that cell movement is positively correlated in all patterns at distances between 10-50 *µm*, with higher correlation left-to-right for *streams* (≈0.2), left-to-right/front-to-back for *flocks* (≈0.2-0.4), and a lower correlation for *swarms* (≈0.1) (**Fig. Supplementary 17F and S18, S19, S20).** We ascertained that tumor cells within oncostreams migrate in a directional manner ("streams (↑↓)" and "flocks (↑↑)"), while non-oncostream cells move randomly without directional alignment as "swarms". Thus, our analyses strongly indicate that within the tumor core of high-grade glioma cells are dynamically heterogeneous and display organized collective migratory behavior associated with tumor histological and genetic features.

Oncostreams increase the intratumoral spread of tumoral and non-tumoral cells

Pair-wise correlation analysis showed that oncostream glioma cells are collectively organized. To test the underlying nature of collective oncostream motility, we analyzed adherent junction

markers. Tumors with oncostreams were negative for E-cadherin, whereas N-cadherin was strongly expressed (**Fig. Supplementary 21A**), suggesting that these fascicles move in a manner akin to collective migration of mesenchymal cells of the neural crest[33,49]. Although, no difference in N-cadherin were found within oncostreams and the surrounding areas, N-cadherin was elevated in TCGA-GBM (Grade IV) tumors compared to TCGA-LGG (Grade III and II). High levels of N-cadherin correlate with lower survival in HGG patients and mesenchymal transformation (**Fig. Supplementary 21B-C**).

On the other hand, oncostream growth and motility is unlikely to be due to glioma proliferation. BrdU staining showed no differences between oncostream and non-oncostream regions, and in the oncostreams, the mitotic plane was always perpendicular to the main axis as expected (**Fig. Supplementary 21D-E**)**.** These results are also supported by the RNA-Seq data of dissected oncostreams, where proliferation genes were not differentially expressed (**Fig. 4 A-C**).

Collective motion could affect the distribution of other cells within the tumor. Since oncostreams are heterogeneous, we inquired about their pro-tumoral role by potentially spreading cells throughout the tumor. We designed co-implantation experiments using human glioma stem cells (MSP-12), and highly aggressive and oncostream-forming glioma cells (GL26) co-implanted into immunosuppressed mice. Implantation of MSP-12 cells alone generated slow-growing tumors (median survival of 6-8 months). At 21 days post-implantation, MSP-12 cells remained restricted to the injection area with an average distance of 28.9±7.73 µm from the actual injection site. Surprisingly, when MSP-12 cells were co-implanted with GL26-citrine cells, MSP-12 cells spread throughout the tumor, moving along oncostreams to much longer distances (83.7±23.74 µm) from the injection site (**Fig. 7I-K**). Cellular cytoplasmic processes from MSP-12 cells implanted alone displayed a random distribution. However, in co-implanted tumors, such processes from MSP-12 cells are completely aligned with glioma GL26 cells within oncostreams **(Fig. 7K-L and Fig. Supplementary 21F)**. These results strongly suggest that oncostreams function as intra-tumoral highways facilitating the rapid distribution of slow-moving glioma cells and/or non-tumor cells throughout the tumor mass. These findings could help explain the dispersal and intratumoral mixing of diverse clonal populations as demonstrated in previous studies, supporting an important potential role of oncostreams in determining spatial cellular heterogeneity.

Dynamic interactions at the tumor border: oncostreams foster glioma aggressiveness through collective invasion of the normal brain parenchyma

Furthermore, we asked whether oncostreams participate in glioma invasion. The analysis of histological sections showed that multicellular fascicles of elongated and aligned cells are found invading from the tumor border into the normal brain parenchyma (**Supplementary Fig. 22A**). Formation of streams around blood vessels was also observed (**Supplementary Fig. 22A**). These patterns of invasion are also detected using our deep learning methods (**Fig. Supplementary 22B**). We then used our glioma explant model to analyze the invasion dynamics by time-lapse confocal imaging at the tumor border (**Fig. 8A and S23**). We implanted glioma NPA GFP+ cells into tdTomato (mT/mG) mice so tumor borders could be delineated. We observed that glioma cells that extended from the tumor border to the normal brain parenchyma used different dynamic patterns, moving as isolated random cells and/or as collective migratory structures moving directionally, and resembled oncostream structures similar to those in the tumor core (**Fig. 8B-G and Fig. Supplementary 23-26**).

To objectively distinguish between different dynamic patterns, we determined the angle velocity distribution, and the likelihood that distributions corresponded to either a *stream*, a *flock*, or a *swarm*. We found *streams* along the perivascular niche or invading brain parenchyma without following any pre-existing brain structures, as well as cells invading as *flocks*, and *swarms* (**Fig. 8E-F, and Fig. Supplementary 23 D, G, H and S24-S26 A, C, D**). Glioma cells moving along blood vessels or directly into the brain as single cells is consistent with previous studies[10]. The correlation of position and pairwise correlation supports the existence of invading collective motion structures in NPA tumors with high expression of COL1A1 (**Fig. S22D-E and Fig. Supplementary S23J-K and-S24-S26 F-G**). We also determined the participation of collagen fibers in oncostreams invasion. Immunofluorescence analysis on explant slices showed that collagen fibers are aligned along multicellular fascicles of glioma cells invading the normal brain. These data show how collagen fibers serve as scaffolds for collective tumoral cell invasion (**Fig. Supplementary S27**).

Our data indicate the existence of a complex framework of collective motion patterns at the glioma border, that is consistent with previous descriptions[34]. Although the patterns observed at the NPA tumor border are similar to those of the tumor core, cell speed differed between the areas. Cells in the tumor core displayed significantly lower average speeds (*stream*: 4.26; *flock*: 5.95, *swarm*: 6.27

µm/hr) compared to cells at the tumor border or those invading the normal parenchyma (*stream*: 7.95; *flock*: 7.55, *swarm*: 8.01 µm/hr) (**Fig. Supplementary 28A-B**).

Then, we asked whether the knockdown of COL1A1 in NPA gliomas affects changes in the patterns of migration and invasion. Analysis of tumor cells (GFP+) at the tumor borders of GEMM of gliomas comparing NPA and NPA-shCOL1A1 showed a difference in the apparent invasion patterns. The analysis of tumor borders revealed an increase in the sinuosity of NPA tumors, a finding compatible with NPA tumors exhibiting a higher proportion of collective invasion into the normal brain when compared to NPA-shCOL1A1 tumors (**Fig. 8G-I and Supplementary Fig. S29**).

Moreover, the time lapse-confocal imaging and migration analysis of NPA-shCOL1A1 explants showed that tumor cells invade the normal brain parenchyma as isolated cells (**Supplementary Fig. 30-33**). Velocity angle, velocity vector and the likelihood analysis indicated that the overall distribution corresponded predominantly to *swarm* random patterns (**Supplementary Fig. 30-33 D, E, F)**. Further analysis of Relative Position Correlation and Pairwise correlation supports the presence of low density of cells compatible with single cell invasion patterns in NPAshCOL1A1 tumors with low expression of collagen (**Supplementary Fig. S30-33 G, H**).

We conclude that oncostreams (*streams* and *flocks*) are organized collective migratory structures enriched in COL1A1 that participate in the dynamic organization of the tumor microenvironment within the tumor core and at the tumor invasive border of high-grade gliomas, and facilitate invasion into the normal brain, impacting the malignant behavior of gliomas. Depletion of collagen1A1 eliminates oncostreams and their associated functions.

Intravital imaging of glioma reveals the existence of oncostreams' collective motion patterns *in vivo* and their contribution to invasion

To determine whether our previously described collective migration patterns of glioma cells *ex vivo* occur also *in vivo* we performed high resolution time lapse intravital imaging using two photon microscopy. To do so NPA glioma cells were intracranially implanted in the brains of tdTomato (mT/mG) mice at a depth of 0.8 mm (**Fig 9A**). To visualize cell migration we established a cranial window above the injection site (**Fig 9B**). After 7-15 days of tumor growth we acquired z-stack images to obtain a 3D orthogonal view of the tumor growing within the cortex. This allowed us to establish intravital imaging below the brain surface. Next, we selected a position at a depth of >100 µm and proceeded to acquire time lapse images of the tumor growing in the normal parenchyma

at an interval of 5 minutes, for 8-12 hours (**Fig 9C-D, Supplementary Fig. S35A, S36A-B and S37A-B**).

In some cases, to determine the exact anatomical location of the tumor, following intravital imaging, we perfused-fixed the brain and performed fluorescence immuno-histochemistry analysis on paraffin embedded coronal sections which were imaged by confocal microscopy (**Supplementary Fig. S35B-C**). The location of the area imaged by intravital microscopy (>100 µm) is shown. In this figure we can identify the location of the tumor (GFP+ cells) containing parenchymal blood vessels (TdTomato+), and astrocytic processes (GFAP+) (**Supplementary Fig. S35B-C**). Thus, together with **Supplementary Fig. S35A**, we demonstrate imaging below the brain surface and within the normal parenchyma of the brain.

To analyze the movement of GFP+ glioma cells we used Fiji's plug-in Track-Mate. In **Fig. 9** the imaged area was divided into subregions to determine the existence of migration patterns (**Fig. 9E,F**). The analysis of cell migration *in vivo* showed that the glioma cells exhibit organized, nematically aligned cells moving collectively at a depth of 120 µm. Angle velocity distribution analysis determined the existence of '*stream*' collective motion patterns in the three delimited subregions (for example in **Fig. 9H**), illustrating that aligned cells are moving in opposite directions. To corroborate the existence of '*stream'* patterns we applied our likelihood analysis. For '*streams'* the distribution of the angle velocity and velocity vectors displayed a bi-modal distribution (cells were moving in equal but opposite directions) (**Fig. 9I**), similar to that observed in the explant models. Mean speed varied from 5.50 to 9.95 µm/hour (**Fig. 9 G**). A different tumor imaged at a depth of 145 µm also showed the presence of '*streams'* (**Supplementary Fig. S36).**

To analyze the invasion of glioma cells, we imaged tumor movement at the border with normal brain (**Supplementary Fig. S35**). At an imaging depth of 140 µm, cells at the tumor border displayed '*stream'* collective dynamics (**Supplementary Fig. S35D, Zone A and B**). These motion patterns were determined using the angle velocity distribution analysis, and likelihood distribution analysis (**Supplementary Fig. S35 G-H**). Further analysis of Relative Position and Pairwise correlation supports the presence of high density of cells compatible with collective migration and invasion patterns (**Supplementary Fig. S34B-C, S35I-J and S36I-J**).

Further, to determine whether the knockdown of COl1A1 in NPA gliomas alters the migration patterns observed for NPA gliomas we analyzed a NPAshCOL1A1 tumor by two-photon microscopy at a depth of 110 µm (**Supplementary Fig. S37 A-B**).

The intravital migration analysis of NPAshCOL1A1 gliomas showed that glioma cells migrate without a preferred direction, and invade the normal brain parenchyma as '*swarms*' (**Supplementary Fig. S37 C-G**). The alignment analysis, Relative Position, and Pairwise correlation confirm the presence of not-aligned low-density cells compatible with single cell invasion patterns in gliomas with COL1A1 downregulation (**Supplementary Fig. S37 E, J, K).**

The collective motion patterns found *in vivo* resemble the collective motion patterns described in the *ex vivo* explant model. Our results show that glioma cells expressing collagen are organized in collective dynamic patterns at the tumor core and the tumor invasive border, in tumor explants and in *in vivo* intravital models of gliomas analyzed by two photon microscopy.

DISCUSSION

Mesenchymal transformation is a hallmark of tumor heterogeneity that is associated with a more aggressive phenotype and therapeutic resistance[13,18,21]. Mesenchymal transformation involves fibroblast-like morphological changes associated with active migration and gain of expression of mesenchymal genes as previously described[21,22].

Herein we present a comprehensive study that defines the morphological, cellular, dynamic, and molecular properties of multicellular mesenchymal-like structures within gliomas. These structures are fascicles of aligned spindle-like cells found throughout the tumors and represent areas of mesenchymal transformation. We interpret these structures to be the histological expression of areas of collective motion of glioma cells. For the sake of simplicity, we have referred to these areas of mesenchymal transformation as oncostreams.

Oncostreams are areas of mesenchymal transformation and are identified histologically as fascicles of aligned and elongated cells. When examined dynamically, we found that tumor cells move by collective motion within the tumor core and at the invading border. The capacity to identify areas of collective motion in histological sections has allowed us to characterize the molecular organization of such dynamic structures. We thus describe the overall molecular mechanisms that govern the organization and function of these structures and demonstrate the causal role of individual mediators. Surprisingly, we discovered that COL1A1 is central to the structural and dynamic characteristics of oncostreams. Indeed, the loss of COL1A1 expression from tumor cells

disrupts the structural and functional characteristics of oncostreams, resulting in a complete loss of mesenchymal areas within gliomas and a reduction in glioma malignant behavior (**Fig. 10**).

The analysis of the gene ontologies over-represented within oncostreams indicates that oncostreams denote areas enriched for "positive regulation of cell migration", and in mesenchymal related genes. Interestingly, COL1A1 appeared as a central hub of oncostream organization and mesenchymal transformation. We postulate that oncostreams are the histopathological expression of patterns of collective motion (i.e., streams and flocks) in high grade glioma tumors. Different strategies of cell migration encountered in our gliomas are reminiscent of migratory characteristics observed during embryonic development[31,32,49]. In developmental biology, collective motion is represented by cells moving together in clusters, sheets, streams, or other multicellular arrangements[28,31,32].

Our studies of oncostream dynamics at the tumor core are compatible with the results of Ralitza et al[50]. This group studied *ex-vivo* explant slices of spontaneous intestinal carcinoma, and showed that cells within the tumor core were highly dynamic and display directionally correlated cell motion[50], similar to our results described herein. Recent *in silico* based mathematical modelling of glioma cell dynamics by our group, showed that only elongated cells, but not spherical cells, are able to form organized aligned cellular structures in a cell-density dependent manner[47]. Our modeling studies strongly support our *in-vivo* and *ex-vivo* data described in this manuscript.

Moreover, it has been described that increased matrix cross-linking, enzymatic remodeling and parallel orientation of matrix collagen fibers stiffens tissue, modifies cell morphology and promotes cell migration and invasion[36,38,39,51,52]. Our results support the proposal that oncostreams serve as highways to spread tumor, and non-tumor cells, throughout the tumor. Indeed, oncostream fascicles contain higher amounts of macrophages/microglia and mesenchymal cells. Dispersal of tumor and non-tumoral cells throughout the tumors could help explain the mixing of different clonal populations seen in molecular studies of high-grade gliomas[10].

This study contributes to explaining how a particular feature of intratumoral heterogeneity, namely mesenchymal transformation, affects HGG progression. Our data indicate that the density of oncostreams plays a potential role in overall glioma malignant behavior in mouse and human gliomas.

Spatially resolved transcriptional analysis using laser capture microdissection provided novel insights into the molecular mechanisms that regulate oncostream functions. Oncostreams were

defined by a unique transcriptomic signature that matched our immunohistochemical studies. COL1A1 overexpression within oncostreams was complemented with the overexpression of extracellular matrix proteins such as MMP9, MMP10, ADAMTS2, which are known to remodel and participate in the reorganization of collagen fibers. Oncostream fascicles were correspondingly enriched in COL1A1 when assessed by immunohistochemistry.

Within the extracellular matrix, collagen fibers constitute a scaffold for the organization of the tumor microenvironment and thus promote tumor infiltration and invasion. While collagen was previously thought to be a passive barrier that could reduce tumor invasion, it has now been shown that collagen fibers can serve as mechanical and biochemical tracks that facilitate cellular migration and tumor progression[36-38,53]. Previously, multi-cancer computational analysis found that within a mesenchymal transformation signature in different cancers including gliomas, COL1A1 was one of the top differentially expressed genes[18,22,54]. COL1A1 is overexpressed in high grade malignant gliomas and its expression levels are inversely correlated with patient survival[55] as indicated in https://www.cancer.gov/tcga. In our mouse glioma models and in human gliomas, tumors with higher density of oncostreams also express higher levels of COL1A1. COL1A1 is a consistently differentially expressed gene in the glioma mesenchymal signature identified in malignant gliomas and in glioma stem cells as described in previous studies[4,9,56]. Overall, our data are in agreement with a recent study by Puchalski *et al.*, which assigned genetic and transcriptional information to the most common morphological hallmarks of a glioma, emphasizing the importance of integrative histo-molecular studies[8].

Surprisingly, our data indicate a remarkable plasticity of the mesenchymal phenotype in gliomas, similar to other studies[13,15]. Genetic inhibition of COL1A1 within glioma cells depleted COL1A1 from tumors, eliminated oncostream structures, reduced the glioma malignant phenotype, and prolonged animal survival. Our findings are comparable with results from various studies that investigated the *in-vitro* and *in-vivo* consequences of collagen depletion, inhibition of collagen cross-linking or collagen synthesis inhibition on normalizing tumor ECM. In these studies, inhibition of collagen led to changes in the ECM which improved drug penetration, efficacy, as well as tumor access of therapeutic nano-particles or gene based therapies[57-61]. In addition, COL1A1 inhibition within glioma cells induced cell intrinsic and extrinsic changes in the TME. COL1A1 inhibition not only inhibits tumor cell proliferation and migration but also decreased the infiltration of microglia/macrophages, endothelial cells proliferation, and perivascular mesenchymal-like cells. As

595 previously shown by other studies, glioblastomas exhibit complex interactions between tumoral and
596 non-tumoral cells, (including macrophages, immune cells, and endothelial cells), which influence
597 tumor growth, transformation, invasion, and response to treatment[9,13,62]. However, a major
598 remodeling of the tumor mesenchymal phenotype in response to inhibition of COL1A1 has not been
599 described earlier.
600 Moreover, we found that multicellular oncostream fascicles are detected in both *ex vivo* and *in*
601 *vivo* glioma models, and that oncostreams facilitate tumor cell invasion, thereby increasing glioma
602 aggressiveness. Our findings strongly support the importance of collective motility of glioma cells in
603 the progression of tumor growth and invasion of normal brain parenchyma, and is supported by
604 earlier studies of normal and pathological conditions[27,34,35,63-67].
605 In summary, our observations suggest that oncostreams are morphologically and molecularly
606 distinct structures that represent areas of collective motion that contribute to tumor growth and
607 invasion. These malignant dynamic structures overexpress COL1A1. COL1A1 knockdown eliminates
608 oncostreams, reduces the mesenchymal phenotype, modifies the TME and slows tumor
609 progression. Our findings open new vistas to understanding tumor mesenchymal transformation
610 and its therapeutic treatment. We propose that depletion of COL1A1 within glioma cells is a
611 promising approach to reprogram mesenchymal transformation in glioma tumors, and could be
612 harnessed as a novel therapeutic approach, and reduce the glioma malignant phenotype.

613

614 **METHODS**

615 **Glioma cell lines and culture conditions:**

616 Mouse glioma cells (NPA, NPD, NPAshCol1A1, NPDshCol1A1 and GL26) and human glioma cells
617 (MSP-12, SJGBM2) were maintained at 37 °C with 5% CO2 and their respective media as described
618 before[41-44]. Mouse NPA, NPD, NPAshCol1A1, NPDshCol1A1 neurospheres were derived from
619 genetically engineered tumor using the Sleeping Beauty (SB) transposase system as previously
620 described[41-44]. Mouse GL26 glioma cells were generated by Sugiura K and obtained from the frozen
621 stock maintained by the National Cancer Institute (Bethesda, MD)[25]. MSP-12 human glioma cell lines
622 were provided by Christine Brown, City of Hope, and SJGBM2 human glioma cells were provided by
623 Children's Oncology Group (COG) Repository, Health Science Center, Texas Tech University.

624 **Intracranial implantable syngeneic mouse gliomas:**

Glioma tumors were generated by stereotactic intracranial implantation into the mouse striatum of 3.0 x 10^4 mouse glioma cells (either, NPA, NPD or, GL26) in C57BL/6 mice, or human glioma cells in immune-deficient NSG mice (SJGBM2) as described before[42-44,68]. To test whether oncostream tumor cells help move other cells throughout the tumor we generated a co-implantation glioma model by intracranial implantation of highly malignant GL26-citrine cells with low aggressive human MSP12 glioma cells at a ratio of 1:30 (1,000 GL26-citrine cells and 30,000 MSP12 cells) in immune-deficient NSG mice. As controls, NSG mice were implanted with 30,000 MSP12 cells alone or 1,000 GL26-citrine cells alone as controls. Experiments were conducted according to the guidelines approved by the Institutional Animal Care (IACUC) and Use Committee at the University of Michigan. Stereotactic implantation was performed as previously described[42].

Genetically engineered mouse glioma models (GEMM)

We used genetically engineered mouse glioma models for survival analysis and histopathological analysis. Murine glioma tumors harboring different genetic drivers were generated using the Sleeping Beauty (SB) transposon system as described before[41-44]. Genetic modifications were induced in postnatal day 1 (P01) male and female wild-type C57BL/6 mice (Jackson Laboratory), according to IACUC regulations. shRNA targeting the COL1A1 gene was cloned as describe in detail in Supplementary Methods.

Analysis of oncostreams in human glioma tissue

Oncostream presence was analyzed in unidentified H&E sections of paraformaldehyde-fixed paraffin-embedded (PFPE) human glioma samples obtained from primary surgery from the University of Michigan Medical School Hospital. To determine the presence of oncostreams in a large cohort of human glioma tissues we used the biospecimens from “The Cancer Genome Atlas Research Network” (TCGA) from the Genomic Data Commons Data Portal, National Cancer Institute, NIH (https://portal.gdc.cancer.gov). We analyzed primary Glioblastoma multiforme (TCGA-GBM) and Low-Grade Glioma (TCGA-LGG) databases. We selected cases that have available the Slide Image and diagnostic Slides. The diagnostic slides are available for TCGA-GBM: 389 patients and TCGA-LGG: 491 patients. The presence of oncostreams was scored on 100 TCGA-GBM Grade IV tissue samples and 120 TCGA-LGG samples.

Cell aspect ratio and alignment analysis in H&E tumor sections

Images were obtained using bright-field microscopy of H&E stained paraffin sections (Olympus BX53 Upright Microscope). Tumors were imaged using 40X and 20X objectives. Images were processed using the program ImageJ as indicated in detail in supplementary methods.

Deep learning analysis for oncostreams detection on H&E staining of glioma tissue

A fully convolutional neural network (fCNN) was trained in order to identify and segment oncostreams in histologic images[69]. We implemented a U-Net architecture to provide semantic segmentation of glioma specimens using deep learning[70-72]. Our oncostream dataset consisted of images from mouse tissues and open-source images from The Cancer Genome Atlas (TCGA). A total of 109 hematoxylin and eosin (H&E) stained histologic mouse images and 64 from TCGA were reviewed and oncostreams were manually segmented by the study authors (AC, A.E.A and P.R.L.). Images from both datasets were then augmented by randomly sampling regions within each image to generate unique patches (~ 300 patches/image). The location and scale of each patch was randomly chosen to allow for oncostream segmentation to be scale invariant. The analysis is explained in further detail in Supplementary Methods.

Immunohistochemistry on paraffin embedded brain tumors

This protocol was performed as described before[42] and as is detailed in Supplementary Methods. Primary antibodies were diluted at the concentration indicated in Supplementary Table 8. Images were obtained using bright-field microscopy from five independent biological replicates (Olympus BX53 Upright Microscope). Ten different fields of each section were selected at random for study to include heterogeneous tumor areas. For immunofluorescence on paraffin embedded sections from brain tumors images were acquired with a laser scanning confocal microscope (LSM 880, Axio Observer, Zeiss, Germany). Integrated density was determined for the analysis of Col1a1 expression using Image J. For immunohistochemistry on vibratome brain tumor sections were left in 4% paraformaldehyde fixation for 48 hours and then transferred to PBS 0.1% sodium azide for an additional 24 hours at 4°C. A Leica VT100S vibratome was used to obtain 50 µm coronal brain sections. The immunohistochemistry protocol was performed as previously described[73,74].

Laser capture microdissection (LCM) of brain tumors

Malignant glioma tumors were induced by intracranial implantation of dissociated NPA neurospheres in C57BL/6 mice as described above. LCM approach to analyze differential mRNA expression of intra-tumoral glioma heterogeneity was performed as described elsewhere[75].

RNA-Sequencing and bioinformatics analysis

RNA was isolated for laser microdissected tissues using the RNeasy Plus Micro Kit following the manufacturer recommendations (Qiagen). Before library preparation, RNA was assessed for quality using the TapeStation System (Agilent, Santa Clara, CA) using manufacturer's recommended protocols. We obtained a RIN between 6 to 7 after laser microdissection of glioma tissue. A RIN of 6 was determined to be suitable for cDNA library preparation. 0.25 ng to 10 ng of total RNA was used for cDNA library preparation using a kit suitable for RNA isolation at pico-molar concentrations (MARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian) following manufacturer recommended protocol (Clontech/Takara Bio #635005). Sequencing was performed by the UM DNA Sequencing Core, using the Illumina Hi-Seq platform and Bioinformatic analysis were executed by the UM Bioinformatics Core. Differentially expressed genes of all tumors were used for gene ontology (GO), Pathways analysis and genes analysis using iPathwayGuide (Advaita Corporation 2021). Network analysis of the DE genes were achieved using Cytoscape and Reactome App. Network was clustered by Reactome Functional Interaction (FI). Analysis of the expression of COL1A1 in normal tissue and in human gliomas were performed using the dataset of TCGA-GBM and TCGA-LGG from Gliovis (http://gliovis.bioinfo.cnio.es)[46].

Tumor explant glioma model and time-lapse confocal imaging

For the analysis of glioma dynamics, we generated tumors by intracranial implantation of 3 x 10^4 NPA neurospheres which were used to carry out a 3D explant slice culture glioma model. C57BL6 mice were used for the dynamic analyses of the tumor core and B6.129(Cg)-Gt(ROSA)26Sortm4(ACTB-tdTomato,-EGFP)Luo/J- transgenic mice (Jackson laboratory, STOCK 007676) were used for invasion analysis. The red color correspond to a cell membrane-localized tdTomato (mT) fluorescence expression which is widespread in all cells and tissues. Mice were euthanized at 19 days' post-implantation for NPA tumors and 31 days' post-implantation for NPAshCOL1A1 tumors. Brains were removed, dissected, and embedded in a 4% agarose solution and kept on ice for 5 minutes. Then, brains were submerged in ice-cold and oxygenated media (DMEM High-Glucose without phenol red, GibcoTM, USA) and sectioned in a Leica VT100S vibratome (Leica, Buffalo Grove, IL) set to 300 µm in the z-direction. All steps were performed under sterile conditions in a BSL2 laminar flow hood. Brain tumor sections were transferred to laminin-coated Millicel Cell Culture Insert (PICM0RG50, Millipore Sigma, USA) (**Supplementary Fig. S23A**). Tumor slices were maintained in D-MEM F-12 media supplemented with 25% FBS, Penicillin-

Streptomycin 10.000 U/ML at 37 °C with a 5% CO2 atmosphere. After 4-18 hours' media were replaced with DMEM-F12 media supplemented with B27 2%, N2 1%, Normocin 0.2 %, Penicillin-Streptomycin 10.000 U/ML and growth factors EGF and FGF 40 ng/ml. For time-lapse imaging, slices were placed in an incubator chamber of a single photon microscope. We utilized an inverted Zeiss LSM880 laser scanning confocal microscope with AiryScan (Carl Zeiss, Jena, Germany), equipped with an incubation chamber kept at 37 °C with a 5% CO2. To validate the depth of imaging, in some experiments (see **Supplementary Fig. S23A-C**), before time lapse imaging, we obtained high resolution z-stacks with approximate dimensions of z=143.81, x=850.19, y=850.19. Time-lapse images were then obtained every ten minutes for 100-300 cycles, selecting the imaging plane to fall at the middle of the z-stack to avoid imaging at the bottom of the explant. Following movie acquisition, sections were fixed in 4% paraformaldehyde (PFA) for 2 days. Fixed sections were embedded in 2% agarose for H&E and immunohistochemistry analysis. Sections were processed and embedded in paraffin at the University of Michigan Microscopy & Image Analysis Core Facility using a Leica ASP 300 paraffin tissue processor/Tissue-Tek paraffin tissue embedding station (Leica, Buffalo Grove IL). Tumor explants were used for collagenase treatment. Sections were then treated for 48 hours with collagenase (C2399, MilliporeSigma, USA) at a concentration of 5, 10, or 15 units/ml or vehicle control. Following treatment, sections were fixed in 4% paraformaldehyde (PFA) for 2 days.

Cranial window implantation and two photon intravital live imaging *in vivo*:

Craniotomy and cranial window implantations were performed following previously described protocols by us and others[25,76,77]. The protocol was conducted according to the guidelines approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Michigan. Briefly, mice were anesthetized and placed in a stereotactic frame. A craniotomy of 3x3 mm size was made over the right hemisphere between bregma and lambda. $5x10^4$ GFP^+ NPA glioma cells were intracranially implanted at a depth of 0.8mm ventral, near the center of the craniotomy overlying the brain cortex of B6.129(Cg)-Gt(ROSA)26Sortm4(ACTB-tdTomato-EGFP)Luo/J mice. These mice were utilized to identify (in red) normal brain tissue, and thus establish tumor borders. The cranial window was covered with two round microscope cover glasses, and a metal head bar was positioned on the skull posterior to the cranial window. One week post NPA tumor cells' injection and cranial window implantation, and two-weeks post NPAshCOL1A1 implantation, intravital live imaging was performed using a two-photon microscope (Bruker Technology) with a 20X water

immersion objective (Olympus, NA 1.0) for 8-12 hours. To avoid imaging at the brain surface, first we acquired high-resolution 3D z-stacks spanning 0-330 µm depth from the surface of the brain. Z-stacks were then imported to Imaris viewer version 9.8 (Bitplane, Imaris, Oxford Instruments, MA, USA) to obtain 3D images. We then used Orthoslicer3D to reveal the depth of the imaging position in relation to the surface of the brain for time-lapse data acquisition. The detailed methodology is available in Supplementary Methods.

Mathematical analysis of tumor cell movement

To determine the movement of cells in different areas of the tumor we performed localized statistical analysis in different zones of the tumor. We selected localized areas based on the organization of cells in clusters, group of cells moving together with similar distribution. Raw data of 4 movies from the tumor core and 4 movies from the tumor border were analyzed for 293 cycles (core) and 186 cycles (border) for a frame rate of $\Delta t = 10$ min between image acquisition. To track cell motion, we used the software Fiji with the plugin TrackMate[78]. Analysis was performed as indicated in detail in Supplementary Methods.

Classification of glioma migration patterns

To classify the collective cellular motion behavior of the three types of patterns called flock, stream, and swarm illustrated in Supplementary Figure S10F we used as criteria the orientation of each cell described by its unique angle velocity denoted θi. More precisely, we transformed the Angle Velocity Distribution graph into a histogram where we examined the distribution of all the values θi. A schematic representation of these distributions is depicted in Figure 4G. Considering a data-set θn $n=1…N$ of orientations where N is the total number of cells, $\theta n \in [0, 2\pi]$ is the direction of the cell n. We tested three types of distributions ρ to describe the dataset and gave a likelihood in each case as described in Supplementary Methods. The Akaike Weight (AW) indicates which pattern has the highest likelihood in each experimental situation[79].

Statistical Analysis

All in vivo experiments were performed using independent biological replicates, as indicated in the text and figures for each experiment. Data are shown as the mean ± SEM. Any difference was considered statistically significant when $p < 0.05$. In experiments that included one variable, the one-way ANOVA test was used. In experiments with two independent variables, the two-way ANOVA test was employed. A posterior Tukey's multiple comparisons test was used for mean comparisons. Student's t-test was used to compare unpaired data from two samples. Survival data were entered into Kaplan-Meier survival curves plots, and statistical analysis was performed using the Mantel log-rank test. Median survival is expressed as MS. Significance was determined if $p<0.05$. All analyses were conducted using GraphPad Prism (version 8.0.0) or SAS (2021 SAS Institute, Cary, NC). Each statistical test used is indicated within the figure legends.

Data availability: All data associated with this study are in the paper and/or the Supplementary Information. RNA-Seq data was deposited at the NCBI's Gene Expression Omnibus (GEO) with identifier GSE188970. Source data are provided with this paper. Further information and requests for resources and reagents should be directed to and will be fulfilled by the corresponding author P.R. Lowenstein.

Code Availability: The analysis of oncostreams in mouse and human glioma tissue was performed using U-Net architecture to provide semantic segmentation of specimens using deep learning. Public GitHub repository for the project code can be found at https://github.com/MLNeurosurg/DeepStreams.

Analysis of glioma cells dynamics was performed using the Julia Programing Language. Link for this project Script and their dependencies can be found at public GitHub repository https://github.com/smotsch/analysis_glioma.

REFERENCES

1 Parsons, D. W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science (New York, N.Y.)* **321**, 1807-1812, doi:10.1126/science.1164382 (2008).

2 Louis, D. N. *et al.* The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathologica* **114**, 97-109, doi:10.1007/s00401-007-0243-4 (2007).

3 Ceccarelli, M. *et al.* Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell* **164**, 550-563, doi:10.1016/j.cell.2015.12.028 (2016).

4 Verhaak, R. G. *et al.* Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **17**, 98-110, doi:10.1016/j.ccr.2009.12.020 (2010).

5 Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 4009-4014, doi:10.1073/pnas.1219747110 (2013).

6 Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell* **155**, 462-477, doi:10.1016/j.cell.2013.09.034 (2013).
7 Nicholson, J. G. & Fine, H. A. Diffuse Glioma Heterogeneity and Its Therapeutic Implications. *Cancer discovery* **11**, 575-590, doi:10.1158/2159-8290.Cd-20-1474 (2021).
8 Puchalski, R. B. *et al.* An anatomic transcriptional atlas of human glioblastoma. *Science (New York, N.Y.)* **360**, 660-663, doi:10.1126/science.aaf2666 (2018).
9 Wang, Q. *et al.* Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* **32**, 42-56.e46, doi:10.1016/j.ccell.2017.06.003 (2017).
10 Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (New York, N.Y.)* **344**, 1396-1401, doi:10.1126/science.1254257 (2014).
11 Bergmann, N. *et al.* The Intratumoral Heterogeneity Reflects the Intertumoral Subtypes of Glioblastoma Multiforme: A Regional Immunohistochemistry Analysis. *Front Oncol* **10**, 494, doi:10.3389/fonc.2020.00494 (2020).
12 Garofano, L. *et al.* Pathway-based classification of glioblastoma uncovers a mitochondrial subtype with therapeutic vulnerabilities. *Nat Cancer* **2**, 141-156, doi:10.1038/s43018-020-00159-4 (2021).
13 Kim, Y. *et al.* Perspective of mesenchymal transformation in glioblastoma. *Acta Neuropathol Commun* **9**, 50, doi:10.1186/s40478-021-01151-4 (2021).
14 Hara, T. *et al.* Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* **39**, 779-792.e711, doi:10.1016/j.ccell.2021.05.002 (2021).
15 Neftel, C. *et al.* An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* **178**, 835-849.e821, doi:10.1016/j.cell.2019.06.024 (2019).
16 Majc, B. *et al.* Epithelial-to-mesenchymal transition as the driver of changing carcinoma and glioblastoma microenvironment. *Biochim Biophys Acta Mol Cell Res* **1867**, 118782, doi:10.1016/j.bbamcr.2020.118782 (2020).
17 Azam, Z., To, S. T. & Tannous, B. A. Mesenchymal Transformation: The Rosetta Stone of Glioblastoma Pathogenesis and Therapy Resistance. *Adv Sci (Weinh)* **7**, 2002015, doi:10.1002/advs.202002015 (2020).
18 Behnan, J., Finocchiaro, G. & Hanna, G. The landscape of the mesenchymal signature in brain tumours. *Brain* **142**, 847-866, doi:10.1093/brain/awz044 (2019).
19 Yang, J. *et al.* Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* **21**, 341-352, doi:10.1038/s41580-020-0237-9 (2020).
20 Williams, E. D., Gao, D., Redfern, A. & Thompson, E. W. Controversies around epithelial-mesenchymal plasticity in cancer metastasis. *Nat Rev Cancer* **19**, 716-732, doi:10.1038/s41568-019-0213-x (2019).
21 Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318-325, doi:10.1038/nature08712 (2010).
22 Cheng, W. Y., Kandel, J. J., Yamashiro, D. J., Canoll, P. & Anastassiou, D. A multi-cancer mesenchymal transition gene expression signature is associated with prolonged time to recurrence in glioblastoma. *PLoS One* **7**, e34705, doi:10.1371/journal.pone.0034705 (2012).
23 Ray, A., Morford, R. K., Ghaderi, N., Odde, D. J. & Provenzano, P. P. Dynamics of 3D carcinoma cell invasion into aligned collagen. *Integr Biol (Camb)* **10**, 100-112, doi:10.1039/c7ib00152e (2018).
24 Pickup, M. W., Mouw, J. K. & Weaver, V. M. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep* **15**, 1243-1253, doi:10.15252/embr.201439246 (2014).
25 Baker, G. J. *et al.* Mechanisms of glioma formation: iterative perivascular glioma growth and invasion leads to tumor progression, VEGF-independent vascularization, and resistance to antiangiogenic therapy. *Neoplasia* **16**, 543-561, doi:10.1016/j.neo.2014.06.003 (2014).
26 Deisboeck, T. S. & Couzin, I. D. Collective behavior in cancer cell populations. *Bioessays* **31**, 190-197, doi:10.1002/bies.200800084 (2009).

27 Friedl, P., Locker, J., Sahai, E. & Segall, J. E. Classifying collective cancer cell invasion. *Nat Cell Biol* **14**, 777-783, doi:10.1038/ncb2548 (2012).
28 Friedl, P. & Mayor, R. Tuning Collective Cell Migration by Cell-Cell Junction Regulation. *Cold Spring Harb Perspect Biol* **9**, doi:10.1101/cshperspect.a029199 (2017).
29 Méhes, E. & Vicsek, T. Collective motion of cells: from experiments to models. *Integr Biol (Camb)* **6**, 831-854, doi:10.1039/c4ib00115j (2014).
30 Rørth, P. Fellow travellers: emergent properties of collective cell migration. *EMBO Rep* **13**, 984-991, doi:10.1038/embor.2012.149 (2012).
31 Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat Rev Mol Cell Biol* **10**, 445-457, doi:10.1038/nrm2720 (2009).
32 Scarpa, E. & Mayor, R. Collective cell migration in development. *J Cell Biol* **212**, 143-155, doi:10.1083/jcb.201508047 (2016).
33 Theveneau, E. & Mayor, R. Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Dev Biol* **366**, 34-54, doi:10.1016/j.ydbio.2011.12.041 (2012).
34 Alieva, M. *et al.* Intravital imaging of glioma border morphology reveals distinctive cellular dynamics and contribution to tumor cell invasion. *Scientific reports* **9**, 2054, doi:10.1038/s41598-019-38625-4 (2019).
35 Haeger, A., Krause, M., Wolf, K. & Friedl, P. Cell jamming: collective invasion of mesenchymal tumor cells imposed by tissue confinement. *Biochim Biophys Acta* **1840**, 2386-2395, doi:10.1016/j.bbagen.2014.03.020 (2014).
36 Brett, E. A., Sauter, M. A., Machens, H. G. & Duscher, D. Tumor-associated collagen signatures: pushing tumor boundaries. *Cancer Metab* **8**, 14, doi:10.1186/s40170-020-00221-w (2020).
37 Mammoto, T. *et al.* Role of collagen matrix in tumor angiogenesis and glioblastoma multiforme progression. *Am J Pathol* **183**, 1293-1305, doi:10.1016/j.ajpath.2013.06.026 (2013).
38 Fang, M., Yuan, J., Peng, C. & Li, Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol* **35**, 2871-2882, doi:10.1007/s13277-013-1511-7 (2014).
39 Payne, L. S. & Huang, P. H. The pathobiology of collagens in glioma. *Mol Cancer Res* **11**, 1129-1140, doi:10.1158/1541-7786.MCR-13-0236 (2013).
40 Pointer, K. B. *et al.* Association of collagen architecture with glioblastoma patient survival. *J Neurosurg* **126**, 1812-1821, doi:10.3171/2016.6.JNS152797 (2017).
41 Calinescu, A. A. *et al.* Transposon mediated integration of plasmid DNA into the subventricular zone of neonatal mice to generate novel models of glioblastoma. *Journal of visualized experiments : JoVE*, doi:10.3791/52443 (2015).
42 Comba, A. *et al.* Fyn tyrosine kinase, a downstream target of receptor tyrosine kinases, modulates antiglioma immune responses. *Neuro-oncology* **22**, 806-818 (2020).
43 Koschmann, C. *et al.* ATRX loss promotes tumor growth and impairs nonhomologous end joining DNA repair in glioma. *Sci Transl Med* **8**, 328ra328, doi:10.1126/scitranslmed.aac8228 (2016).
44 Núñez, F. J. *et al.* IDH1-R132H acts as a tumor suppressor in glioma via epigenetic up-regulation of the DNA damage response. **11**, eaaq1427, doi:10.1126/scitranslmed.aaq1427 %J Science Translational Medicine (2019).
45 Brat, D. J. *et al.* Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N Engl J Med* **372**, 2481-2498, doi:10.1056/NEJMoa1402121 (2015).
46 Bowman, R. L., Wang, Q., Carro, A., Verhaak, R. G. & Squatrito, M. GlioVis data portal for visualization and analysis of brain tumor expression datasets. *Neuro-oncology* **19**, 139-141, doi:10.1093/neuonc/now247 (2017).
47 Jamous, S., Comba, A., Lowenstein, P. R. & Motsch, S. Self-organization in brain tumors: How cell morphology and cell density influence glioma pattern formation. *PLoS computational biology* **16**, e1007611 (2020).

48 Bialek, W. *et al.* Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 4786-4791, doi:10.1073/pnas.1118633109 (2012).

49 Szabó, A. & Mayor, R. Mechanisms of Neural Crest Migration. *Annu Rev Genet* **52**, 43-63, doi:10.1146/annurev-genet-120417-031559 (2018).

50 Staneva, R. *et al.* Cancer cells in the tumor core exhibit spatially coordinated migration patterns. *Journal of cell science* **132**, doi:10.1242/jcs.220277 (2019).

51 Egeblad, M., Rasch, M. G. & Weaver, V. M. Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol* **22**, 697-706, doi:10.1016/j.ceb.2010.08.015 (2010).

52 Provenzano, P. P. *et al.* Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med* **4**, 38, doi:10.1186/1741-7015-4-38 (2006).

53 Shintani, Y., Hollingsworth, M. A., Wheelock, M. J. & Johnson, K. R. Collagen I promotes metastasis in pancreatic cancer by activating c-Jun NH(2)-terminal kinase 1 and up-regulating N-cadherin expression. *Cancer research* **66**, 11745-11753, doi:10.1158/0008-5472.CAN-06-2322 (2006).

54 Kim, H., Watkinson, J., Varadan, V. & Anastassiou, D. Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics* **3**, 51, doi:10.1186/1755-8794-3-51 (2010).

55 Glioma through the looking GLASS: molecular evolution of diffuse gliomas and the Glioma Longitudinal Analysis Consortium. *Neuro-oncology* **20**, 873-884, doi:10.1093/neuonc/noy020 (2018).

56 Behnan, J. *et al.* Differential propagation of stroma and cancer stem cells dictates tumorigenesis and multipotency. *Oncogene* **36**, 570-584, doi:10.1038/onc.2016.230 (2017).

57 Dolor, A. & Szoka, F. C., Jr. Digesting a Path Forward: The Utility of Collagenase Tumor Treatment for Improved Drug Delivery. *Mol Pharm* **15**, 2069-2083, doi:10.1021/acs.molpharmaceut.8b00319 (2018).

58 Diop-Frimpong, B., Chauhan, V. P., Krane, S., Boucher, Y. & Jain, R. K. Losartan inhibits collagen I synthesis and improves the distribution and efficacy of nanotherapeutics in tumors. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 2909-2914, doi:10.1073/pnas.1018892108 (2011).

59 Kumar, S., Henning-Knechtel, A., Magzoub, M. & Hamilton, A. D. Peptidomimetic-Based Multidomain Targeting Offers Critical Evaluation of Aβ Structure and Toxic Function. *J Am Chem Soc* **140**, 6562-6574, doi:10.1021/jacs.7b13401 (2018).

60 Kato, M., Hattori, Y., Kubo, M. & Maitani, Y. Collagenase-1 injection improved tumor distribution and gene expression of cationic lipoplex. *Int J Pharm* **423**, 428-434, doi:10.1016/j.ijpharm.2011.12.015 (2012).

61 Lee, S. *et al.* Extracellular matrix remodeling in vivo for enhancing tumor-targeting efficiency of nanoparticle drug carriers using the pulsed high intensity focused ultrasound. *J Control Release* **263**, 68-78, doi:10.1016/j.jconrel.2017.02.035 (2017).

62 Hambardzumyan, D., Gutmann, D. H. & Kettenmann, H. The role of microglia and macrophages in glioma maintenance and progression. *Nat Neurosci* **19**, 20-27, doi:10.1038/nn.4185 (2016).

63 Gritsenko, P. G. & Friedl, P. Adaptive adhesion systems mediate glioma cell invasion in complex environments. *Journal of cell science* **131**, doi:10.1242/jcs.216382 (2018).

64 Gritsenko, P. G. *et al.* p120-catenin-dependent collective brain infiltration by glioma cell networks. *Nat Cell Biol* **22**, 97-107, doi:10.1038/s41556-019-0443-x (2020).

65 Peglion, F., Llense, F. & Etienne-Manneville, S. Adherens junction treadmilling during collective migration. *Nat Cell Biol* **16**, 639-651, doi:10.1038/ncb2985 (2014).

66 Lim, D. A. & Alvarez-Buylla, A. The Adult Ventricular-Subventricular Zone (V-SVZ) and Olfactory Bulb (OB) Neurogenesis. *Cold Spring Harb Perspect Biol* **8**, doi:10.1101/cshperspect.a018820 (2016).

67 Waclaw, B. *et al.* A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature* **525**, 261-264, doi:10.1038/nature14971 (2015).

68 Kamran, N. *et al.* Current state and future prospects of immunotherapy for glioma. *Immunotherapy* **10**, 317-339, doi:10.2217/imt-2017-0122 (2018).

69 Shelhamer, E., Long, J. & Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans Pattern Anal Mach Intell* **39**, 640-651, doi:10.1109/tpami.2016.2572683 (2017).

70 Sharma, M., Pachori, R. & Acharya, U. R. Adam: a method for stochastic optimization. *Pattern Recognit. Lett* **94**, 172-179 (2017).

71 Hollon, T. C. *et al.* Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat Med* **26**, 52-58, doi:10.1038/s41591-019-0715-9 (2020).

72 Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Springer International Publishing* 234-241 (2015).

73 Wilson, T. J., Zamler, D. B., Doherty, R., Castro, M. G. & Lowenstein, P. R. Reversibility of glioma stem cells' phenotypes explains their complex in vitro and in vivo behavior: Discovery of a novel neurosphere-specific enzyme, cGMP-dependent protein kinase 1, using the genomic landscape of human glioma stem cells as a discovery tool. *Oncotarget* **7**, 63020-63041, doi:10.18632/oncotarget.11589 (2016).

74 Yadav, V. N. *et al.* CXCR4 increases in-vivo glioma perivascular invasion, and reduces radiation induced apoptosis: A genetic knockdown study. *Oncotarget* **7**, 83701-83719, doi:10.18632/oncotarget.13295 (2016).

75 Comba, A. *et al.* Laser Capture Microdissection of Glioma Subregions for Spatial and Molecular Characterization of Intratumoral Heterogeneity, Oncostreams, and Invasion. *JoVE (Journal of Visualized Experiments)*, e60939 (2020).

76 Huber, D. *et al.* Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature* **484**, 473-478, doi:10.1038/nature11039 (2012).

77 Chen, Z., Ross, J. L. & Hambardzumyan, D. Intravital 2-photon imaging reveals distinct morphology and infiltrative properties of glioblastoma-associated macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 14254-14259, doi:10.1073/pnas.1902366116 (2019).

78 Tinevez, J. Y. *et al.* TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80-90, doi:10.1016/j.ymeth.2016.09.016 (2017).

79 Kenneth P. Burnham, D. R. A. *Model selection and multimodel inference: A practical information-theoretic approach*. 2 edn, (Springer, 2002).

ACKNOWLEDGMENTS

We thank all members of our laboratory for advice and comments on this work. This work was supported by National Institutes of Health, National Institute of Neurological Disorders and Stroke (NIH/NINDS) grants: R37-NS094804, R01-NS105556, R21-NS107894, R21-NS091555; R01-NS074387 to M.G.C.; National Institute of Neurological Disorders and Stroke (NIH/NINDS) grants: R01-NS076991, R01-NS096756, R01-NS082311, R01-NS122234, R01-NS127378 to P.R.L.; National Institute of Biomedical Imaging and Bioengineering (NIH/NIBI): R01-EB022563; National Cancer Institute (NIH/NCI) U01CA224160; Rogel Cancer Center at The University of Michigan G023089 to M.G.C. Ian's Friends Foundation grant G024230, Leah's Happy Hearts Foundation grant G013908, Pediatric Brain Tumor Foundation grant G023387 and ChadTough Foundation grant G023419 to P.R.L. RNA Biomedicine grant: F046166 to M.G.C. Health and Human Services, National Institutes of Health, UL1 TR002240 to Michigan Institute for Clinical and Health Research (MICHR), Postdoctoral Translational Scholars Program (PTSP), Project F049768 to A.C.

AUTHORS INFORMATION

Affiliations:

Dept. of Neurosurgery, University of Michigan Medical School, Ann Arbor, 48109, MI, USA
Andrea Comba, Syed M. Faisal, Patrick J. Dunn, Todd C. Hollon, Wajd N. Al-Holou, Maria Luisa Varela, Daniel B. Zamler, Anna E. Argento, Phillip E. Kish, Clifford Abel II, Maria G Castro, Pedro R. Lowenstein

Dept. of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, 48109 MI, USA
Andrea Comba, Syed M. Faisal, Patrick J. Dunn, Maria Luisa Varela, Daniel B. Zamler, Anna E. Argento, Clifford Abel II, Maria G Castro, Pedro R. Lowenstein

Rogel Cancer Center, University of Michigan medical School, Ann Arbor, 48109, MI, USA
Andrea Comba, Syed M. Faisal, Patrick J. Dunn, Maria Luisa Varela, Daniel B. Zamler, Clifford Abel II, Maria G Castro, Pedro R. Lowenstein

School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, USA
Sebastien Motsch

Ophthalmology & Visual Science, University of Michigan Medical School, Ann Arbor, 48109 MI, USA
Phillip E. Kish, Alon Kahana

Dept. of Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA
Celina G. Kleer

Kresge Hearing Research Institute, Department of Otolaryngology-Head & Neck Surgery, University of Michigan medical School, Ann Arbor, MI 48109, USA
Gunnar L Quass, Pierre F. Apostolides

Dept. of Molecular & Integrative Physiology, University of Michigan medical School, Ann Arbor, MI 48109, USA
Pierre F. Apostolides

Departments of Hematology & Hematopoietic Cell Transplantation and Immuno-Oncology, City of Hope, Duarte, CA, USA
Christine E. Brown

Contributions: Conception and design: A. Comba, M.G. Castro, P. R. Lowenstein. Development of methodology: A. Comba, M.S. Faisal, P. J. Dunn, A. E. Argento, T. Hollon, W.N. Al-Holou, M.L. Varela, D.B. Zamler, S. Motsch, P. R. Lowenstein. Acquisition of data, analysis, and interpretation: A. Comba, M.S. Faisal, P. J. Dunn, A. E. Argento, T. Hollon, M.L. Varela, D.B. Zamler, Clifford Abel II, M.G. Castro, S. Motsch, P. R. Lowenstein. Human histopathology analysis and identification of oncostreams: A. Comba, C. Kleer, A. E. Argento, P. R Lowenstein. Laser microdissection protocol: A. Comba, P. R. Lowenstein, P. E. Kish, Alon Kahana. Development and establishment of intravital imaging using

multiphoton microscopy: Comba, M.S. Faisal, P. R. Lowenstein, G. L. Quass, P. F. Apostolides. Development and experimental assistance with human glioma cell lines. development and assistance: C.E. Brown. Manuscript writing: A. Comba, S. Motsch, P. R. Lowenstein. Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A. Comba, P. R. Lowenstein. Study supervision: M. G. Castro and P. R. Lowenstein. All authors reviewed the final version of the manuscript.

Corresponding author:

Correspondence to Pedro R. Lowenstein

ETHICS DECLARATIONS

Competing interests: All authors of this paper declare no potential conflicts of interest.

Figure Legends

Fig. 1. Oncostreams are multicellular fascicles present in mouse and human gliomas
A) Representative 5 µm H&E microtome sections from gliomas showing that fascicles of spindle-like glioma cells (oncostreams, outlined by the dotted line) are present in a GEMM of gliomas NPA (NRAS/shATRx/shp53) and the GL26 intracranial implantable model of glioma. Scale bars: 50 µm. **B)** Representative H&E microtome sections of human glioma and human xenografts showing the presence of oncostreams. Scale bar: 20 µm. **C-D)** Histograms showing the cellular shape analysis (aspect ratio) (**C**) and angle orientation (alignment) for the corresponding images (**D**) show areas of oncostreams (OS) formed by elongated and aligned cells and areas with no oncostreams (No-OS) as rounded and not-aligned cells. **E-G)** Immunostaining shows that tumor cells, mesenchymal cells (ACTA2+), microglia/macrophages (IBA1+ and CD68+), are aligned within, the main orientation axis of oncostreams. Bar graphs show the quantification of ACTA+, n=3 (**E**), IBA1+, n=5 (**F**) and CD68+, n=5 (**G**) cells within oncostreams areas in NPA tumors. 6-13 areas of oncostreams per tumor section per animal were imaged. Scale bar: 20 µm. Error bars represent ± SEM; unpaired t-test analysis, *p<0.05, **p<0.001. **H)** Angle orientation shows the alignment of ACTA+, IBA1+ and CD68+ cells within oncostreams for the corresponding images.

Fig. 2. Oncostreams density positively correlates with tumor aggressiveness in GEMM of gliomas.
A) Genetic makeup of NPA and NPAI tumors. **B)** Kaplan–Meier survival curves of NPA and NPAI mouse gliomas show that animals bearing IDH1-R132H mutant tumors (NPAI) have prolonged median survival (MS): **NPA** (MS: 86 days; n: 18) versus **NPAI** (MS: 213 days; n:12). Log-rank (Mantel-Cox) test; ****p<0.0001. **C-D)** Deep learning method for oncostream detection in H&E stained mouse glioma sections: **C)** Representative images of oncostreams manually segmented on H&E stained sections of NPA gliomas and NPAI tumors. The output of our trained model for each image is shown below (probability heat maps), for tissues containing oncostreams (NPA), and without oncostreams (NPAI), scale bar = 50 µm. **D)** 10-14 random fields per tumor section per animal were imaged, n=9 NPA and n=12 NPAI, and quantified using deep learning analysis. Error bars represent ± SEM; unpaired t-test

analysis, *p<0.05. **E)** Angle histogram plots show aligned cells in NPA tumors vs non-aligned cells in NPAI tumors for the representative images showed in figure.

Fig. 3. The density of oncostreams positively correlates with tumor aggressiveness in human gliomas. A) TCGA tumors were analyzed from different grade: GBM-Grade IV (100 tumors). LGG-Grade III (70 tumors) and LGG-Grade II (50 tumors). Pie charts show percentage of tumors displaying oncostreams in relation to tumor grade. Oncostreams are present in 47% of GBM grade IV tumors, 8.6 % of LGG grade III, and are absent from LGG grade II. **B)** Manual identification of oncostreams in H&E images are shown for human gliomas with WHO grades IV, III, II from TCGA. **C)** Deep learning analysis for human gliomas. Our algorithm was able to detect oncostreams in grade IV and III gliomas but not in grade II gliomas. **D)** Angle histogram plots show the alignment of cells in H&E histology sections of Grade IV and Grade III gliomas' oncostreams and random alignment in grade II glioma sections lacking oncostreams. Angle histogram correspond to the representative images. **E-F)** Immuno-fluorescence staining of SOX2+ tumor cells (green), glial fibrillary acidic protein (GFAP+) cells (red), and microglia/macrophage (IBA1+) cells (red) in high-grade human glioblastoma (GBM) (WHO Grade IV), IDH-WT **(E)** and in low-grade glioma (LGG) (WHO Grade III), IDH-mutant **(F)**, showing oncostreams heterogeneity and cellular alignment of these cells in human high-grade gliomas but not in low grade gliomas (arrows). Scale bars: 50 µm.

Fig. 4. Oncostreams are defined by a unique gene expression signature related to mesenchymal transformation and migration. A) (a) Schematic representation of spatial transcriptomic analysis of glioma oncostreams using Laser Capture Microdissection (LCM). Glioma tumors were generated by intracranial implantation of NPA tumor cells in C57BL6 mice. (b-c) Oncostream areas (red outline) were identified and dissected from surrounding glioma tissue (black outline) in mouse glioma samples using a LCM microscope. **B)** A volcano plot displays differentially expressed (DE) genes from oncostream *vs* no-oncostream areas. DE genes were selected based on a fold change of ≥ 1.5 and a q-value (false discovery rate (FDR) corrected p-value) of ≤ 0.05. Upregulated genes (red dots) and downregulated genes (green dots) are shown. Relevant genes related to mesenchymal migration are labeled on the graph. **C)** Heat map illustrates gene expression patterns for oncostream vs no-oncostream areas in NPA glioma tumors (n=3 biological replicates/group). Differentially upregulated genes (16) are represented in red and downregulated genes (n=27) are represented in green (q-value ≤ 0.05 and fold change ≥ ± 1.5). **D)** Functional enrichment analysis of overrepresented GO terms (biological processes) obtained when comparing oncostream vs no-oncostream DE genes. p-value corrected for multiple comparisons using the FDR method. Cutoff FDR<0.05. Blue: Downregulated GOs. Red: upregulated GOs. **E)** Bar graphs show DE genes annotated to the most relevant enriched GOs biological process: "Positive regulation of cell motility", "Regulation of cell migration" and "Collagen metabolic process."

Fig. 5. COL1A1 is a central hub in oncostream organization and glioma malignancy. A) Network analysis of the DE genes comparing oncostreams versus no-oncostreams DE genes. Genes with a higher degree of connectivity are highlighted with larger nodes. Clusters of nodes with the same color illustrate modules of highly interacting of genes in the network. **B)** Immunofluorescence analysis of COL1A1 expression in GEMM of glioma tissues comparing NPA (IDH1-WT) vs NPAI (IDH1mut). Representative confocal images display COL1A1 expression in green (Alexa 488) and nuclei in blue (DAPI). Scale bar: 50 µm. **C)** Bar graphs represent COL1A1 quantification in terms of fluorescence integrated density. NPA n=5 and NPAI n=6 animals for each experimental condition

were used for the analysis. Ten fields of each tumor section were selected at random. Error bars represent ±SEM. t-test, **p<0.01. **D)** Immunofluorescence analysis of COL1A1 expression in human GBM and LGG tumors. COL1A1 expression in green (Alexa 488) and nuclei in blue (DAPI). Scale bar: 50 µm. **E)** Bar graphs represent COL1A1 quantification as fluorescence integrated density. 5 (LGG) and 8 (GBM) tumor samples were used for the analysis. Ten fields of each tumor section were selected at random. Error bars represent ±SEM. t-test, **p<0.01. **F-G)** GEMM of glioma with COL1A1 inhibition. **F)** Kaplan–Meier survival curve comparing NPA (MS: 68 days; n: 14) vs NPAshCOL1A1 (MS: 123 days; n: 28) **G)** Kaplan–Meier survival curve comparing NPD (MS: 74 days; n=15) versus NPDshCOL1A1 (MS: 98 days; n=17). Log-rank (Mantel-Cox) test. **** p<0.0001, **p<0.0126. **H)** Immunofluorescence analysis of COL1A1 expression in GEMM of glioma controls (NPA and NPD) and Col1A1 downregulation (NPAshCOL1A1 and NPDshCOL1A1). Representative confocal images of COL1A1 expression in green (Alexa 488) and nuclei in blue (DAPI). Arrows indicate COL1A1 enriched perivascular cells. Scale bar: 50 µm. **I)** Bar graphs represent COL1A1 quantification in terms of fluorescence integrated density. 5-7 tumor samples for each experimental condition were used for the analysis. Ten fields of each tumor section were selected at random. Error bars represent ±SEM. t-test, **p<0.01. **J)** Representative images of the histopathological identification of oncostreams in H&E tissue sections comparing the COL1A1 knockdown tumors with their respective controls. Scale bars: 50 µm. **K)** Quantitative analysis of oncostream areas using deep learning analysis. 4-12 tumor samples for each experimental condition were used for the analysis. Error bars represent ± SEM; unpaired t-test analysis, *p<0.05.

Fig. 6. Knockdown of COL1A1 within glioma cells modifies the tumor microenvironment
Immunohistochemical analysis (**A**, **C** and **E**) of GEMM of glioma controls (**NPA** and **NPD**) and COL1A1 downregulation (**NPAshCOL1A1** and **NPDshCOL1A1**). **A)** Representative images of PCNA expression. Scale bar: 20 µm. **B)** Bar graphs represent the quantification of PCNA+ cells numbers (cells/mm^2) using QuPath positive cell detection. Error bars represent ±SEM, (NPA: n=8, NPAshCOL1A1: n=5, NPD: n=8, NPDshCOL1A1: n=4), t-test, *p<0.05. **C)** Representative images of CD68 expression. Scale bar: 20 µm**. D)** Bar graphs represent CD68+ cell quantification (cells/mm^2) using QuPath positive cell detection. Error bars represent ±SEM, (NPA: n=8, NPAshCOL1A1: n=5, NPD: n=6, NPDshCOL1A1: n=4), t-test, *p<0.05, ns: no significant. **E)** Representative images of CD31 expression. Scale bar: 20 µm**. F)** Bar graphs represent CD31+ cells quantification (cells/mm^2) using QuPath positive cells detection. Error bars represent ±SEM, (NPA: n=8, NPAshCOL1A1: n=5, NPD: n=6, NPDshCOL1A1: n=4), t-test, **p<0.01, *p<0.05. **G)** Immunofluorescence analysis of GEMM of glioma controls (**NPA** and **NPD**) and COL1A1 downregulation (**NPAshCOL1A1** and **NPDshCOL1A1**). Representative images of ACTA2 expression in red (Alexa 555) and nuclei in blue (DAPI). Scale bar: 50 µm. **H)** Bar graphs represent ACTA2 quantification in terms of fluorescence integrated density. Error bars represent ±SEM, (NPA: n=6, NPAshCOL1A1: n=5, NPD: n=4, NPDshCOL1A1: n=3), t-test, **p<0.01, ns: no significant.

Fig. 7. Collective dynamics of oncostreams increase cell spreading within the tumor core. A) Experimental setup: NPA-GFP glioma cells were intracranially implanted in C57BL6 mice. Explant slice cultures of growing tumors were used for confocal time-lapse imaging of the tumor core. **B)** Single representative time-lapse confocal image of glioma cells within the tumor core (Movie #1). **C)** Tracking analysis of individual cell paths performed using the Track-Mate plugin from Image-J. **D)** Preferred directions of cells within three zones (A-C) superimposed onto a representative time lapse-

image. **E)** Speed distribution and mean speed (*μm/hr*) in Zones A, B and C. **F)** Distribution of angle velocity for each zone. The Angle Velocity of each cell is denoted ϑ. The plot shows the proportion of cells moving in angle direction ϑ for each zone. **G-H)** Classification of collective motion patterns: *stream*, *flock* or *swarm*. The distribution is *uni-modal* for a flock (only one peak) and *bi-model* for a stream (two peaks = 2 preferred angle velocity). For a *swarm*, the distribution is *flat* (no preferred angle velocity). In **(G)** Angle Velocity was transformed to a histogram; these data were then used to calculate the likelihood that a particular distribution of velocity angles corresponds to either a *stream*, *flock*, or *swarm*. The results are given in **(H)** for each zone. The frequency distribution of the data (shown in black) uses a *non-parametric* estimation (kernel density estimator). We tested three types of distributions, ρ, to describe the data-sets and give a likelihood for each case. The best fit was then determined by the Akaike weight (AW). **I)** Co-implantation of highly malignant GL26-citrine cells (green) and human MSP-12 glioma stem cells (ratio 1:30), and MSP-12 cells alone (control – left image). Immunohistochemistry of human nuclei (black) denote MSP-12 cells. Arrows show the distribution of MSP-12 cells within the brain or the tumor. Scale bar: 100 μm. **J)** Quantification of the distance of MSP-12 from the site of implantation. n=3 for control and n=5 for co-implantation (MSP-12+GL26). Error bars ± SEM; t-test, *p<0.05. **K)** Immunofluorescence images of human-nestin (red) labeling MSP-12 cells, and GL26-citrine cells. Note that MSP-12 cells have a multipolar morphological structure when alone, but a bipolar, elongated structure when aligned to GL26-citrine cells. Scale bar: 47.62 μm. **L)** Angle histogram plots quantify the alignment of MSP-12 within oncostreams, and the random alignment of MSP-12 cells when implanted alone (with dashed overlays of the other condition's alignment).

Fig. 8. Collective invasion of COL1A1 enriched oncostreams contributes to malignant glioma behavior. A) Schematic representation of the experimental setup and location of imaging and quantification of tumor borders using td/mtTomato mice (Movie #5). **B)** Representative time-lapse scanning confocal image of glioma cells at the tumor border. This image was taken from border movie #1 and shows the subdivision into different zones. **C)** Preferred direction of cells within different zones superimposed onto a representative time lapse-image. **D)** Histogram of speed distribution and mean speed (*μm/h*) of Zones A, B, C and D. **E)** Angle Velocity distribution analysis (ϑ) performed by zones. Plot shows overall direction and magnitude of cell movement. **F)** Likelihood analysis of the dynamic patterns at the tumor border. Graph of density estimation ρ *flock* (red), ρ *stream* (yellow) and ρ *swarm* (blue). The estimation of the black line (data) uses a *non-parametric* estimation. AW: 0 or AW:1. **G)** Immunofluorescence analysis of GFP expression in GEMM of glioma controls (NPA), and NPAshCOL1A1. Representative confocal images of the tumor borders. GFP expression is shown in green (Alexa 488) and nuclei in blue (DAPI). Dotted lines show tumor borders. Stars show tumor cell invasion patterns. Notice the absence of collective invasion patterns in NPAshCOL1A1. Scale bar: 50 μm. **H**) The analysis of tumor borders was determined using the Allen-Cahn equation. Images were split into two values (−1 and +1) representing the inside and outside of the tumor to analyze the sinuosity of the borders. Illustration of the sinuosity of a curve: it is defined as the ratio between the length of the curve *L* and the distance between the two extreme points. The sinuosity is close to 1 for a straight line. **I)** Sinuosity of the border for all experiments. 4-10 images of each tumor border were obtained. NPA: n=6 and NPAshCOL1A1: n=5 tumors for each experimental condition were used for the analysis. We detected a decrease of the sinuosity in COL1A1 knockdown tumors. t-test unequal variance, *p=0.0297.

Fig. 9. Intravital two-photon imaging reveals the collective patterns of glioma oncostream dynamics *in vivo*.

A) Schematic representation of the cortical site of glioma cell implantation for intravital two-photon imaging. The inset shows a coronal section of the brain illustrating that glioma cells were implanted at 0.8 mm depth, with the glass cranial window (CW) positioned on top of the implantation site. **B)** Representative photograph of the head of an animal implanted with a cranial window showing the metallic head-bar (HB) positioned on the skull posterior to the cranial window. It affixed using dental cement to stabilize the imaging plane and minimize motion artifacts during time-lapse imaging. a: anterior; p: posterior. **C)** A high-resolution 3D z-stack spanning up to 300 µm depth (starting at the brain's surface) was acquired on the multiphoton microscope, imported into the Imaris viewer and used to reconstruct this 3D image. 300 x-y frames from the brain's surface were taken at a depth increment of 1 µm (voxel size=1) at a resolution of 1024x1024 pixels. XYZ axes of the 3D image are shown in white (596x596x300 µm), and the yellow line shows the exact imaging plane for time-lapse data acquisition *in vivo* (at 120 µm depth). Red fluorescent protein: normal brain parenchyma. Green fluorescent protein: tumor cells. **D)** This panel represents the X-Y and the Y-Z plane of the reconstructed 3D image (shown in **C**) using the Orthoslicer 3D function of the Imaris viewer software to illustrate the depth of the imaging plane. The X-Y plane shown at 120 µm depth illustrates the actual imaging position for movie #14 shown in **E**. **E)** Single representative time-lapse two-photon image of glioma cells within the tumor core *in vivo* (Movie #14) and imaged at a depth of 120 µm, showing Zones A, B, and C. **F)** Individual cell trajectories of the *in vivo* time-lapse experiment. G) Speed distribution and mean speed (µm/hr) for Zones A, B, and C, as indicated in **(E)**. **H)** Angle Velocity distribution for each zone's in the *in vivo* time-lapse movie #14. The Angle Velocity of each cell is denoted θ. The plot shows the proportion of cells moving in the angle direction θ for each zone. **I)** Likelihood analysis of the dynamic patterns determined for each zone of Movie #14 obtained by intravital imaging. The frequency distribution *ρ flock* (red), *ρ stream* (yellow) and *ρ swarm* (blue) are shown. The estimation of the black line (data) uses a *non-parametric* assessment (kernel density estimator) to determine the structure of each zone. AW: 0 or AW:1.

Fig. 10. Oncostreams are COL1A1-rich multicellular dynamic mesenchymal structures that regulate glioma invasion and malignancy. Summary representation of mesenchymal dynamic fascicles (oncostreams) present in high grade gliomas. Our study reveals that oncostreams display directional collective motility patterns including streams and flocks. Non-directional collective motion (swarms) are represented by round cells that move do not have a preferred direction of motion. Directional dynamic patterns function as tumoral highways to facilitate the intra-tumoral spread of cells and participate in local invasion of normal brain. Oncostreams are areas of mesenchymal transformation defined by a molecular signature enriched in COL1A1. COL1A1 knockdown disrupts oncostream organization, decreases intratumoral heterogeneity and significantly increases animal survival. Our study reveals that oncostreams are anatomically and molecularly distinctive, are areas of mesenchymal transformation organized through interactions with the COL1A1 matrix, move by collective motion, and regulate glioma growth and invasion.

SUPPLEMENTARY INFORMATION

Supplementary Material and Methods

Supplementary Figures

Supplementary Tables

Supplementary Videos

Fig. 1

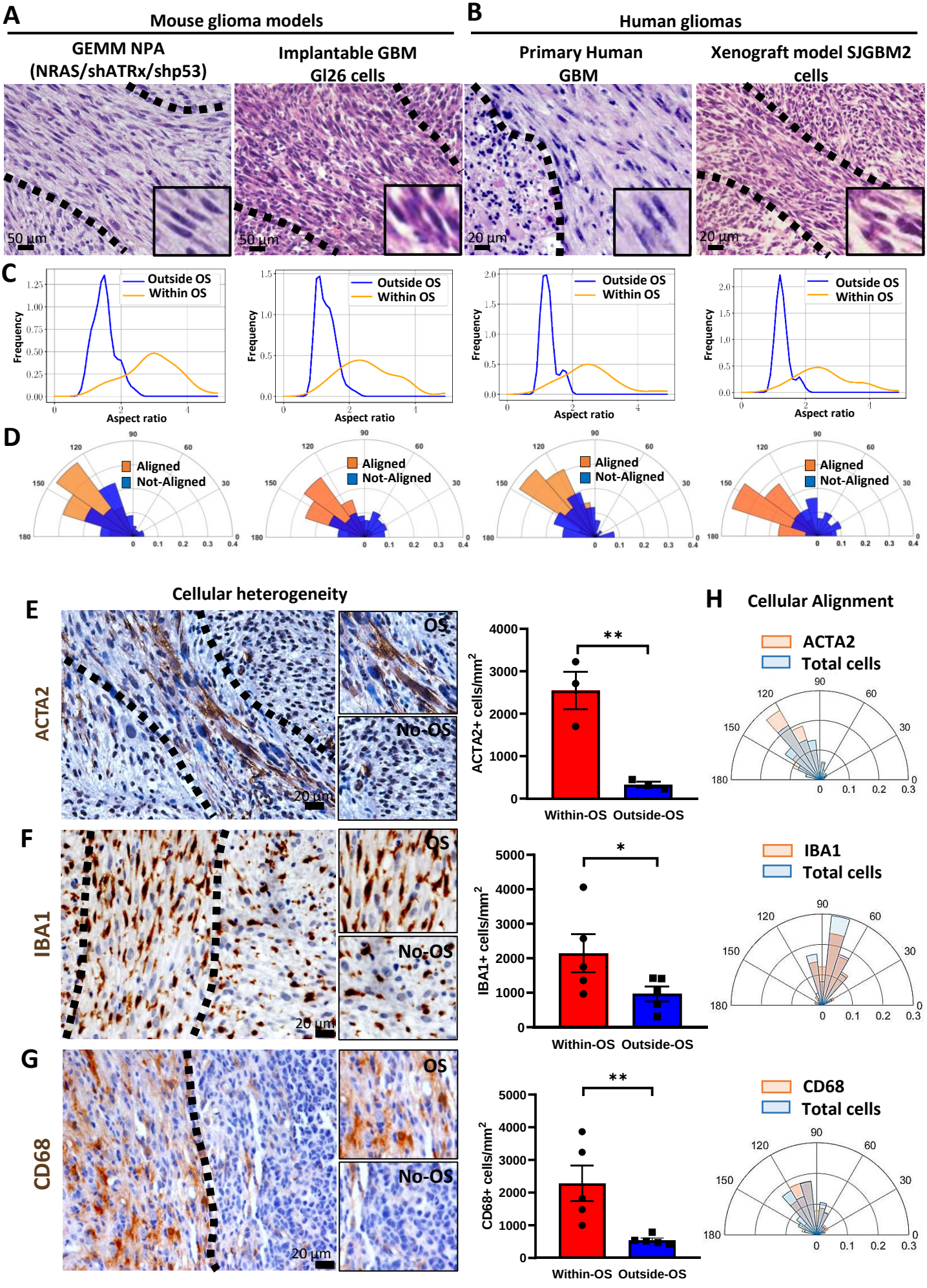


Fig. 1. Oncostreams are multicellular fascicles present in mouse and human gliomas

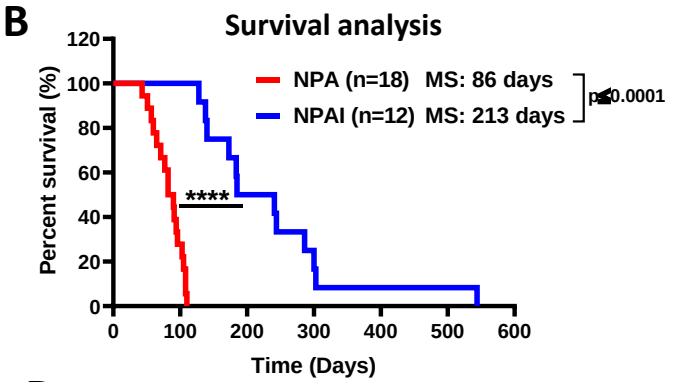
A) Representative 5 µm H&E microtome sections from gliomas showing that fascicles of spindle-like glioma cells (oncostreams, outlined by the dotted line) are present in a GEMM of gliomas NPA (NRAS/shATRx/shp53) and the GL26 intracranial implantable model of glioma. Scale bars: 50 µm. **B)** Representative H&E microtome sections of human glioma and human xenografts showing the presence of oncostreams. Scale bar: 20 µm. **C-D)** Histograms showing the cellular shape analysis (aspect ratio) (**C**) and angle orientation (alignment) for the corresponding images (**D**) show areas of oncostreams (OS) formed by elongated and aligned cells and areas with no oncostreams (No-OS) as rounded and not-aligned cells. **E-G)** Immunostaining shows that tumor cells, mesenchymal cells (ACTA2+), microglia/macrophages (IBA1+ and CD68+), are aligned within, the main orientation axis of oncostreams. Bar graphs show the quantification of ACTA+, n=3 (**E**), IBA1+, n=5 (**F**) and CD68+, n=5 (**G**) cells within oncostreams areas in NPA tumors. 6-13 areas of oncostreams per tumor section per animal were imaged. Scale bar: 20 µm. Error bars represent ± SEM; unpaired t-test analysis, *$p<0.05$, **$p<0.001$. **H)** Angle orientation shows the alignment of ACTA+, IBA1+ and CD68+ cells within oncostreams for the corresponding images.

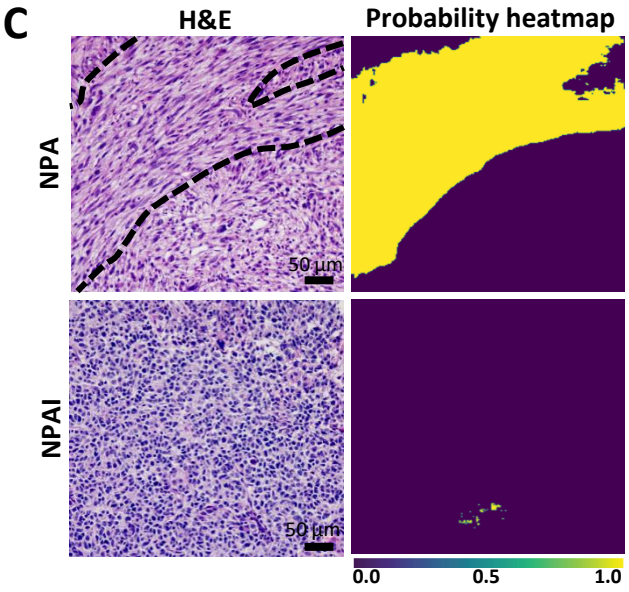
Fig. 2

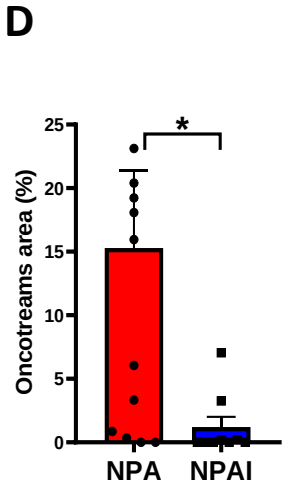
A

GEMM of gliomas

Tumor	Genetic modifications
NPA	NRAS-GV12/shP53-GFP/shATRX-GFP
NPAI	NRAS-GV12/shP53-GFP/shATRX-GFP/IDH1-R132H







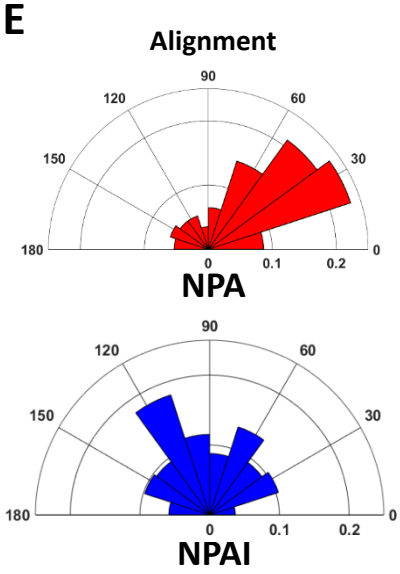


Fig. 2. Oncostreams density positively correlates with tumor aggressiveness in GEMM of gliomas. **A)** Genetic makeup of NPA and NPAI tumors. **B)** Kaplan–Meier survival curves of NPA and NPAI mouse gliomas show that animals bearing IDH1-R132H mutant tumors (NPAI) have prolonged median survival (MS): **NPA** (MS: 86 days; n: 18) versus **NPAI** (MS: 213 days; n:12). Log-rank (Mantel-Cox) test; ****p<0.0001. **C-D)** Deep learning method for oncostream detection in H&E stained mouse glioma sections: **C)** Representative images of oncostreams manually segmented on H&E stained sections of NPA gliomas and NPAI tumors. The output of our trained model for each image is shown below (probability heat maps), for tissues containing oncostreams (NPA), and without oncostreams (NPAI), scale bar = 50 µm. **D)** 10-14 random fields per tumor section per animal were imaged, n=9 NPA and n=12 NPAI, and quantified using deep learning analysis. Error bars represent ± SEM; unpaired t-test analysis, *p<0.05. **E)** Angle histogram plots show aligned cells in NPA tumors vs non-aligned cells in NPAI tumors for the representative images showed in figure.

Fig. 3

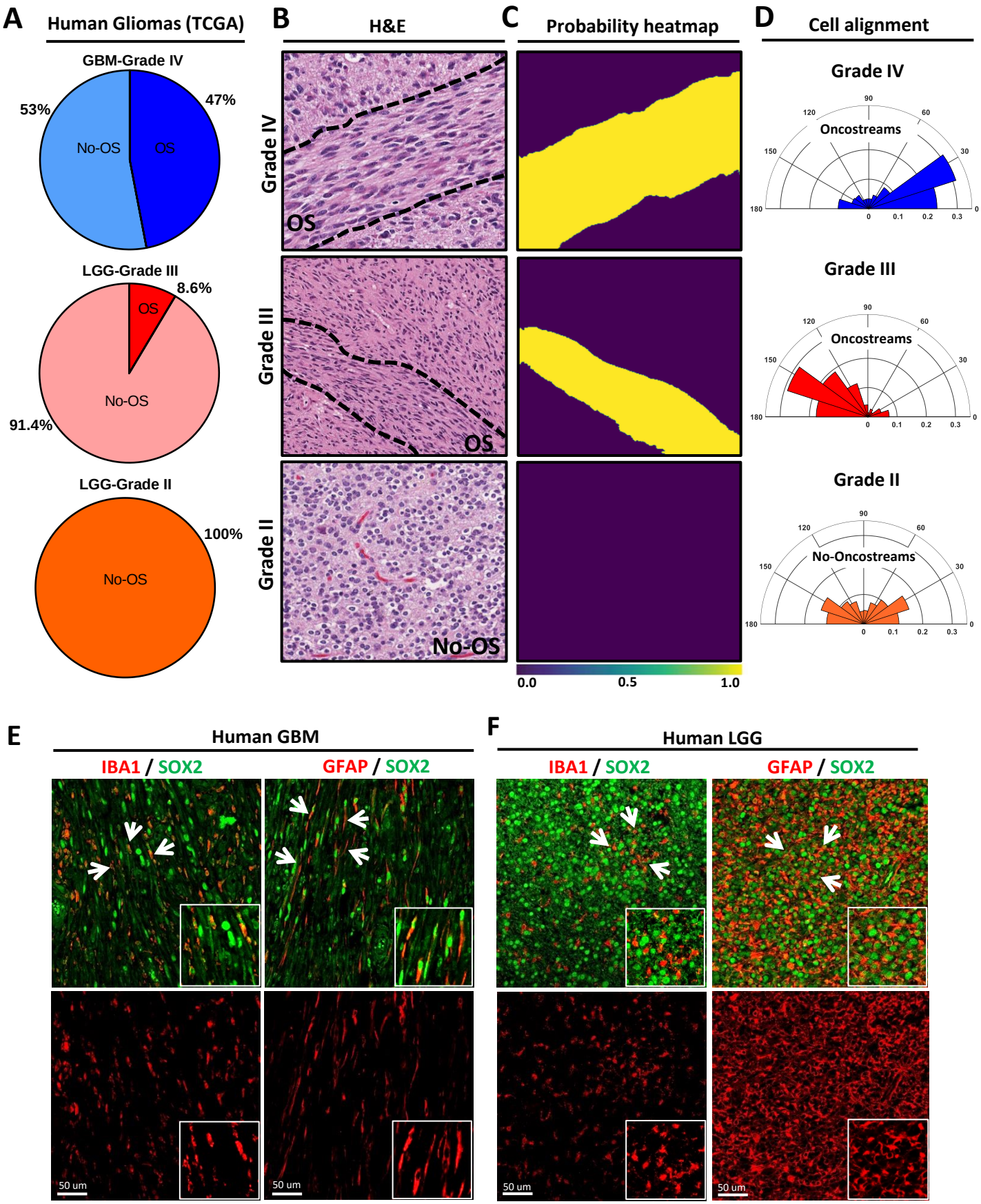


Fig. 3. The density of oncostreams positively correlates with tumor aggressiveness in human gliomas. A) TCGA tumors were analyzed from different grade: GBM-Grade IV (100 tumors). LGG-Grade III (70 tumors) and LGG-Grade II (50 tumors). Pie charts show percentage of tumors displaying oncostreams in relation to tumor grade. Oncostreams are present in 47% of GBM grade IV tumors, 8.6 % of LGG grade III, and are absent from LGG grade II. **B)** Manual identification of oncostreams in H&E images are shown for human gliomas with WHO grades IV, III, II from TCGA. **C)** Deep learning analysis for human gliomas. Our algorithm was able to detect oncostreams in grade IV and III gliomas but not in grade II gliomas. **D)** Angle histogram plots show the alignment of cells in H&E histology sections of Grade IV and Grade III gliomas' oncostreams and random alignment in grade II glioma sections lacking oncostreams. Angle histogram correspond to the representative images. **E-F)** Immuno-fluorescence staining of SOX2+ tumor cells (green), glial fibrillary acidic protein (GFAP+) cells (red), and microglia/macrophage (IBA1+) cells (red) in high-grade human glioblastoma (GBM) (WHO Grade IV), IDH-WT **(E)** and in low-grade glioma (LGG) (WHO Grade III), IDH-mutant **(F)**, showing oncostreams heterogeneity and cellular alignment of these cells in human high-grade gliomas but not in low grade gliomas (arrows). Scale bars: 50 µm.

Fig. 4

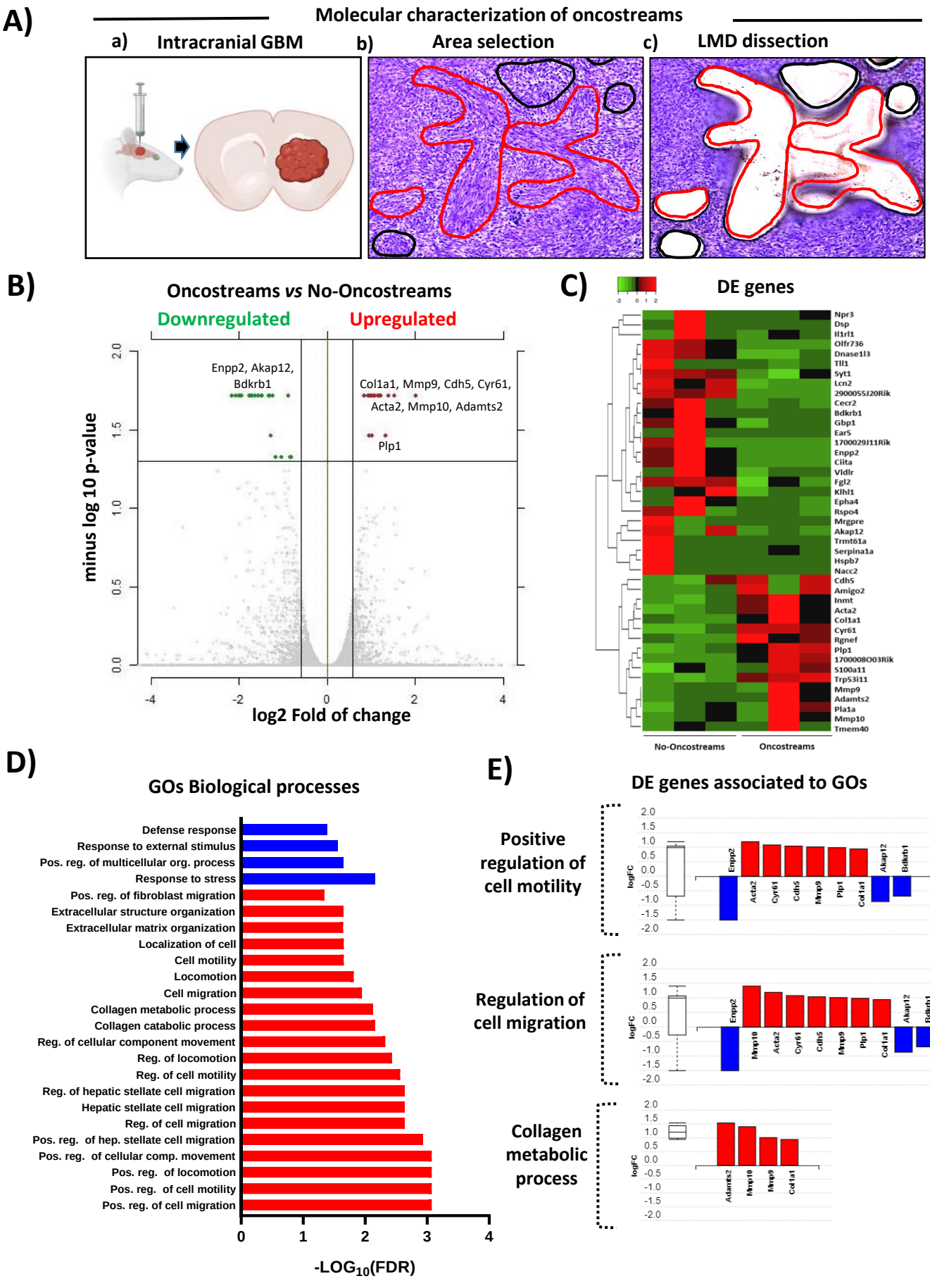


Fig. 4. Oncostreams are defined by a unique gene expression signature related to mesenchymal transformation and migration. A) (a) Schematic representation of spatial transcriptomic analysis of glioma oncostreams using Laser Capture Microdissection (LCM). Glioma tumors were generated by intracranial implantation of NPA tumor cells in C57BL6 mice. (b-c) Oncostream areas (red outline) were identified and dissected from surrounding glioma tissue (black outline) in mouse glioma samples using a LCM microscope. **B)** A volcano plot displays differentially expressed (DE) genes from oncostream *vs* no-oncostream areas. DE genes were selected based on a fold change of ≥ 1.5 and a q-value (false discovery rate (FDR) corrected p-value) of ≤ 0.05. Upregulated genes (red dots) and downregulated genes (green dots) are shown. Relevant genes related to mesenchymal migration are labeled on the graph. **C)** Heat map illustrates gene expression patterns for oncostream vs no-oncostream areas in NPA glioma tumors (n=3 biological replicates/group). Differentially upregulated genes (16) are represented in red and downregulated genes (n=27) are represented in green (q-value ≤ 0.05 and fold change ≥ ± 1.5). **D)** Functional enrichment analysis of overrepresented GO terms (biological processes) obtained when comparing oncostream vs no-oncostream DE genes. p-value corrected for multiple comparisons using the FDR method. Cutoff FDR<0.05. Blue: Downregulated GOs. Red: upregulated GOs. **E)** Bar graphs show DE genes annotated to the most relevant enriched GOs biological process: “Positive regulation of cell motility”, “Regulation of cell migration” and “Collagen metabolic process.”

Fig. 5

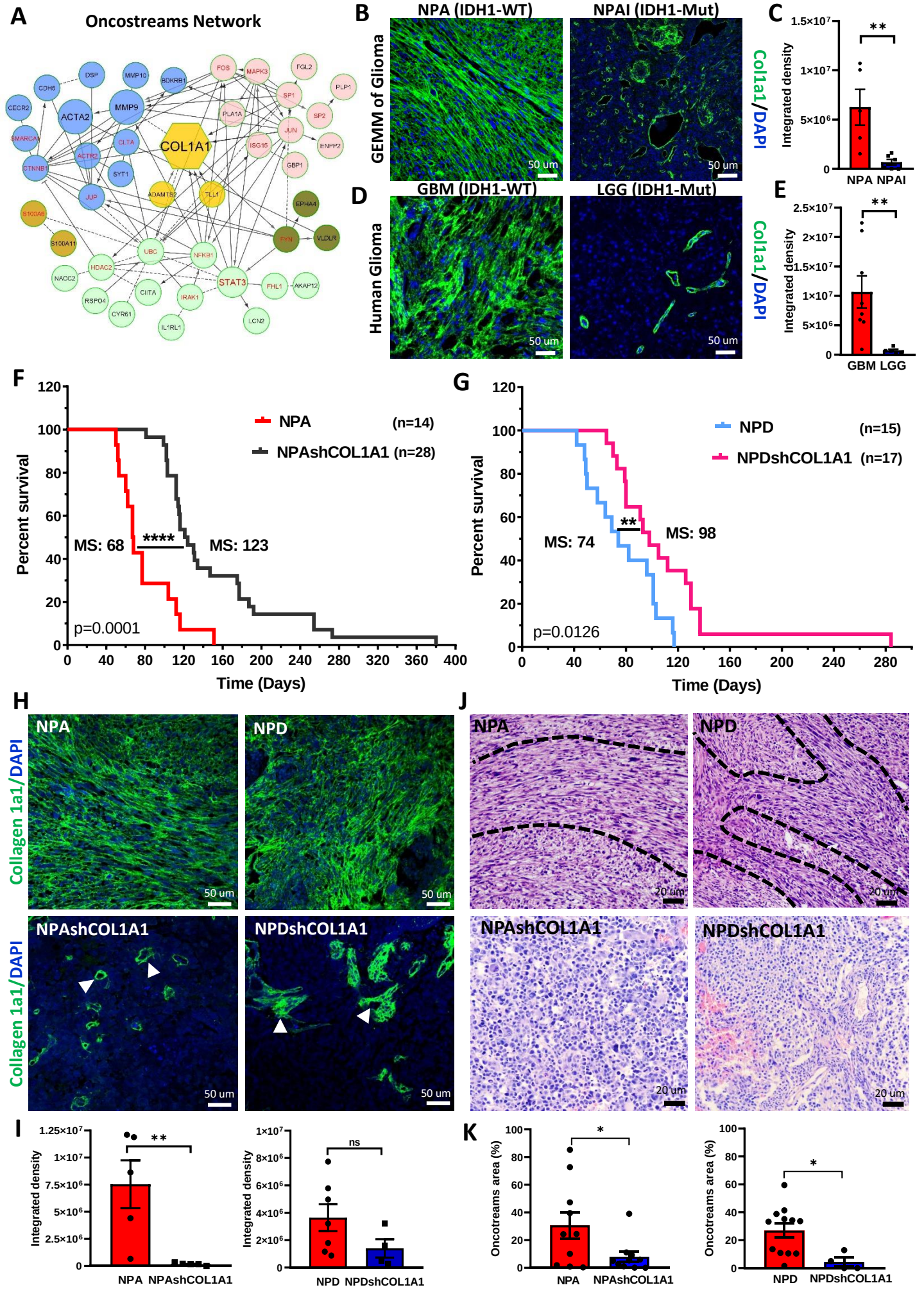


Fig. 5. COL1A1 is a central hub in oncostream organization and glioma malignancy. A) Network analysis of the DE genes comparing oncostreams versus no-oncostreams DE genes. Genes with a higher degree of connectivity are highlighted with larger nodes. Clusters of nodes with the same color illustrate modules of highly interacting of genes in the network. **B)** Immunofluorescence analysis of COL1A1 expression in GEMM of glioma tissues comparing NPA (IDH1-WT) vs NPAI (IDH1mut). Representative confocal images display COL1A1 expression in green (Alexa 488) and nuclei in blue (DAPI). Scale bar: 50 µm. **C)** Bar graphs represent COL1A1 quantification in terms of fluorescence integrated density. NPA n=5 and NPAI n=6 animals for each experimental condition were used for the analysis. Ten fields of each tumor section were selected at random. Error bars represent ±SEM. t-test, **p<0.01. **D)** Immunofluorescence analysis of COL1A1 expression in human GBM and LGG tumors. COL1A1 expression in green (Alexa 488) and nuclei in blue (DAPI). Scale bar: 50 µm. **E)** Bar graphs represent COL1A1 quantification as fluorescence integrated density. 5 (LGG) and 8 (GBM) tumor samples were used for the analysis. Ten fields of each tumor section were selected at random. Error bars represent ±SEM. t-test, **p<0.01. **F-G)** GEMM of glioma with COL1A1 inhibition. **F)** Kaplan–Meier survival curve comparing NPA (MS: 68 days; n: 14) vs NPAshCOL1A1 (MS: 123 days; n: 28) **G)** Kaplan–Meier survival curve comparing NPD (MS: 74 days; n=15) versus NPDshCOL1A1 (MS: 98 days; n=17). Log-rank (Mantel-Cox) test. **** p<0.0001, **p<0.0126. **H)** Immunofluorescence analysis of COL1A1 expression in GEMM of glioma controls (NPA and NPD) and Col1A1 downregulation (NPAshCOL1A1 and NPDshCOL1A1). Representative confocal images of COL1A1 expression in green (Alexa 488) and nuclei in blue (DAPI). Arrows indicate COL1A1 enriched perivascular cells. Scale bar: 50 µm. **I)** Bar graphs represent COL1A1 quantification in terms of fluorescence integrated density. 5-7 tumor samples for each experimental condition were used for the analysis. Ten fields of each tumor section were selected at random. Error bars represent ±SEM. t-test, **p<0.01. **J)** Representative images of the histopathological identification of oncostreams in H&E tissue sections comparing the COL1A1 knockdown tumors with their respective controls. Scale bars: 50 µm. **K)** Quantitative analysis of oncostream areas using deep learning analysis. 4-12 tumor samples for each experimental condition were used for the analysis. Error bars represent ± SEM; unpaired t-test analysis, *p<0.05.

Fig. 6

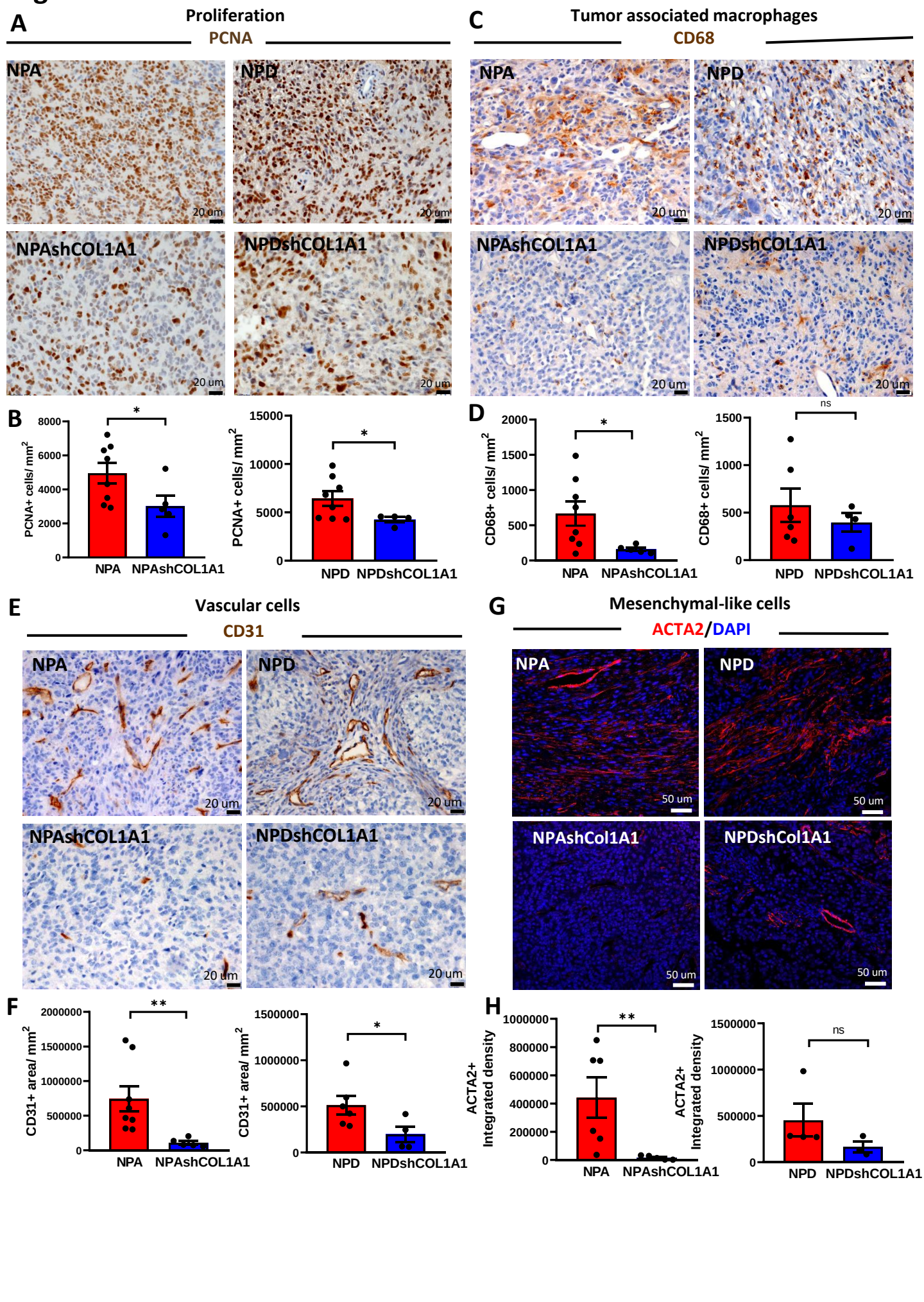


Fig. 6. Knockdown of COL1A1 within glioma cells modifies the tumor microenvironment
Immunohistochemical analysis (**A**, **C** and **E**) of GEMM of glioma controls (**NPA** and **NPD**) and COL1A1 downregulation (**NPAshCOL1A1** and **NPDshCOL1A1**). **A)** Representative images of PCNA expression. Scale bar: 20 µm. **B)** Bar graphs represent the quantification of PCNA+ cells numbers (cells/mm^2) using QuPath positive cell detection. Error bars represent ±SEM, (NPA: n=8, NPAshCOL1A1: n=5, NPD: n=8, NPDshCOL1A1: n=4), t-test, *p<0.05. **C)** Representative images of CD68 expression. Scale bar: 20 µm**. D)** Bar graphs represent CD68+ cell quantification (cells/mm^2) using QuPath positive cell detection. Error bars represent ±SEM, (NPA: n=8, NPAshCOL1A1: n=5, NPD: n=6, NPDshCOL1A1: n=4), t-test, *p<0.05, ns: no significant. **E)** Representative images of CD31 expression. Scale bar: 20 µm**. F)** Bar graphs represent CD31+ cells quantification (cells/mm^2) using QuPath positive cells detection. Error bars represent ±SEM, (NPA: n=8, NPAshCOL1A1: n=5, NPD: n=6, NPDshCOL1A1: n=4), t-test, **p<0.01, *p<0.05. **G)** Immunofluorescence analysis of GEMM of glioma controls (**NPA** and **NPD**) and COL1A1 downregulation (**NPAshCOL1A1** and **NPDshCOL1A1**). Representative images of ACTA2 expression in red (Alexa 555) and nuclei in blue (DAPI). Scale bar: 50 µm. **H)** Bar graphs represent ACTA2 quantification in terms of fluorescence integrated density. Error bars represent ±SEM, (NPA: n=6, NPAshCOL1A1: n=5, NPD: n=4, NPDshCOL1A1: n=3), t-test, **p<0.01, ns: no significant.

Fig. 7

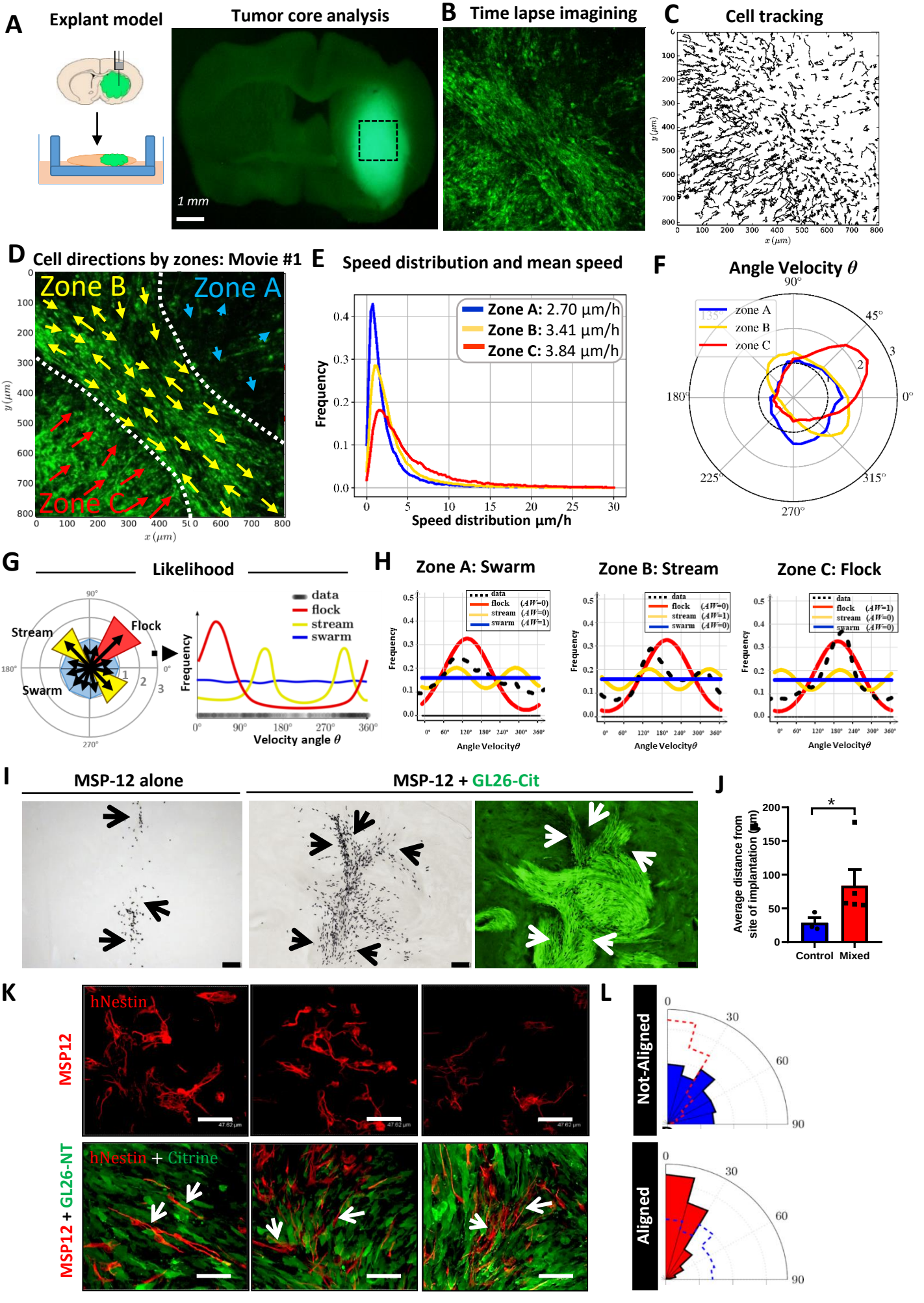


Fig. 7. Collective dynamics of oncostreams increase cell spreading within the tumor core. A) Experimental setup: NPA-GFP glioma cells were intracranially implanted in C57BL6 mice. Explant slice cultures of growing tumors were used for confocal time-lapse imaging of the tumor core. **B)** Single representative time-lapse confocal image of glioma cells within the tumor core (Movie #1). **C)** Tracking analysis of individual cell paths performed using the Track-Mate plugin from Image-J. **D)** Preferred directions of cells within three zones (A-C) superimposed onto a representative time lapse-image. **E)** Speed distribution and mean speed ($\mu m/hr$) in Zones A, B and C. **F)** Distribution of angle velocity for each zone. The Angle Velocity of each cell is denoted ϑ. The plot shows the proportion of cells moving in angle direction ϑ for each zone. **G-H)** Classification of collective motion patterns: *stream*, *flock* or *swarm*. The distribution is *uni-modal* for a flock (only one peak) and *bi-model* for a stream (two peaks = 2 preferred angle velocity). For a *swarm*, the distribution is *flat* (no preferred angle velocity). In **(G)** Angle Velocity was transformed to a histogram; these data were then used to calculate the likelihood that a particular distribution of velocity angles corresponds to either a *stream*, *flock*, or *swarm*. The results are given in **(H)** for each zone. The frequency distribution of the data (shown in black) uses a *non-parametric* estimation (kernel density estimator). We tested three types of distributions, ρ, to describe the data-sets and give a likelihood for each case. The best fit was then determined by the Akaike weight (AW). **I)** Co-implantation of highly malignant GL26-citrine cells (green) and human MSP-12 glioma stem cells (ratio 1:30), and MSP-12 cells alone (control – left image). Immunohistochemistry of human nuclei (black) denote MSP-12 cells. Arrows show the distribution of MSP-12 cells within the brain or the tumor. Scale bar: 100 µm. **J)** Quantification of the distance of MSP-12 from the site of implantation. n=3 for control and n=5 for co-implantation (MSP-12+GL26). Error bars ± SEM; t-test, *p<0.05. **K)** Immunofluorescence images of human-nestin (red) labeling MSP-12 cells, and GL26-citrine cells. Note that MSP-12 cells have a multipolar morphological structure when alone, but a bipolar, elongated structure when aligned to GL26-citrine cells. Scale bar: 47.62 µm. **L)** Angle histogram plots quantify the alignment of MSP-12 within oncostreams, and the random alignment of MSP-12 cells when implanted alone (with dashed overlays of the other condition's alignment).

Fig. 8

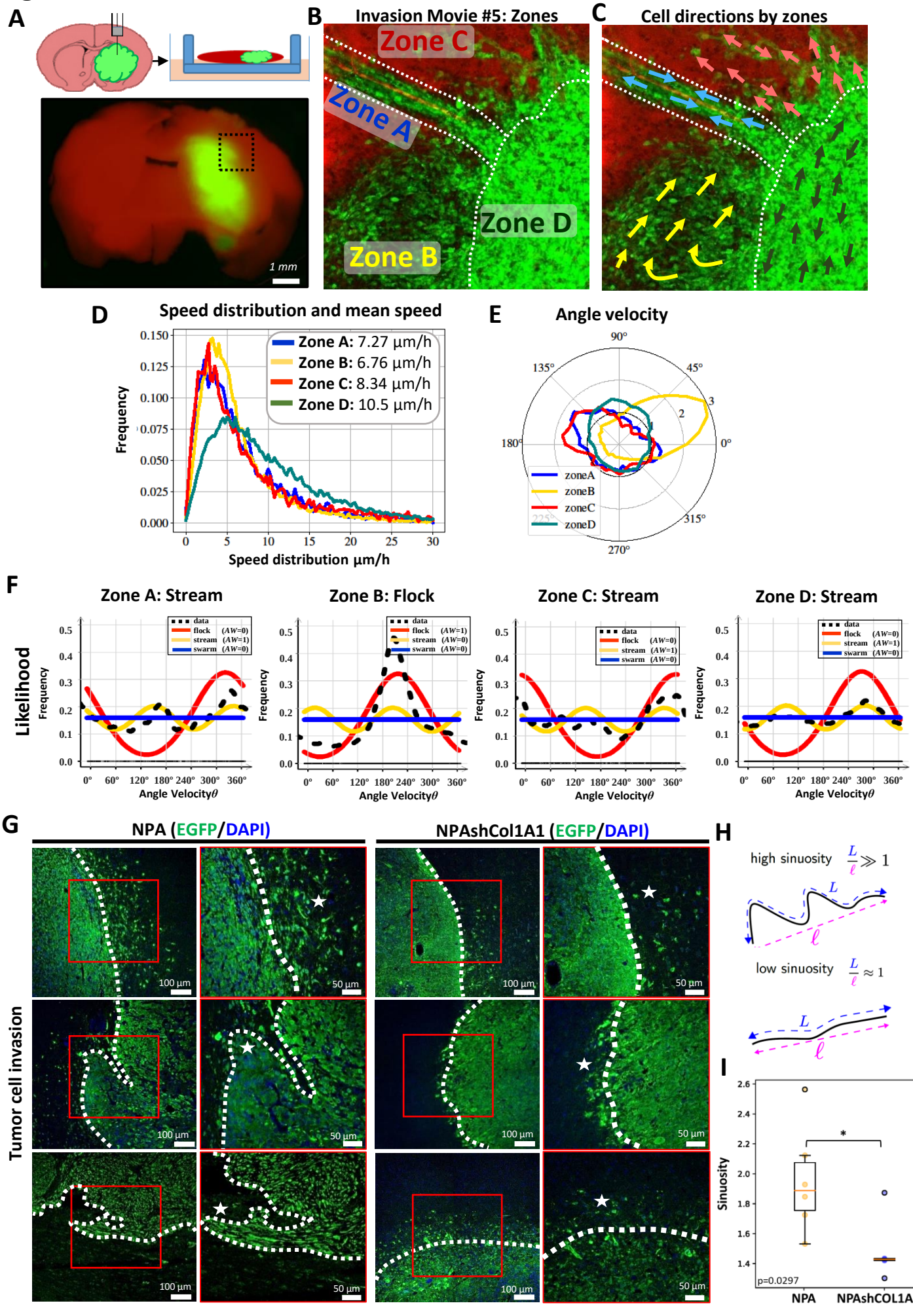
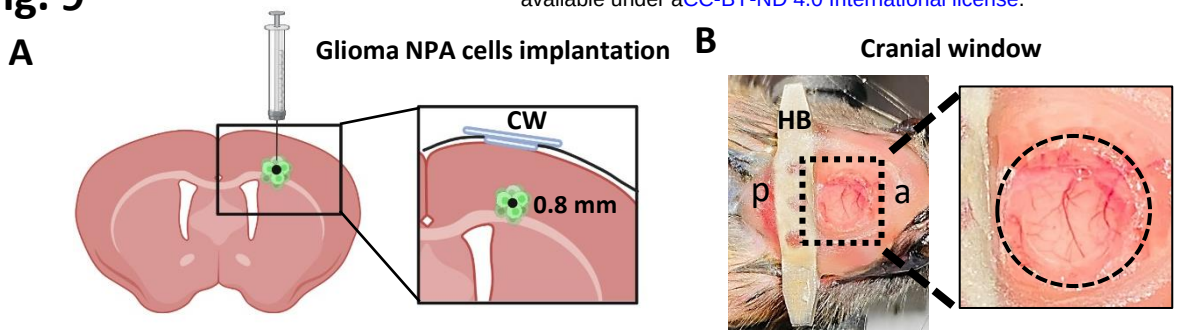
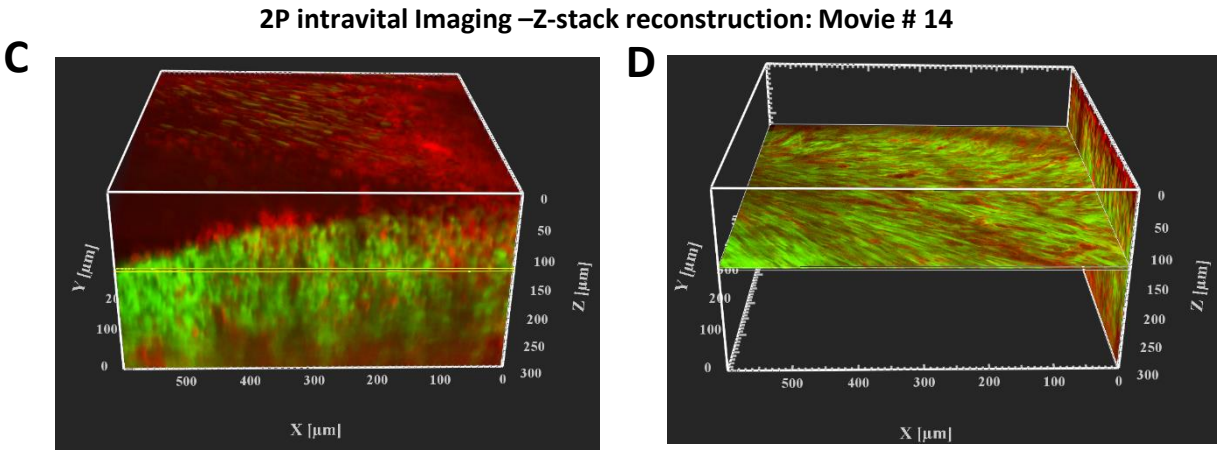
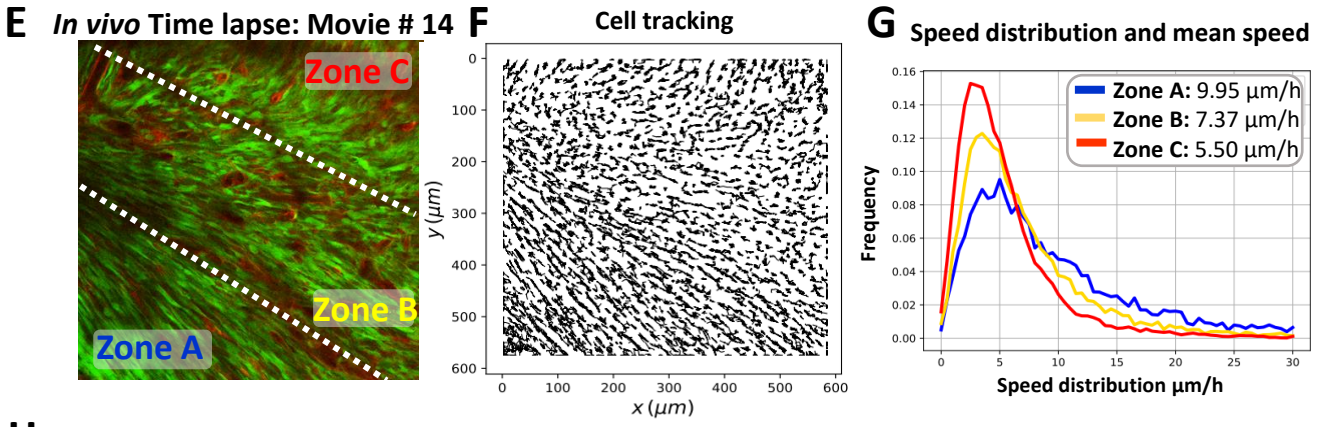


Fig. 8. Collective invasion of COL1A1 enriched oncostreams contributes to malignant glioma behavior. A) Schematic representation of the experimental setup and location of imaging and quantification of tumor borders using td/mtTomato mice (Movie #5). **B)** Representative time-lapse scanning confocal image of glioma cells at the tumor border. This image was taken from border movie #1 and shows the subdivision into different zones. **C)** Preferred direction of cells within different zones superimposed onto a representative time lapse-image. **D)** Histogram of speed distribution and mean speed ($\mu m/h$) of Zones A, B, C and D. **E)** Angle Velocity distribution analysis (ϑ) performed by zones. Plot shows overall direction and magnitude of cell movement. **F)** Likelihood analysis of the dynamic patterns at the tumor border. Graph of density estimation *ρ flock* (red), *ρ stream* (yellow) and *ρ swarm* (blue). The estimation of the black line (data) uses a *non-parametric* estimation. AW: 0 or AW:1. **G)** Immunofluorescence analysis of GFP expression in GEMM of glioma controls (NPA), and NPAshCOL1A1. Representative confocal images of the tumor borders. GFP expression is shown in green (Alexa 488) and nuclei in blue (DAPI). Dotted lines show tumor borders. Stars show tumor cell invasion patterns. Notice the absence of collective invasion patterns in NPAshCOL1A1. Scale bar: 50 µm. **H**) The analysis of tumor borders was determined using the Allen-Cahn equation. Images were split into two values (−1 and +1) representing the inside and outside of the tumor to analyze the sinuosity of the borders. Illustration of the sinuosity of a curve: it is defined as the ratio between the length of the curve *L* and the distance between the two extreme points. The sinuosity is close to 1 for a straight line. **I)** Sinuosity of the border for all experiments. 4-10 images of each tumor border were obtained. NPA: n=6 and NPAshCOL1A1: n=5 tumors for each experimental condition were used for the analysis. We detected a decrease of the sinuosity in COL1A1 knockdown tumors. t-test unequal variance, *p=0.0297.

Fig. 9







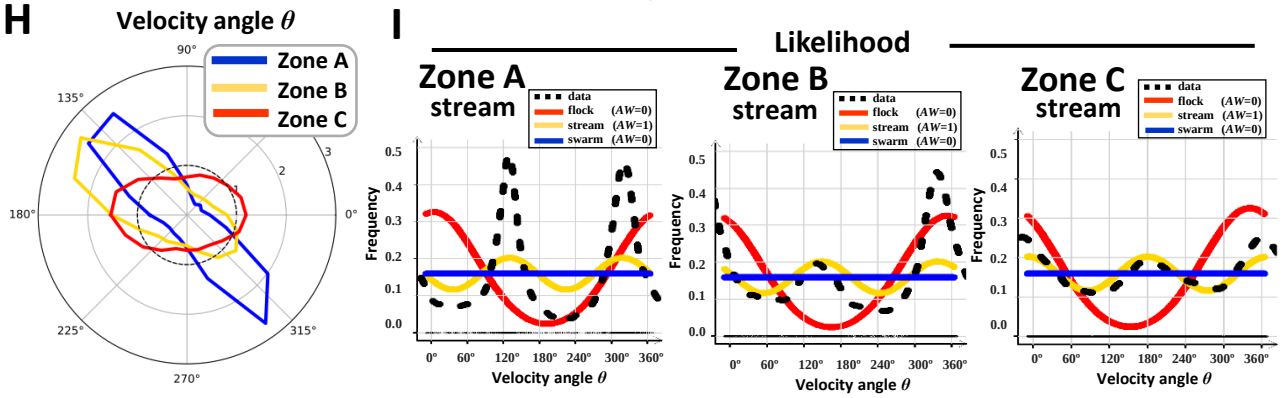


Fig. 9. Intravital two-photon imaging reveals the collective patterns of glioma oncostream dynamics *in vivo*.

A) Schematic representation of the cortical site of glioma cell implantation for intravital two-photon imaging. The inset shows a coronal section of the brain illustrating that glioma cells were implanted at 0.8 mm depth, with the glass cranial window (CW) positioned on top of the implantation site. **B)** Representative photograph of the head of an animal implanted with a cranial window showing the metallic head-bar (HB) positioned on the skull posterior to the cranial window. It affixed using dental cement to stabilize the imaging plane and minimize motion artifacts during time-lapse imaging. a: anterior; p: posterior. **C)** A high-resolution 3D z-stack spanning up to 300 µm depth (starting at the brain's surface) was acquired on the multiphoton microscope, imported into the Imaris viewer and used to reconstruct this 3D image. 300 x-y frames from the brain's surface were taken at a depth increment of 1 µm (voxel size=1) at a resolution of 1024x1024 pixels. XYZ axes of the 3D image are shown in white (596x596x300 µm), and the yellow line shows the exact imaging plane for time-lapse data acquisition *in vivo* (at 120 µm depth). Red fluorescent protein: normal brain parenchyma. Green fluorescent protein: tumor cells. **D)** This panel represents the X-Y and the Y-Z plane of the reconstructed 3D image (shown in **C**) using the Orthoslicer 3D function of the Imaris viewer software to illustrate the depth of the imaging plane. The X-Y plane shown at 120 µm depth illustrates the actual imaging position for movie #14 shown in **E**. **E)** Single representative time-lapse two-photon image of glioma cells within the tumor core *in vivo* (Movie #14) and imaged at a depth of 120 µm, showing Zones A, B, and C. **F)** Individual cell trajectories of the *in vivo* time-lapse experiment. G) Speed distribution and mean speed (µm/hr) for Zones A, B, and C, as indicated in **(E)**. **H)** Angle Velocity distribution for each zone's in the *in vivo* time-lapse movie #14. The Angle Velocity of each cell is denoted θ. The plot shows the proportion of cells moving in the angle direction θ for each zone. **I)** Likelihood analysis of the dynamic patterns determined for each zone of Movie #14 obtained by intravital imaging. The frequency distribution *ρ flock* (red), *ρ stream* (yellow) and *ρ swarm* (blue) are shown. The estimation of the black line (data) uses a *non-parametric* assessment (kernel density estimator) to determine the structure of each zone. AW: 0 or AW:1.

Fig. 10

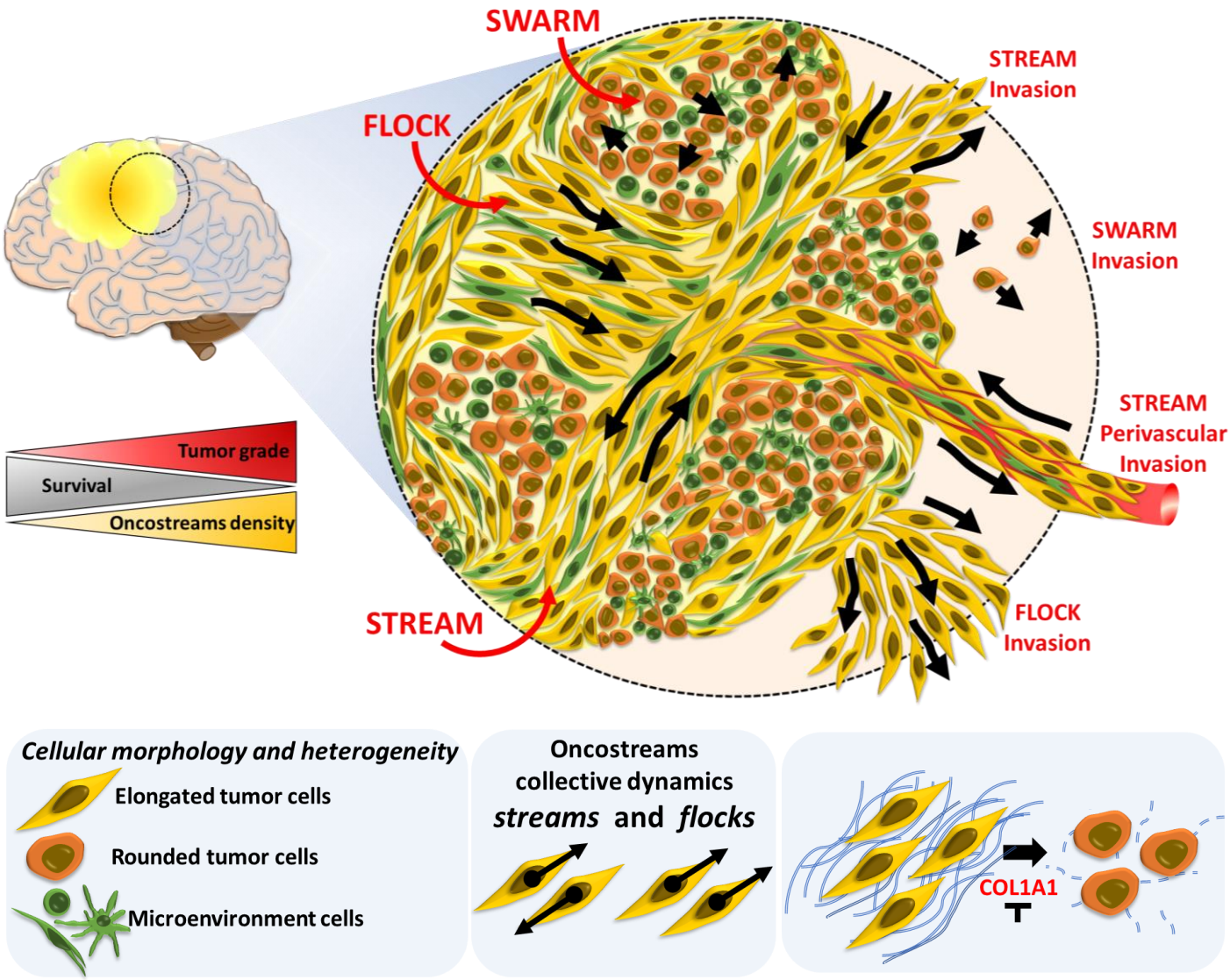


Fig. 10. Oncostreams are COL1A1-rich multicellular dynamic mesenchymal structures that regulate glioma invasion and malignancy. Summary representation of mesenchymal dynamic fascicles (oncostreams) present in high grade gliomas. Our study reveals that oncostreams display directional collective motility patterns including streams and flocks. Non-directional collective motion (swarms) are represented by round cells that move do not have a preferred direction of motion. Directional dynamic patterns function as tumoral highways to facilitate the intra-tumoral spread of cells and participate in local invasion of normal brain. Oncostreams are areas of mesenchymal transformation defined by a molecular signature enriched in COL1A1. COL1A1 knockdown disrupts oncostream organization, decreases intratumoral heterogeneity and significantly increases animal survival. Our study reveals that oncostreams are anatomically and molecularly distinctive, are areas of mesenchymal transformation organized through interactions with the COL1A1 matrix, move by collective motion, and regulate glioma growth and invasion.