Supplementary Methods

Measurements of cell deformation

Definition of the mean nuclear deformation during invasion. Given the micrometer-scale resolution of the images, we consider each nucleus $k \in \{1, \dots, N\}$ under study as a continuous media in a compact domain $\Omega_{k,t}$ that deforms over time, $t \in [0, T]$. After tracking and segmenting the nucleus, we extract the displacement fields between successive time frames (all in 3D, see below) [59]. We then use the resulting data to measure the deformation of each nucleus at each point $\boldsymbol{x}$ of their domain and each time t of the invasion. To this end, we compute the second order strain tensor, defined as the symmetric part of the Jacobian of the displacement field $\boldsymbol{u}$:

$$\forall (\boldsymbol{x}, t) \in \Omega_{k,t} \times [0, T], \boldsymbol{\epsilon}(\boldsymbol{x}, t) := \frac{1}{2}(\nabla \boldsymbol{u}(\boldsymbol{x}, t) + \nabla \boldsymbol{u}^T(\boldsymbol{x}, t))$$

Other higher order definitions of the strain such as Green's tensor are too sensitive to noise in our setting. The resulting strain tensor can be in turn used to compute strain invariants that characterize the mechanical behavior of the deforming domain quantitatively. In particular, we chose the Frobenius norm of the strain tensor as a measure of local deformation because it accounts for both hydrostatic and shear distortions:

$$\forall (\boldsymbol{x}, t) \in \Omega_{k,t} \times [0, T], d_k(\boldsymbol{x}, t) = \sqrt{\sum_{i,j} \boldsymbol{\epsilon}_{ij}(\boldsymbol{x}, t)^2}$$

At last, we compute the mean value of this locally defined deformation measure over the spatial domain and over the invasion time to characterize the global deformation of one nucleus labeled k:

$$D_k = \int_{t \in [0,T]} \int_{\boldsymbol{x} \in \Omega_{k,t}} \frac{d_k(\boldsymbol{x}, t)}{V_k(t) T},$$

where $V_k(t)$ is the volume of the nucleus at time t.

We now comment on why we use the strain and not the displacement. It is not enough that an object moves for it to deform, for example rigid translations do not induce any deformation yet are part of the displacement field. Rather, it is the relative displacements between adjacent points, whether they draw to one another or spread apart, that characterize deformation. The strain tensor quantifies this notion.

Numerical computation. We segment each deforming nucleus by a combination of HK-means and Active Contours [60]. The sets of resulting voxels define the nuclear domains $\Omega_{k,t}$. We take $N_{voxels} n_x n_y n_z$ to be the volume $V_k(t)$ of the segmented nucleus at time t, where n_x, n_y and n_z are the respective dimensions in micrometers of the voxels and N_{voxels} is the number of voxels belonging to the segmented domain. We use the Horn-Schunck optical flow method to compute at each voxel the displacement field u, a

framework that we proved to be consistent in the estimation of strain in a previous study [54]. It operates under two chief assumptions: brightness constancy states that the image intensities are conserved from one frame to the next; and smoothness asserts that the displacements of each pixel are small enough, typically about one voxel or less, and that their variations are bounded. The former is secured by the use of stable markers. The latter is also fulfilled because GBM nuclei deform slowly and we employ a fast enough confocal microscope. Under these circumstances, the 3D displacements of the nuclei are obtained by solving the following variational problem:

$$\underset{\boldsymbol{u}}{\operatorname{argmin}} \left(\frac{\partial I}{\partial t} + \boldsymbol{u} \cdot \nabla I \right)^2 + \alpha ||\nabla \boldsymbol{u}||^2$$

where I is the image intensity function and $\alpha > 0$ a regularization parameter. This problem is undertaken through an iterative scheme drawn from the corresponding Euler-Lagrange equations. We employ the well-known trick of post processing the derived displacements with a median filter to remove noise before differentiation [61]. The corresponding strain tensor is calculated by convolution with a Sobel filter, and the mean deformation of each nucleus by averaging these computations over the segmentation masks. In this way, each nucleus, be it control or IF-depleted, leader or follower, is thus assigned a deformation value. We compare the resulting values through Student's t-tests and Cohen's coefficient to conclude that the leading IF-depleted nuclei deform significantly more than the control ones (see this and other results in the main text).

Supplementary Figures

Supplementary Figure 1- Identification of neoplastic cells for each individual patient.

A Uniform Manifold Approximation and Projection (UMAP) analysis shows the different cell populations present within each tumor per patient (Patient 1-4 (Darmanis et al., 2017), Patient 5-8 (Yuan et al., 2018)). B. Heatmaps of the average gene expression per chromosome per suspected neoplastic cell normalized by the average gene expression per chromosome of all suspected non-neoplastic cells. Scaling was performed on rows.

Supplementary Figure 2 – Heterogeneity of IF expression in glioblastoma patient samples and correlation to functional gene expression signatures

A Uniform Manifold Approximation and Projection (UMAP) analysis shows the different neoplastic cell populations present within the eight8 different tumours of patients after integration of the data. **B** Glial fibrillary acidic protein (GFAP), Vimentin (VIM), nestin (NES), and synemin (SYNM) are heterogeneously expressed within these cell populations. **C** Hierarchical clustering of all neoplastic cells based on GFAP, VIM, NES, and SYNM expression levels reveals 10 different expression patterns. **D** Pie charts showing the distribution of cells in each IF cell cluster for each individual patient. **E** Intermediate filament score (sum of normalized IF gene expression levels) for each IF cell cluster. **F-I** Expression of migration-related genes (GO ID GO:0030335, Supplementary Table 3) (F), extracellular matrix organization (GO:0030198, Supplementary Table 3) (G), response to mechanical stimulus (GO:0009612, Supplementary Table 3) (H), cell division (GO:0051301, Supplementary Table 3) (I) in the different IF expression pattern-groups show a strong correlation with total intermediate filament expression levels. The color represent the average expression level within the cluster. The size of the circle represents the percentage of cells expressing the gene. **J** Patient legend of Figure 1E-G and (K). **K** Violin plot of the cell division (GO:0051301) module score in cells with a 'no filaments' or 'filaments' IF signature. Diamonds represent the average expression per patient (see J). Two-sided t-tests were realised on patient averages (N = 99) and the Cohen's d effect size coefficient is given.

Supplementary Figure 3 – Characteristics of IF-depleted cells: mRNA expression, spheroid radii and radial invasion index in Matrigel

A Quantification of *VIM*, *GFAP*, *NES* and *SYNM* mRNA levels in control (ctl) and IF knockout (ko) cells by quantitative real-time PCR analysis. mRNA levels were normalized to reference genes *GAPDH* and *Alu-J* (*N* = 3). P-values are the result of a two-sided paired *t*-test on the normalized gene expression values and Cohen's d effect size coefficient is given. Bars represent the average percentage of mRNA expression from control. Black diamonds represent the percentage of mRNA expression from control for each individual experiment. **B, C** Quantification of the GBM spheroid radius (excluding invading cells) at the start and after 40 hours of invasion in Matrigel. Dots represent individual spheroids (*n* = 29, 30 spheroids). The data was acquired from three biologically independent experiments and analyzed using a two-tailed unpaired *t*-tests. Cohen's d effect size coefficients are given on the graphs. **D** Schematic of Radial Invasion Index (RII) calculations. R_0 = radius at start of invasion. **E** RII of control and ko cells at 40 hours of invasion in Matrigel. An RII > 1 indicates invasion. Dots represent individual spheroids (*n* = 29, 30 spheroids). The data was acquired from three biologically independent experiments and analyzed using a two-tailed unpaired *t*-tests. Cohen's d effect size coefficients are given on the graphs. Statistics: Error bars represent SEM.

Supplementary Figure 4 – IF depletion does not impact migration speed in channels, nuclear mechanical properties or nuclear localization of YAP in 2D

A Schematic of the optical tweezers-based micro rheometer to measure the mechanical properties of the nucleus. **B, C** Rheological measurements of the nucleus from a viscoelastic analysis of the indentation phase. The spring constant K_n (*n* = 43, 46) and viscosity η_n (*n* = 37, 41) of ctl and ko nuclei were obtained by fitting the displacement of the bead x_b relative to the nucleus x_n $(x_b - x_n)$ curves with a viscoelastic model (see Methods). Data were acquired from three biologically independent experiments and analysed statistically using two-tailed unpaired *t*-tests. **D, E** Maximum indentation of the bead in the nucleus before ejection of the bead from the optical trap (D) and time (in seconds) at which the maximum indentation of ctl and ko nuclei was reached (E) (*n* = 47, 47). Data were acquired from three biologically independent experiments and analysed statistically using two-tailed unpaired *t*-tests. **F** Maximum indentation (µm) plotted against the time at which the maximum indentation of the nuclei was reached (s). **G** Immunostaining of sparsely plated ctl and ko cells for YAP and actin (labelled with phalloidin) on fibronectin-coated substrates of different rigidities. **H** Quantification of nuclear/cytoplasmic mean grey value for YAP immunostaining of cells on fibronectin-coated substrates of different rigidities. Grey dots represent values from individual cells (*n* = 35, 33, 107, 102, 85, 88, 91, 88, 110, 111 cells). Pink shapes represent averages of biological replicates (*N* = 2, 2, 5, 5, 4, 4, 4, 4, 5, 5). Two-tailed unpaired *t*-tests between ctl and ko cells plated on the same substrate indicated no significant difference in YAP nuclear/cytoplasmic ratio. Cohen's d effect size coefficients are given on the graphs. Statistics: Error bars represent SEM.

Supplementary Figure 5 – IF-depletion reorganizes the actin cytoskeleton but has a minor or no impact on 2D cell migration, cell and nuclear size

A Averaged fluorescence intensity images of ctl and IF-depleted cells plated on fibronectin-coated crossbow shaped micropatterns of 1000 µm^2 and labelled with Hoechst (nucleus) and phalloidin (actin). Overlays illustrate the location of the nucleus and actin distribution on the crossbow shaped micropattern. Mean intensity of all cells is shown. Nuclear intensity range: 0 to 400. Actin intensity range: 0 to 600 (*n* = 153, 102, *N* = 3). *Scale bar = 10 µm.* **B** Typical examples of a ctl cell and an IF-depleted cell plated on fibronectin-coated crossbow shaped micropatterns of 1000 µm^2 labelled with Hoechst (nucleus) and phalloidin (actin). *Scale bar = 10 µm.* **C** Phase-contrast images of GBM cell wound healing after 21 hours of migration. *Scale bar = 100 µm.* **D** Average migration speed over 8 hours of migration for ctl and ko cells measured by nuclear tracking of migrating cells. Grey dots represent values from individual cells (*n* = 47, 47 cells). Pink shapes represent averages of biological replicates (*N* = 3, two-tailed unpaired *t*-test). **E-G** Cell area, nuclear area and nuclear/cell ratio of ctl and ko cells sparsely plated on fibronectin coated glass coverslips (*n* = 85, 105, *N* = 3). Two-tailed unpaired *t*-test did not show significant differences. Cohen's d effect size coefficients are given on the graphs. Statistics: Error bars represent SEM.

Supplementary Tables

Supplementary Table 1: Glioblastoma cells with different IF gene expression signatures and total IF gene expression levels

Characteristics of IF cell clusters with different IF gene expression signatures.

FC = Fold Change of gene expression in cell cluster compared to all other cells (positive value = higher expression, negative value = lower expression); adj. p-value = adjusted p-value of a differential gene expression test in the cell cluster compared to all other cells

Supplementary Table 2: Lists of significantly differentially expressed genes between each IF cell cluster and all other cells.

Results of differential gene expression analysis between cells within each IF cell cluster compared to all other neoplastic cells.

gene_name = HGNC gene symbol, p_val = p-value of differentially gene expression test, avg_logFC = logFC between cells within the IF cell cluster (column 'cluster') and all other neoplastic cells, pct.1 = percentage of cells expressing the gene in the IF cell group in column 'cluster', pct.2 = percentage of cells expressing the gene in all other neoplastic cells, p_val_adj = p_val corrected for multiple testing, cluster = IF group for which the comparison to all other neoplastic cells was made.

Supplementary Table 3: List of overrepresented gene ontology clusters by differentially expressed genes between each IF cell cluster and other cells.

List represents GO clusters with a Fold Change of significant/expected (FoldE) > 1, a minimum of 20 significantly overrepresent genes (Significant) and an adjusted p-value of < 0.05 (padj). Lists represent gene ontology clusters overrepresented by genes with a significantly higher or lower expression in each IF cluster.

GO.ID = gene ontology ID code, Term = GO biological process term, Annotated = total number of genes belonging to the term, significant = number of differentially expressed genes belonging to the term, expected = number of differentially expressed genes expected to be overrepresented by chance, elimFisher = p-value from the Fisher's Extract test, sig.genes.in.term = gene names of significantly overrepresented genes, padj = Bonferonni corrected p-value, FoldE = Fold Change significant/expected genes, cluster = IF cluster name, pos/neg = overrepresented genes are increased in IF cluster (pos) or decreased (neg).

Supplementary Table 4: List of tumor periphery markers (analysis of data Darmanis et al., 2017)

List of significant differentially expressed gene between cells from the tumour core (cluster 3) and tumour periphery (cluster 2). Genes were considered significantly different at an adjusted p-value of < 0.05 (p_val_adj). Genes with a significantly higher expression in the tumour periphery (cluster 2, average log Fold Change (avg_logFC) > 0) were defined as tumour periphery markers.

gene = gene name, p_val = p-value of differential gene expression analysis, p_val_adj = corrected p-value for multiple testing, cluster = tumour core (3) or periphery (2) cell cluster.

Movies

Supplementary Video 1

Live-imaging of ctl (left) and IF-depleted cells (right) invading from a GBM spheroid into a 3D extracellular matrix composed of 75% Matrigel.

Supplementary Video 2

Example of an invading GBM cell trail in a 3D extracellular matrix composed of 75% Matrigel. A cell leading the invasion (leader cell, purple arrow) and a cell following the leader cell (follower cell, blue arrow) are shown.

Supplementary Video 3

Migration of ctl (left) and IF-depleted cells (right) through constrictions of 1.5 µm width. Phase contrast (top), Red CellTracker (middle) and sirDNA (bottom) labelling is shown. *Time = hours.*

Supplementary Video 4

Live-imaging of ctl (left) and IF-depleted cells (right) invading into the zebrafish brain after injection in the optic tectum. *Time = hours.*

Supplementary Video 5

One of the segmented nuclei rendered in 3D. The arrows on the mesh show the movement within the nucleus and its direction. The solid color encodes the magnitude of the displacements (blue low, red high). In the video, the nucleus is being deformed from the left. The local differences (gradients) in displacement create the strain that we use to differentiate cell populations. The strain is computed as described in the material and methods.