# DynaMorph: learning morphodynamic states of human cells with live imaging and sc-RNAseq

**Zhenqin Wu**[1,*], **Bryant B. Chhun**[2,*], **Galina Schmunk**[3,*], **Chang N. Kim**[3], **Li-Hao Yeh**[2], **Tomasz Nowakowski**[3,#,✉], **James Zou**[4,#,✉], **and Shalin B. Mehta**[2,#,✉]

[1]Department of Chemistry, Stanford University
[2]Chan Zuckerberg Biohub
[3]Department of Anatomy, University of California, San Francisco
[4]Department of Biomedical Data Science, Stanford University
[*]M.W., B.B.C., and G.S., contributed equally
[#] T.N., J.Z., and S.B.M., contributed equally
✉correspondence: `tomasz.j.nowakowski@gmail.com, jamesz@stanford.edu, shalin.mehta@czbiohub.org`

**Morphological states of human cells are widely imaged and analyzed to diagnose diseases and to discover biological mechanisms. Morphodynamics of cells capture their functions more fully than their morphology. Discovery of morphodynamic states of human cells is challenging, because genetic labeling or manual annotation may not be feasible. We propose a computational framework, DynaMorph, that combines quantitative label-free imaging and deep learning for automated discovery of morphodynamic states. As a case study, we apply DynaMorph to study the morphodynamic states of live primary human microglia, which are mobile immune cells of the brain that exhibit complex functional states. DynaMorph identifies two distinct morphodynamic states of microglia under perturbation by cytokines and glioblastoma supernatant. We find that microglia actively transition between the two states. Moreover, single-cell RNA-sequencing of the perturbed microglia shows that the morphodynamic states correspond to distinct transcriptomic clusters of the cells, revealing how perturbations alter gene expression and phenotype. DynaMorph can broadly enable automated discovery of functional states of cellular systems.**

**Keywords: Label-free imaging, deep learning, human microglia, morphodynamic states, single-cell RNA-sequencing.**

## Introduction

Organs and tissues of the human body consist of an astonishing diversity of cell types, classically described by their anatomical location, morphology, dynamic behavior, gene expression, and protein expression. These descriptions of cell types and cell states span multiple biological scales – the scales of molecules, organelles, cells, and tissue. Building holistic models of types and states of human cells requires mapping temporal changes in these descriptions in response to functionally-relevant and disease-relevant perturbations. Morphometry of human cells is widely used to analyze healthy and disease states of cells in clinical pathology and to discover fundamental biological mechanisms. However, automated analysis of morphodynamic states of human cells still remains an unsolved problem, because functional states are difficult to label either with molecular markers or manual annotations. While recent advances in single cell genomics have significantly advanced our understanding of molecular diversity of types of human cells, morphological and behavioral states of cells cannot be inferred from gene expression data alone. Identifying morphodynamic states in the context of disease can help elucidate the natural and induced changes in cell behavior, potentially informing novel therapeutics or diagnostics. As a result, quantitative analysis of morphodynamic states of human cells with high-throughput methods is a timely area of research.

Recent work on analysis of morphological states of cells has relied on images of fixed cells labeled with a panel of fluorescent markers (1), live three-dimensional imaging of the membrane labeled with genetic markers (2), and phase contrast imaging of live cells (3–6). The morphological states have been analyzed with low dimensional representations computed with geometric or biophysical models (3, 7–11), supervised learning of morphological labels (4, 12–17), and, recently, self-supervised learning of latent representations of morphology (5, 6). These analytical approaches have been inspired by the need for quantitative descriptions of specific, complex biological functions, such as motility of single cells (2, 3, 7, 8, 18), collective cell migration (9, 11), cell cycle (4, 12, 13), spatial gene expression (17), and spatial protein expression (14, 16).In addition, data-driven integration of the morphology and gene expression (13, 17, 19–22) is now enabling rapid analysis of functional roles of genes.

Despite this progress, we currently lack high-throughput and data driven approaches for analysis of morphodynamic states of human cells, including their temporal dynamics, due to few key technological limitations: a) Measurements of morphodynamics of live human cells is challenging because they are difficult to label with consistency and without perturbing the cellular behavior. Dozens of molecular reporters have been developed to visualize cell structure based on gene expression studies, but introducing molecular reporters in primary human cells is sometimes not possible or highly disruptive. b) Defining cell states from multidimensional imaging-based datasets remains challenging, because identifying statistically significant cell behavior by human visual system is often hard and annotation of cell states in high dimensional data is time consuming, even when recogniz-

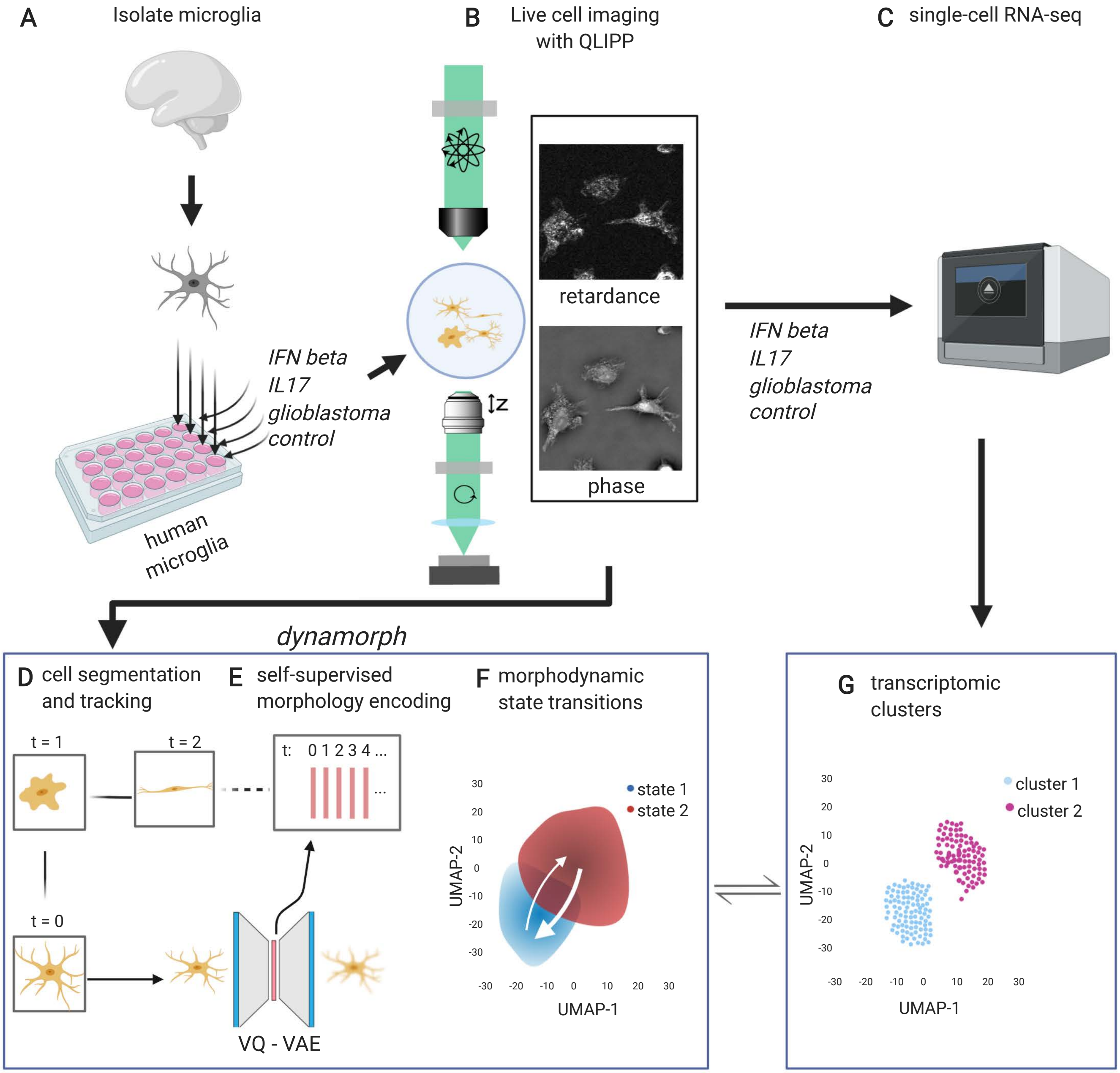


**Fig. 1. DynaMorph enables automated discovery of morphodynamic transitions in human microglia: (A)** Human microglia are isolated from brain tissue and plated in 24-well plates and perturbed with cytokines of relevance to infection (IFN beta, IL17) or cancer (glioblastoma extract), (B) Morphodynamics of perturbed microglia, along with control cells, are imaged using quantitative label-free imaging with phase and polarization imaging (QLIPP), which measures isotropic and anisotropic optical path lengths of cells. **(C)** Gene expression profiles of cells from each condition were measured using 10x single-cell RNA-sequencing (sc-RNAseq) data at the end of the imaging experiment. **(D)** Cells were tracked by detecting the bounding boxes around cells with a sparsely-supervised deep learning model and by linking the cells across frames. **(E)** A generalizable and quantitative representation of morphological states was learned from the thousands of tracked cells using a self-supervised model that reconstructs cell morphology. **(F)** Morphological states and transitions among states under each perturbation were revealed via dimensionality reduction (PCA and UMAP) algorithms and clustering of most significant features. **(G)** The correlations among the morphodynamic and transcriptomic states of the cells were analyzed by comparing the morphodynamic states and transcriptomic clusters under each perturbation.

able. c) While cell states can be expressed in terms of gene expression measured with single-cell RNA sequencing, similar vocabulary for description of cell states from their behavior remains to be developed. In this paper, we explore use of quantitative label-free measurements of cellular morphodynamics and deep learning to overcome these limitations and pursue automated discovery of functionally-relevant morphodynamic states of the human microglia.

Microglia are the resident macrophages of the central nervous system that are involved in brain development and homeostasis, as well as immune responses (23). Microglia's response to secreted cytokines or viral infections elicits profound changes at the transcriptomic and cell behavioral levels that are unique for different perturbations (24, 25). Microglia survey brain parenchyma with their highly motile processes and respond to changes in brain homeostasis by altering cell morphology and motility (26, 27), but large-scale features of their dynamic behavior are not well characterized partly due to the lack of tractable molecular labeling tools. For other types of motile cells, either learned features or the manually curated features (3, 5, 6, 18) have been used to describe complex cell dynamics, including random search, persistent migration towards target, changes in cell shape due to interaction with other cells and targets, and endocytosis, among others. However, changes in the expression of a small number of genes or morphological features of relevance to the other cell types may not reflect the full spectrum of microglia phenotypes and functions. Outstanding questions remain: Which states mediate microglia response to immune signals or infection? Are these states unique? What is the relationship between these morphodynamic states and their transcriptomic profile during immune response? We seek to answer these questions in the context of various immune responses as a proof-of-concept of our measurement and analysis method.

We acquired reproducible measurements of cellular architecture and dynamics of microglia under immunogenic perturbations (Figure 1A) using quantitative label-free imaging with phase and polarization (QLIPP) (16). To identify morphodynamic states for human microglia, we developed DynaMorph, a deep-learning enabled framework that automatically learns the morphodynamic states and transitions among states from high-dimensional live cell imaging data. The main modules of DynaMorph are as follows: A deep convolutional network for cell segmentation; cell tracking (Figure 1D); self-supervised model for learning a quantitative and generalizable latent representation of morphology (Figure 1E); and lastly, clustering of morphology and motion descriptors for identification of morphological states and transitions among states (Figure 1F).

By applying DynaMorph to live-cell videos of microglia perturbed with pro-inflammatory, anti-inflammatory cytokines, and glioblastoma supernatant, we show that the model generalizes well across new conditions and identifies biologically meaningful morphodynamics states. The model provides a quantitative description of cell morphology and enables discovery of infrequent transitions among cell states. The same set of cells were also profiled with single-cell RNA sequencing after imaging (Figure 1C) to analyze gene expression clusters. We discovered intriguing correlations between gene expression and morphodynamic clusters (Figure 1G).

In the following sections, we describe the key components of the DynaMorph pipeline and describe insights gained from the results.

## Results

**Label-free imaging and tracking of microglia.** Current methods for obtaining morphological measurements of cells require labeling cell membrane with genetic fluorophores, such as Green Fluorescent Protein (GFP) (2), or involve exogenous fluorescent compounds (1). Both approaches can influence the natural behavior of cells and require high-power light sources that may induce phototoxic or phototactic responses. Additionally, transfection efficiency of primary human microglia is usually low, providing a further challenge for cell labeling. To image the natural behavior of microglia over long periods, we used quantitative label-free imaging with phase and polarization (QLIPP) (16). QLIPP measures physical properties of the specimen in terms of its optical phase and retardance. The phase, or optical path length, reports the density of molecular assemblies in a cell. The retardance, or polarization-resolved optical path length, reports the density of anisotropic molecular assemblies such as cytoskeletal networks and lipid bilayers (16).

Cellular membrane is highly dynamic and can be difficult to label densely with fluorescence in primary cells, but cellular morphology can be quantitatively described with phase and retardance (Figure 1B. These measurements capture overall cell morphodynamics without the need to use fluorescent markers. Moreover, they do not rely on pre-defined marker sets, but rather provide an unbiased description of dynamic cell behavior, allowing one to uncover novel cell states. We explored live-cell membrane dyes (CellTracker and diI) for labeling the microglia as shown in Supplementary Figure 1 and found that distribution of membrane dyes was not consistent across cells. The measurements of morphology from phase and retardance were more consistent, quantitative, and gentle on cells for long-term imaging.

We acquired two sets of QLIPP images of cultured primary human microglia, over a 24-hour period. In the first set of experiments, primary human microglia were imaged using standard multi-dimensional acquisition with 52 time points at 27 minute intervals across 27 field of views, the data from which were used to build and develop the segmentation and morphological encoding pipeline. This set will be referred to as the training dataset in the following text. The second set of experiments imaged microglia derived from a biologically independent sample under different perturbations (4 fields of views for each condition). Images were collected for 24 hours, but with a much higher time resolution of 159 time points at 9 minute intervals. The resulting set, referred to as the test dataset, is analyzed for validation and cell state discovery.

In order to identify and track microglia, DynaMorph uses segmentation algorithms that required small amount of human annotation and resulted in robust detection of bounding-boxes around single cells. As illustrated in Supplementary Figure 2, Supplementary Figure 3, and detailed in Methods, we used ilastik (28) for interactive annotation that distinguished microglia from background and other cell types, followed by sematic segmentation of microglia pixels with neural network, followed by clustering to separate the segmentation results into cell masks. We used cell masks to identify bounding boxes and patches of images around each cells. A tracking algorithm was used to match and link static patches from adjacent frames of time-lapse images, which generated trajectories of cells along with motion of the cell throughout the tracking period (Figure 1D). In total, 2016 and 3715 microglia trajectories were extracted from the training and test set, respectively.

Above trajectories formed the corpus of data from which the morphological representation was learned and morphodynamic states were discovered.

**Learning a latent representation of morphology that generalizes across experiments.** Among all perturbations, we observed significant diversity in morphology and behavior of microglia. While microglia are classically characterized as amoeboid, ramified, mobile, or immobile, in our videos they appeared to demonstrate a variety of complex shapes and state transitions that we attempt to capture in our model. We used self-supervised deep learning (Figure 1E) to build an unbiased model of microglia morphology. Recent work has demonstrated that self-supervised learning with autoencoders (5, 6) can provide a quantitative model of cell morphology informed by all of data.

We trained the autoencoder for image compression and reconstruction. We tracked individual microglia cells from the training dataset as described in the previous section and trained a recent variant of autoencoder, vector quantized variational autoencoder (VQ-VAE) (29), with untreated microglia cells from the training set to encode and reconstruct the image patches. Latent space of the VQ-VAE model was then extracted and interpreted as the morphological space of microglia. In our experiments, multiple variants of autoencoder were tested and VQ-VAE appeared as the best-performing structure with the most stable reconstructions of phase and retardance (Supplementary Figure 4). Its intrinsic robustness against noise inherited from the quantization prior also guaranteed that minor imaging noises won't contribute to the latent distribution. The self-supervised encoding led to a noise-robust and interpretable morphological representation of different cells regardless of position or pose, while reconstruction from the encoding ensured that minimal information was lost in the process of encoding.

We also imposed a matching loss in the model during training to minimize frame-to-frame differences between latent vectors of the same cell along its trajectory, based upon the prior that cell architecture rarely changes significantly between consecutive frames. (see Methods for training details) This approach is motivated by the previous work that enforced temporal slowness constraint on the feature representation of adjacent frames (30). It is also related to the family of triplet loss that contrasts similar samples against non-similar samples to learn metrics of similarity (31). In our settings, cell embedding of a certain frame is encouraged to be closer to other frames of the same cell, which helps to regularize the locality of cell trajectories in morphological space, from which rare events such as significant size/density/shape changes due to cell state transition can become highly noticeable.

The learned morphology space of the DynaMorph autoencoder substantially compresses the data in the raw imaging– by over 680 times. Reconstruction loss averaged over all training cell patches is $0.16 \pm 0.08$SD after normalization of both channels. This was further validated on the test dataset, on which model reconstruction loss is $0.18 \pm 0.07$SD with no significant signs of overfitting. Other autoencoder-based models tested in the same procedure all showed higher reconstruction losses as seen from the comparisons in Supplementary Figure 4.

Comparison of reconstructed morphologies from the test set (Figure 2A and Supplementary Figure 5) and training set (Supplementary Figure 6A) along with the analysis of the latent representation described in the next section show that our self-supervised model trained on one experiment generalized well to unseen cells that were treated with multiple perturbations.

**Latent representation provides a quantitative description of morphology.** We sought to evaluate whether the learned representation provides quantitative metric of morphology, i.e., we sought to answer these questions: Are the most significant modes of morphology retrieved from the learned representation interpretable? Does the distance in the latent space capture morphological similarity?

***Interpretation of learned representation.*** We use principal component analysis (PCA) to visualize the learned morphology space. Distributions of the first two PCs are displayed in Figure 2B, in which each individual dot represents an encoded cell patch and the color indicates the size of its segmentation mask. A horizontal gradient of cell size is observed (Supplementary Figure 6B shows similar results for training cells), which suggests quantitative relationships between top PCs and cell properties.

To validate the necessity of morphology encoding, we evaluated direct application of PCA on the patch image inputs through the same procedure. The derived top PCs have weaker and more intertwined correlations (Supplementary Figure 7) and consequently are harder to interpret. The following clustering analysis also yielded worse performance (Supplementary Figure 8 and Supplementary Table 1), indicating that the encoding step helped in denoising the inputs and emphasizing informative properties that constitute the morphology of microglia.

To further interpret the latent space, we tried to link the top PCs with heuristic cell geometric properties. Qualitatively, we first sampled cell patches along PC axes (Figure 2C and Supplementary Figure 9), from which we recognized some significant correlations: PC1 relates with cell sizes; PC2 relates with peak phase and retardance; PC3 and PC4 relate with cell orientations. Then we performed a quantitative comparison based on Spearman's rank correlation coefficients (Figure 2D), which validated the correlations discovered above.

Our results showing that PC1, or the first significant mode of morphology, represents cell size is intriguingly related with the analysis of cell masks of other immune cell types in Chan *et al.* (6). They also show that the most significant shape mode of three different motile cell types is also the cell size. In addition, the second most significant morphological mode in our data is cell density as measured with phase and retardance. This result also comports with the result in Zaritsky *et al.* (5) that light scattering is a significant descriptor

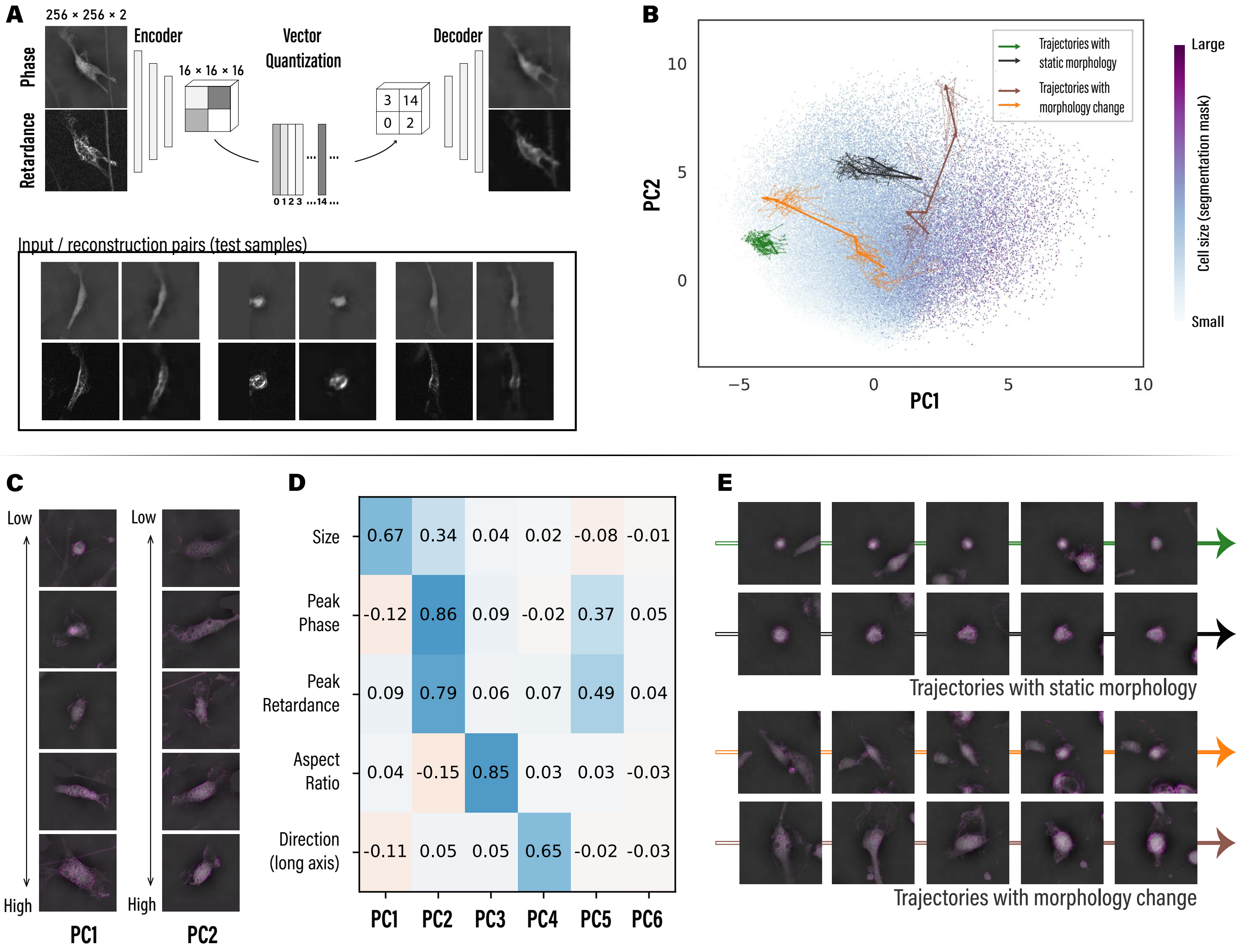


**Fig. 2. DynaMorph learns a generalizable and quantitative representation of cell morphology via self-supervision: (A)** We encoded the morphology of microglia in quantized latent vectors and reconstructed the morphology from the latent vectors by training a convolutional neural network model, vector quantized variational autoencoder (VQ-VAE), under self-supervision. The phase and retardance images from the test experiment (test set) shown here were encoded and reconstructed using models trained on the data from control experiment (training set). Comparison of reconstructions from training set Supplementary Figure 6 illustrates generalizability of the model to the test set. **(B)** Morphological modes that describe the most significant differences were computed with principal component analysis (PCA) of the latent vectors. Resulting top 2 principal components (PC) are visualized, in which each dot represents a single cell patch from the test set and its color indicates the size of cell segmentation mask. 4 representative trajectories: two cells with static morphology (green and black) and two cells undergoing changes in morphology (orange and brown) are plotted in the PCA space as arrows. **(C)** The first two morphological modes (PC1 and PC2) were interpreted by sampling representative patches along each PC axis and are shown here. Clear trends of changing size/cell density could be observed. Note that patches shown combine both phase channel (as grey scale) and retardance channel (as magenta shades). **(D)** We explored correlation between the top 6 PCs and selected geometric properties of the test set using Spearman's rank correlation coefficients. PC1 and PC2 both show positive correlations with cell size; PC2 is also highly correlated with cell's peak phase and retardance that measure isotropic and anisotropic density, respectively; PC3 is correlated with cell's aspect ratio; PC4 is correlated with the orientation of the cell (long axis of the cell body). Note that all results are consistent between training set cells and test set cells (Supplementary Figure 6C). **(E)** Representative cells plotted in **(B)** are visualized. Significant changes in morphology could be linked to jumps in the latent space.

of cell behavior. Thus, our results show that the latent representation automatically identified morphological modes that capture morphological diversity.

Intriguingly, while 80% of the shape variance of ameboid cells analyzed in Chan *et al.* (6) is captured by 5 shape modes, we found that the first 4 morphological modes (or PCs) of microglia account for less than 20% of all variance (Supplementary Figure 10A). The remaining variance might exist in more complex features such as diversity of protrusions, the variations in cell density, location of nuclei in migrating cells, etc. UMAP(32) projections of the patch-level latent vectors show no apparent clustering patterns (Supplementary Figure 10B and C), suggesting that microglia morphology undergoes continuous change and hence no discrete states could be defined based on static frames.

***Trajectories in the learned morphology space characterizes dynamics.*** Latent representations of morphology on static frames were then concatenated to represent the morphological variations throughout the trajectories of each single cell. As a sanity check, the distance between the learned representations of adjacent and non-adjacent frames in the video is compared before and after quantization. We found that the distance between temporally adjacent frames of the cells is substantially less than non-adjacent frames, indicating that our morphology space captures the continuity of cellular dynamics (Supplementary Figure 11).

Among all observed test cells, most trajectories are localized in the morphological space, which likely reflects the fact that for most cells, morphology features remain relatively constant throughout the imaging period. In Figure 2E we illustrate some representative trajectories (marked in green and black) that are in static cell states and have stable appearances. Their PC1 and PC2 indicators are located in confined areas. However, we also observed a number of “leap” events in the PC plane, which coincide with cells undergoing significant transitions in morphological appearances. Representative examples (marked in orange and brown) are also shown in Figure 2E, in which transitions in size or density could be observed. Full video clips of these example cell trajectories could be found in Video Set 1.

**Morphology and motility of microglia under perturbations.** Microglia cells are highly dynamic cells and can rapidly change their morphology and motion in response to external stimuli. We hypothesized that exposing microglia to different types of disease-relevant perturbations could refine the description of biologically-relevant morphodynamic states, or reveal novel states not seen under homeostatic culture conditions.

To test this hypothesis, we used the following conditions to mimic pro- and anti-inflammatory brain states: (1) untreated control (Control), (2) pro-inflammatory cytokine interleukin 17A (IL17), (3) anti-inflammatory cytokine interferon beta (33) (IFN beta), and (4) supernatant from cultured primary human glioblastoma cells (GBM) to model complex inflammatory tumour environment.

Imaging patches from the test conditions were processed and encoded through the DynaMorph pipeline. Video clips of some representative cell trajectories from each condition can be found in Video Set 2. Patch reconstruction and latent space interpretations both show consistent results to the training dataset as discussed above. (Figure 2). The morphology descriptors from the DynaMorph autoencoder and motion descriptors from cell tracking (mean displacement, discussed below), were then concatenated as trajectory feature vectors to represent cell behavior. These vectors could then be clustered and visualized to identify discrete morphodynamic states of the cell population. (Figure 1F)

For each cell, we collected and averaged its PC1 and PC2 values of static frames along the trajectory (Figure 3A). Distributions of these trajectory-averaged PC1 and PC2 values are summarized for each treatment and plotted in Figure 3B and Supplementary Figure 12. It is clearly seen that IL17 and IFN beta subsets have similar distributions on both descriptors, which differ from control and GBM subsets, especially on average PC2 values. This suggests that microglia under IL-17 and IFN-beta treatment tend to have lower density while maintaining a similar size.

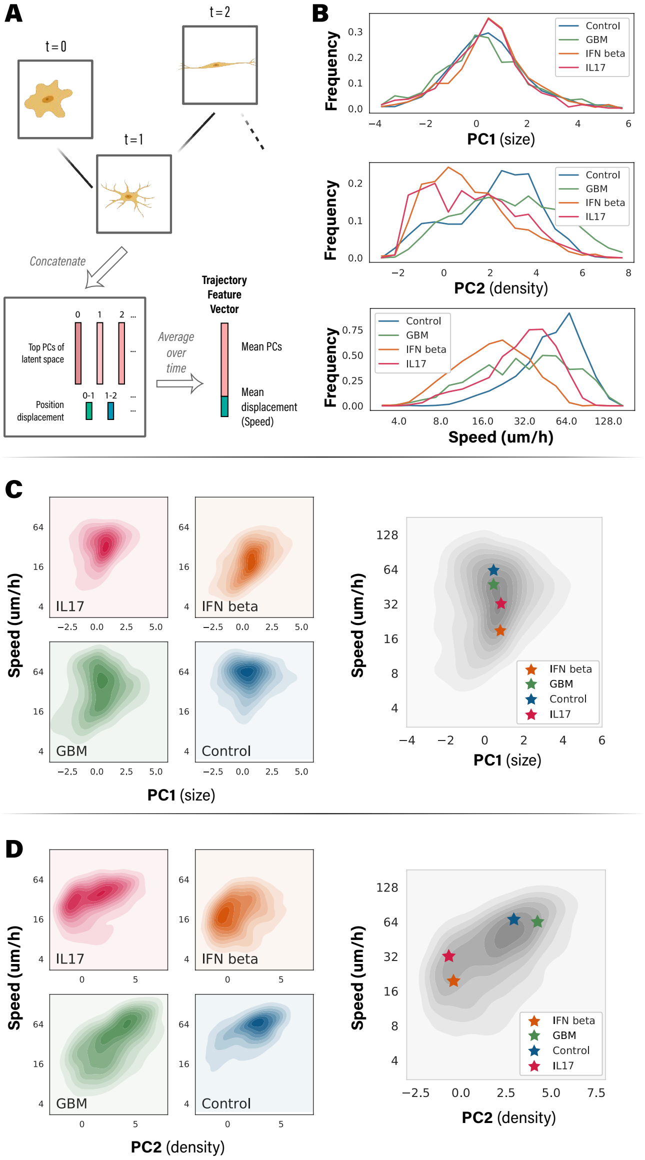


**Fig. 3. DynaMorph reveals differences in morphodynamics of microglia under different perturbations. (A)** To compare the morphodynamics of cell populations under different perturbations, we computed trajectory-averaged features as follows: principal components of the latent vectors of morphology of the frames of a tracked cell were concatenated with position displacements between frames and averaged over each trajectory. **(B)** We analyzed the distributions of trajectory features under multiple perturbations by plotting probability densities of trajectory-averaged PC1, PC2 and mean displacement, i.e., mean speed. The geometric properties of cell size and density correlated highly with PC1 and PC2, respectively. Trajectories under IL17 and IFN beta treatments show significant differences from trajectories under GBM and control treatments in cell density and speed. **(C)** and **(D)** We analyzed changes in average cell size (PC1) and average cell speed, as well as average cell density (PC2) and speed by plotting their joint distributions. Left panels show density plots of distributions under different perturbations. Right panels show the combined distributions, with apexes under each perturbations marked with stars. Separations among IL17/IFN beta and GBM/control conditions observed in **(B)** were confirmed.

To better understand the interplay between morphology and dynamic behavior, we further analyzed the movement patterns of microglia by calculating and averaging position displacements between frames (speed) along trajectories (Figure 3A), generating feature vectors by combining with morphology descriptors. Speed is then plotted individually (Figure 3B) and against top PCs as kernel density estimation (KDE) plots (Figure 3C and D). Microglia cells display a broad range of motility under different conditions, with the fast-moving population (control subset) traveling over 2 times faster than the slow-moving population under an anti-inflammatory IFN beta treatment.

In the joint distribution, apexes of subsets' KDE plot in Figure 3D (right panel) can be separated into two groups: IL17 and IFN beta versus GBM and control. In their individual KDE plots, control and IFN beta subsets form two densely populated groups on different morphology and motility values, while IL17 and glioblastoma have broader distribution that cover both groups. The separation was further revealed in a discriminative study that predict treatment/condition purely based on trajectory feature vectors (Supplementary Figure 13). In the results from gradient boosted decision tree models (34), cells from control subset could be better separated, especially from IFN beta subset, while larger confusions exist for GBM and IL-17 subsets.

Together, this analysis suggests that at least two broad morphodynamic states of microglia can be detected using DynaMorph. In particular, exposure of microglia to IFN beta or IL17, which mimic responses to viral infections, induces a state characterized by reduced density and migration speed. In contrast, exposure to GBM supernatant, which likely contains pro-inflammatory and anti-inflammatory cytokines, induces more subtle, but detectable changes in cell dynamics.

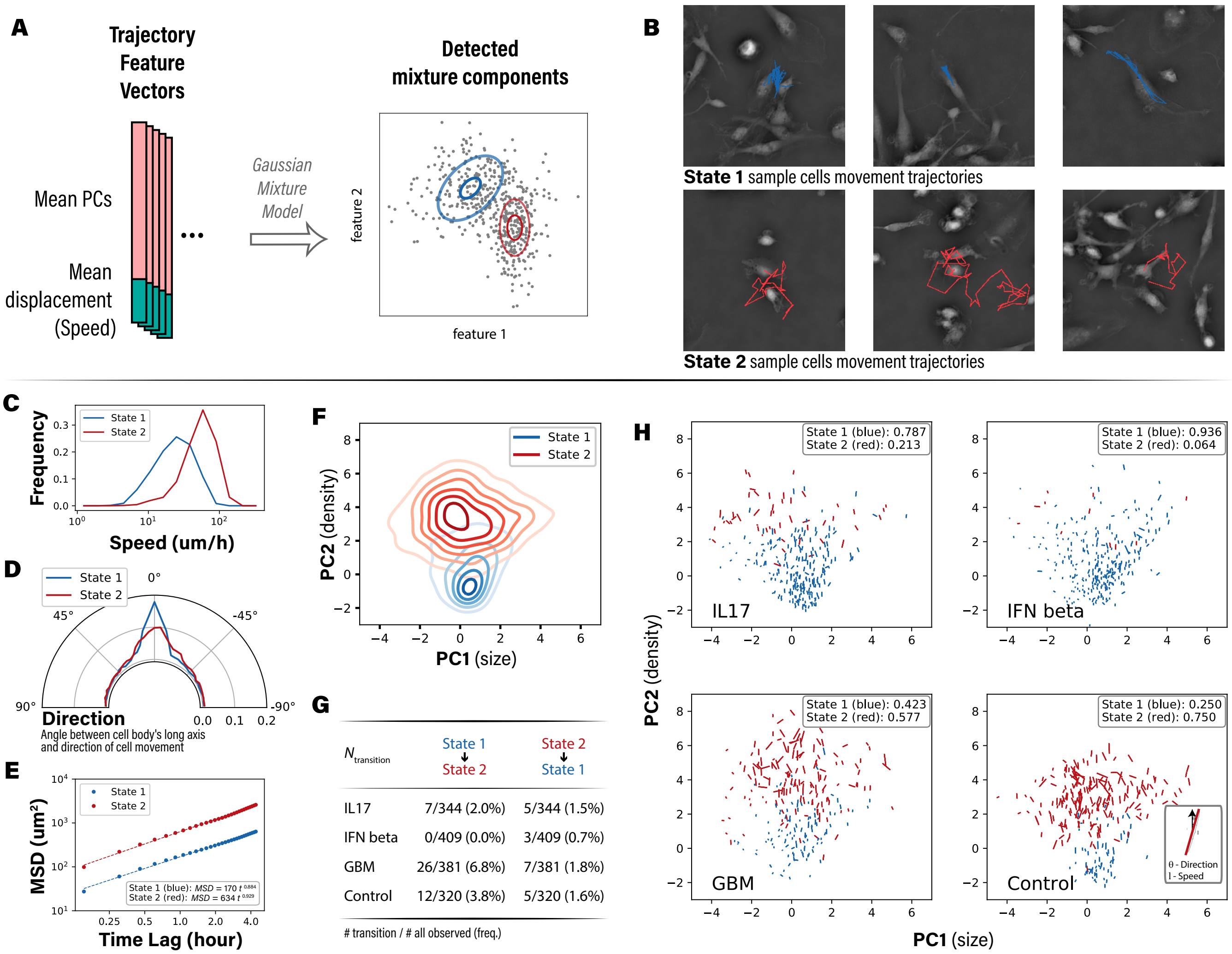
**Identification of morphological states and transitions between them.** To further probe the hypothesis that latent space and motion descriptors of microglia under perturbations are informative, or reveal novel morphodynamic states, we applied unsupervised clustering on the test cell patches to locate and quantify the underlying cell states. For each cell we observed in the test dataset, its trajectory-averaged PCs and mean displacements between frames were gathered and combined as feature vectors (Figure 3A). These cell feature vectors were then normalized and fitted using a Gaussian Mixture Model (GMM) with 2 components (Figure 4A). In the expectation-maximization (EM) procedure, we assumed that cells from each subset are sampled from a Bernoulli distribution dependent on the condition. The parameters of these distributions as well as the centers of mixture components were estimated iteratively. (See Methods for details)

Two components are extracted from the combined test population following the procedure above, which will be referred to as state-1 and state-2 respectively in the following text. Comparing with inter-condition differences, inter-state separation is much stronger (Supplementary Figure 14) and hence has better chances in finding morphodynamic modes of microglia. Representative samples from the two states are illustrated in Figure 4B, with their movement trajectories plotted as colored lines. Representative cell trajectories could also be found in Video Set 3. Morphological distributions of the two states are described in Figure 4F, details could be found in Supplementary Table 1. Qualitatively, based on the appearance and features of the two components, we describe state-1 (blue) cells as large-sized, low density and slow-moving population, and state-2 (red) cells as a higher density and fast-moving population. The two states differ the most on PC2 and mean displacement, and have slight difference on PC1 as well.

A more quantitative view on cell motion demonstrates that state-2 cells, enriched in control subset, are on average 2.3 times faster than state-1 cells (Figure 4C) that appear mostly in IFN beta and IL17 subsets, in line with the subset-level speed differences we noted above. We also consider directionality in the two states, which represents an important feature of cell migration. Note that the movement trajectories of state-1 cell in Figure 4B aligned with the direction of its cell body. We validated this motion pattern on all microglia cells by calculating angles between movement and long axis of cell body in each frame and conducting vector sums over these angles. Resulting trajectory-summed directions are visualized in Figure 4H. Distributions of the directions shows dominating peaks at 0°(Figure 4D), indicating a very high preference for cell body-aligned motion. Notably, state-2 cells have less such tendencies, which is in part due to the morphology distinction as state-2 cells usually appear as dense, round shapes that do not have significant long axes, in contrast to state-1 cells that appear more in oval shapes. Analysis on the Mean Square Displacement (MSD) of the two states were then conducted. Figure 4E and Supplementary Figure 15 plot the log-log fit of MSD over time lag. Both states have similar slopes ($\sim 0.9$) that are close to 1, indicating a near-random motion. State-2 have larger intercept, indicating a larger diffusion constant and hence faster movements. Taken together, the quantitative measurements of speed and directionality for cell populations defined by each GMM state further validate that these populations are unique and descriptive of each perturbation condition.

Among the different perturbation conditions in the test population, cells are widely distributed between two states depending on the treatment (Figure 4H and Supplementary Table 1): majority of the control subset cells are in state-2 ($\sim 75\%$), while almost all IFN beta-treated cells are in state-1 ($> 90\%$); GBM subset falls between the two extremes with near-equal split on the two states. Based on the previous observations that microglia adopt an amoeboid morphology upon exposure to inflammatory stimuli, we speculate that state-2 with faster moving, denser and round-shaped cells represents activated microglia, while state-1 that is characterized by ramified and slower moving cells represents surveying microglia. The treatment of microglia with glioblastoma supernatant lead to an intermediate distribution between the two states, suggesting a complex mixture of both pro- and anti-inflammatory cytokines.

Furthermore, the quantitative definition of states allows us to rigorously characterize state transition events such as the ones shown in Figure 2E (orange and brown trajectories). Given the morphology and motion changes throughout the imaging period, we separated each trajectory into segments (2 hour each) and applied the unsupervised GMM to estimate posterior probabilities of cell states for each segment.



| $N_{transition}$ | State 1 → State 2 | State 2 → State 1 |
|---|---|---|
| IL17 | 7/344 (2.0%) | 5/344 (1.5%) |
| IFN beta | 0/409 (0.0%) | 3/409 (0.7%) |
| GBM | 26/381 (6.8%) | 7/381 (1.8%) |
| Control | 12/320 (3.8%) | 5/320 (1.6%) |

# transition / # all observed (freq.)

**Fig. 4. Cell trajectories are clustered into two states based on the morphodynamic features.** **(A)** We identified two states of the trajectories by pooling their feature vectors and clustering them with a Gaussian Mixture Model (GMM). **B** Representative trajectories from the two states are shown: initial frames are visualized, and the movement of cells in the following frames are marked. **(C–F)** We compared two morphodynamic states with multiple metrics: **(C)** cells in state-1 were found to be slower than cells in state-2 from the probability densities of trajectory-averaged speeds. **(D)** cells in state-1 were found to align more with directions of migration than cells in state-2 from the probability densities of angles between cell body (long axis) and the directions of movement. **(E)** cells in state-1 and state-2 were found to undergo random motion over the time-scale of hours from the mean squared displacement (MSD) curves of the trajectories. Consistent with the speed distribution, state-2 cells travel longer distances on all time scales. **(F)** cells in state-1 had narrower distribution of size and density relative to cells in state-2, as seen from the Kernel density estimates (KDE) of PC1 (cell size) and PC2 (cell density). **(G)** We counted the number of trajectories that transitioned between two states under each perturbation and found rare instances of transition as noted in the table. Note that only trajectories longer than 4.5 hours are considered. **(H)** Metrics shown in **(C)–(F)** were elaborated per perturbation with scatter plots of the metrics: each marker represents a single trajectory, whose length indicates the speed of the cell. Direction of the marker is aligned with the angle between cell body and movement, with a vertical line indicating a perfect alignment. (See legend in the bottom right panel). Fraction of trajectories that occupy either state under each perturbation are noted in the insets.

Cells that have different states assigned to different segments would be of great interest to the study of dynamics and behavior of microglia. We counted the observed transition cases in the test set and report the numbers in Figure 4G. In our analysis, transition events are very rare among cells treated with IFN beta, while cells treated with GBM supernatant are more enriched for such events. While both directions of transitions were observed within the imaging period, cells in state-1 are more likely to transition to state-2 than vice versa. Representative state-transition cells can be found in Video Set 4.

**Correlation between morphodynamic states and transcriptomic clusters.** We examined the transcriptomes of cultured microglia to identify the key molecular sub-types that could correspond to morphodynamic states. In this experiment, cells from three conditions (control, IFN beta and GBM) were processed for single cell mRNA sequencing. Note that the majority of the measured cells ($\sim 90\%$) are from control subset. We generated UMAP projections of both morphodynamic features (Figure 5A) and scRNA-seq results to parallel the comparison, which are plotted in Figure 5B, C, D. A similar two-component clustering was performed on scRNA data, using Leiden clustering (35). The distributions of state/cluster assignments in each subset are plotted as bars in Figure 5E. Notably, cells from the control subset form a large group (cluster 2) and three small isolated groups (combined as cluster 1), in which the isolated groups could be further separated and interpreted through a more fine-grained

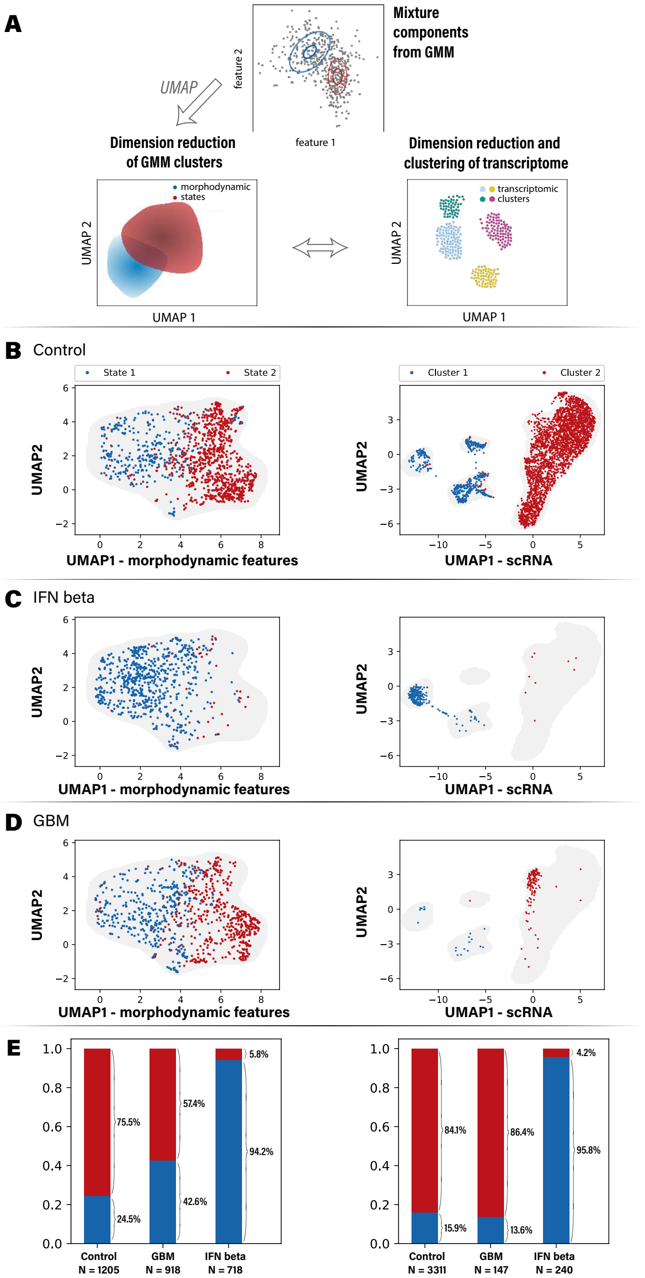


**Fig. 5. Comparison of morphodynamic states and gene expression per perturbation. (A)** We compared relative abundance of state-1 and state-2 with clusters of gene expression by computing consistent low dimensional projections (UMAP) of the trajectory feature vectors and transcriptome. Scatter plots of cell populations from three perturbations (**(B)** Control; **(C)** IFN beta; **(D)** GBM) are shown. Each dot represents a trajectory from imaging or a cell measured for transcriptome, and grey shades show the outlines of the distributions. Left panels show the UMAP projections of combination of morphology features (first 48 PCs) and motility feature (speed, mean displacement) derived from imaging. GMM clustering results are illustrated as the colors of the dots. Right panels show UMAP projections of single-cell RNA sequencing (scRNA-seq) results, on which a similar two-state clustering is performed. **(E)** Distributions of cluster/state assignments of cells from each test condition are plotted as bar graphs. Interestingly, in control and IFN beta conditions, the occupancy ratios between morphodynamic state-1 and state-2 matched with occupancy ratios of gene expression clusters 1 and 2.

clustering procedure. (Supplementary Figure 16A) Cluster 1 accounts for $\sim 15\%$ of all control cells, this striking difference in relative abundance of cells in different clusters coincides with the ratio of cells in states 1 and 2 defined on the morphodynamic space. Interestingly, exposure to IFN beta, which is enriched for morphodynamic state-1, also shows enrichment in transcriptomic cluster 1, appearing as one of the isolated group. In contrast, microglia that were treated with supernatant from GBM cultures, are transcriptomically more similar to the majority of control microglia cells, consistent with more subtle differences in morphodynamic states.

To address this further, we identified the top differentially-expressed genes (Supplementary Figure 16B) for each transcriptomic cluster. Cluster 1, which is the minor group in control subset and enriched for cells exposed to IFN beta, has up-regulated genes encoding antiviral proteins (RSAD2, MX1, etc.) and interferon-induced proteins (IFIT1, IFIT2, IFIT3, etc.). Part of the cluster is also enriched for cytoskeleton and microtubule-related genes (MAP1B, STMN1, TUBA1A, etc.), which would usually cause structural changes in cells. Therefore, the observed shape and density differences between different morphodynamic states could potentially be correlated with the corresponding cell's expression changes.

## Discussion

In this work, we present DynaMorph: a deep learning framework for automatic segmentation, feature extraction and analysis for label-free live cell imaging. Multiple tools, including conventional machine learning (random forest, GMM, DBSCAN), deep convolutional networks (U-Net), and variational autoencoder (VQ-VAE) are incorporated to facilitate/optimize each stage of data processing and analysis. The resulting workflow allows us to analyze cell dynamics and behavior on both individual cell-level and population-level.

Our work formalizes an analytical approach for data-driven discovery of morphodynamic cell states based on the extractions of morphology and motion descriptors. This method enables sensitive detection of transition events induced by exogenous perturbation. Here, we specifically demonstrate the utility of this method to discover and classify morphodynamic states of human primary microglia cultured *in vitro*, and to detect a shift in morphodynamic states upon exposure to polarizing, disease-relevant environments. We have further confirmed, using single cell transcriptomics, that changes in morphodynamic states are paralleled by changes in gene expression, suggesting that morphodynamic features extracted from live imaging might serve as correlates to molecular signatures. In this work, we specifically targeted microglia and unveiled distinct morphodynamic states from its population under different perturbations. DynaMorph can be used with other cell types and other live cell imaging data to delineate morphodynamic distributions. All components in the pipeline are interchangeable based on the needs of analyses and characteristics of the imaging system.

The high-throughput data-driven approach for discovering and classifying morphodynamic cell states developed here has the potential to facilitate unbiased, real-time analysis of imaging data. It further allows filtering and selection of cell-of-interest by identification of critical points of state transition along the trajectory. These specific cells could be used for downstream single cell profiling to discover molecular correlates of those events and consolidate the relationship between morphodynamic behavior and transcriptomic alteration.

It should also be noted that certain caveats still exist in this approach: instance segmentation and cell tracking are limited by the culture density as denser environment will significantly increase the difficulty of segmenting and tracking individual cells. Methods that incorporate more human supervision would be required to better extract and analyze cell behavior in a complex environment. VQ-VAE provides a natural interface to characterize the latent vector states in a discrete fashion, which is lacking in most other autoencoder-based method. This intrinsic advantage could be further examined and linked with other quantitative attributes of the cell. Direct correlation between the single cell morphodynamic behavior and its gene transcriptome profile is lacking. We found strong agreement between the gene expression changes and morphodynamics in response to different perturbations at the resolution of cell population. The transitions in gene expression and morphodynamics driven by the perturbations can be further elucidated by measuring gene expression and morphodynamics with single- cell resolution. Post-imaging staining of cells using markers that discriminate gene expression clusters seen in single cell transcriptomes can provide this information.

Taken together, DynaMorph is a versatile and expressive framework to analyze visual features of live cell imaging. Extraction of morphology and motility descriptors could be coupled with a broad spectra of downstream tasks including anomaly detection, unsupervised state discovery, supervised phenotype classification, etc. Integration with biochemical/transcriptomic assays could further enable a more comprehensive understanding of cell behavior.

## References

1. Mark-Anthony Bray, Shantanu Singh, Han Han, Chadwick T. Davis, Blake Borgeson, Cathy Hartland, Maria Kost-Alimova, Sigrun M. Gustafsdottir, Christopher C. Gibson, and Anne E. Carpenter, "Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes," *Nature Protocols*, vol. 11, no. 9, pp. 1757–1774, Sept. 2016.
2. Meghan K. Driscoll, Erik S. Welf, Andrew R. Jamieson, Kevin M. Dean, Tadamoto Isogai, Reto Fiolka, and Gaudenz Danuser, "Robust and automated detection of subcellular morphological motifs in 3D microscopy images," *Nature Methods*, vol. 16, no. 10, pp. 1037–1044, Oct. 2019.
3. Z. Pincus and J. A. Theriot, "Comparison of quantitative methods for cell-shape analysis," *Journal of Microscopy*, vol. 227, no. 2, pp. 140–156, 2007.
4. David A. Van Valen, Takamasa Kudo, Keara M. Lane, Derek N. Macklin, Nicolas T. Quach, Mialy M. DeFelice, Inbal Maayan, Yu Tanouchi, Euan A. Ashley, and Markus W. Covert, "Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments," *PLoS computational biology*, vol. 12, no. 11, pp. e1005177, Nov. 2016.
5. Assaf Zaritsky, Andrew R. Jamieson, Erik S. Welf, Andres Nevarez, Justin Cillay, Ugur Eskiocak, Brandi L. Cantarel, and Gaudenz Danuser, "Interpretable deep learning of label-free live cell images uncovers functional hallmarks of highly-metastatic melanoma," *bioRxiv*, p. 2020.05.15.096628, May 2020.
6. Caleb K. Chan, Amalia Hadjitheodorou, Tony Y.-C. Tsai, and Julie A. Theriot, "Quantitative comparison of principal component analysis and unsupervised deep learning using variational autoencoders for shape analysis of motile cells," *bioRxiv*, p. 2020.06.26.174474, June 2020.
7. Kinneret Keren, Zachary Pincus, Greg M. Allen, Erin L. Barnhart, Gerard Marriott, Alex Mogilner, and Julie A. Theriot, "Mechanism of shape determination in motile cells," *Nature*, vol. 453, no. 7194, pp. 475–480, May 2008.
8. Wallace F. Marshall, "Origins of cellular geometry," *BMC Biology*, vol. 9, no. 1, pp. 57, Aug. 2011.
9. Shahriar Shadkhoo and Madhav Mani, "The Role of Cytoplasmic Interactions in the Collective Polarization of Tissues and its Interplay with Cellular Geometry," *bioRxiv*, p. 289520, Mar. 2018.
10. P. Rezaie, G. Trillo-Pazos, J. Greenwood, I. P. Everall, and D. K. Male, "Motility and Ramification of Human Fetal Microglia in Culture: An Investigation Using Time-Lapse Video Microscopy and Image Analysis," *Experimental Cell Research*, vol. 274, no. 1, pp. 68–82, Mar. 2002.
11. Assaf Zaritsky, Erik S. Welf, Yun-Yu Tseng, M. Angeles Rabadán, Xavier Serra-Picamal, Xavier Trepat, and Gaudenz Danuser, "Seeds of Locally Aligned Motion and Stress Coordinate a Collective Cell Migration," *Biophysical Journal*, vol. 109, no. 12, pp. 2492–2500, Dec. 2015.
12. Erick Moen, Enrico Borba, Geneva Miller, Morgan Schwartz, Dylan Bannon, Nora Koe, Isabella Camplisson, Daniel Kyme, Cole Pavelchek, Tyler Price, Takamasa Kudo, Edward Pao, William Graf, and David Van Valen, "Accurate cell tracking and lineage construction in live-cell imaging experiments with deep learning," *bioRxiv*, p. 803205, Oct. 2019.
13. Beate Neumann, Thomas Walter, Jean-Karim Hériché, Jutta Bulkescher, Holger Erfle, Christian Conrad, Phill Rogers, Ina Poser, Michael Held, Urban Liebel, Cihan Cetin, Frank Sieckmann, Gregoire Pau, Rolf Kabbe, Annelie Wünsche, Venkata Satagopam, Michael H. A. Schmitz, Catherine Chapuis, Daniel W. Gerlich, Reinhard Schneider, Roland Eils, Wolfgang Huber, Jan-Michael Peters, Anthony A. Hyman, Richard Durbin, Rainer Pepperkok, and Jan Ellenberg, "Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes," *Nature*, vol. 464, no. 7289, pp. 721–727, Apr. 2010.
14. Chawin Ounkomol, Sharmishtaa Seshamani, Mary M. Maleckar, Forrest Collman, and Gregory R. Johnson, "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy," *Nature Methods*, vol. 15, no. 11, pp. 917, Nov. 2018.
15. Eric M. Christiansen, Samuel J. Yang, D. Michael Ando, Ashkan Javaherian, Gaia Skibinski, Scott Lipnick, Elliot Mount, Alison O'Neil, Kevan Shah, Alicia K. Lee, Piyush Goyal, William Fedus, Ryan Poplin, Andre Esteva, Marc Berndl, Lee L. Rubin, Philip Nelson, and Steven Finkbeiner, "In silico labeling: Predicting fluorescent labels in unlabeled images," *Cell*, vol. 173, no. 3, pp. 792–803.e19, Apr. 2018.
16. Syuan-Ming Guo, Li-Hao Yeh, Jenny Folkesson, Ivan Ivanov, Anitha Priya Krishnan, Matthew G. Keefe, David Shin, Bryant Chhun, Nathan Cho, Manuel Leonetti, Tomasz J. Nowakowski, and Shalin B. Mehta, "Revealing architectural order with quantitative label-free imaging and deep learning," *bioRxiv*, p. 631101, Nov. 2019.
17. Bryan He, Ludvig Bergenstråhle, Linnea Stenbeck, Abubakar Abid, Alma Andersson, Åke Borg, Jonas Maaskola, Joakim Lundeberg, and James Zou, "Integrating spatial gene expression and breast tumour morphology via deep learning," *Nature Biomedical Engineering*, pp. 1–8, June 2020.
18. Jacob C. Kimmel, Amy Y. Chang, Andrew S. Brack, and Wallace F. Marshall, "Inferring cell state by quantitative motility analysis reveals a dynamic state system and broken detailed balance," *PLOS Computational Biology*, vol. 14, no. 1, pp. e1005927, Jan. 2018.
19. Anne E. Carpenter and David M. Sabatini, "Systematic genome-wide screens of gene function," *Nature Reviews Genetics*, vol. 5, no. 1, pp. 11–22, Jan. 2004.
20. Keara Lane, David Van Valen, Mialy M. DeFelice, Derek N. Macklin, Takamasa Kudo, Ariel Jaimovich, Ambrose Carr, Tobias Meyer, Dana Pe'er, Stéphane C. Boutet, and Markus W. Covert, "Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-$\kappa$B Activation," *Cell Systems*, vol. 4, no. 4, pp. 458–469.e5, Apr. 2017.
21. Karren Dai Yang, Anastasiya Belyaeva, Saradha Venkatachalapathy, Karthik Damodaran, Adityanarayanan Radhakrishnan, Abigail

Katcoff, G. V. Shivashankar, and Caroline Uhler, "Multi-Domain Translation between Single-Cell Imaging and Sequencing Data using Autoencoders," *bioRxiv*, p. 2019.12.13.875922, Dec. 2019.

22. Kaytlyn A. Gerbin, Tanya Grancharova, Rory Donovan-Maiye, Melissa C. Hendershott, Jackson Brown, Stephanie Q. Dinh, Jamie L. Gehring, Matthew Hirano, Gregory R. Johnson, Aditya Nath, Angelique Nelson, Charles M. Roco, Alexander B. Rosenberg, M. Filip Sluzewski, Matheus P. Viana, Calysta Yan, Rebecca J. Zaunbrecher, Kimberly R. Cordes Metzler, Vilas Menon, Sean P. Palecek, Georg Seelig, Nathalie Gaudreault, Theo Knijnenburg, Susanne M. Rafelski, Julie A. Theriot, and Ruwanthi N. Gunawardane, "Cell states beyond transcriptomics: Integrating structural organization and gene expression in hiPSC-derived cardiomyocytes," *bioRxiv*, p. 2020.05.26.081083, May 2020.
23. Michael W Salter and Beth Stevens, "Microglia emerge as central players in brain disease," *Nature medicine*, vol. 23, no. 9, pp. 1018, 2017.
24. Chintan Chhatbar, Claudia N Detje, Elena Grabski, Katharina Borst, Julia Spanier, Luca Ghita, David A Elliott, Marta Joana Costa Jordao, Nora Mueller, James Sutton, et al., "Type i interferon receptor signaling of neurons and astrocytes regulates microglia activation during viral encephalitis," *Cell reports*, vol. 25, no. 1, pp. 118–129, 2018.
25. Starlee Lively and Lyanne C Schlichter, "Microglia responses to pro-inflammatory stimuli (lps, ifn$\gamma$+ tnf$\alpha$) and reprogramming by resolving cytokines (il-4, il-10)," *Frontiers in Cellular Neuroscience*, vol. 12, pp. 215, 2018.
26. Axel Nimmerjahn, Frank Kirchhoff, and Fritjof Helmchen, "Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo," *Science*, vol. 308, no. 5726, pp. 1314–1318, 2005.
27. Louis-Philippe Bernier, Christopher J Bohlen, Elisa M York, Hyun B Choi, Alireza Kamyabi, Lasse Dissing-Olesen, Jasmin K Hefendehl, Hannah Y Collins, Beth Stevens, Ben A Barres, et al., "Nanoscale surveillance of the brain by microglia via camp-regulated filopodia," *Cell reports*, vol. 27, no. 10, pp. 2895–2908, 2019.
28. Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N. Straehle, Bernhard X. Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, Kemal Eren, Jaime I. Cervantes, Buote Xu, Fynn Beuttenmueller, Adrian Wolny, Chong Zhang, Ullrich Koethe, Fred A. Hamprecht, and Anna Kreshuk, "Ilastik: Interactive machine learning for (bio)image analysis," *Nature Methods*, vol. 16, no. 12, pp. 1226–1232, Dec. 2019.
29. Aaron van den Oord and Oriol Vinyals, "Neural discrete representation learning," in *Advances in Neural Information Processing Systems*, 2017, pp. 6306–6315.
30. Will Zou, Shenghuo Zhu, Kai Yu, and Andrew Y Ng, "Deep learning of invariant features via simulated fixations in video," in *Advances in neural information processing systems*, 2012, pp. 3203–3211.
31. Florian Schroff, Dmitry Kalenichenko, and James Philbin, "Facenet: A unified embedding for face recognition and clustering," *arXiv preprint arXiv:1503.03832*, 2015.
32. Leland McInnes, John Healy, and James Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:1802.03426*, 2018.
33. Thomas Blank and Marco Prinz, "Type i interferon pathway in cns homeostasis and neurological disorders," *Glia*, vol. 65, no. 9, pp. 1397–1406, 2017.
34. Jerome H Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.
35. Vincent A Traag, Ludo Waltman, and Nees Jan van Eck, "From louvain to leiden: guaranteeing well-connected communities," *Scientific reports*, vol. 9, no. 1, pp. 1–12, 2019.
36. Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention.* Springer, 2015, pp. 234–241.
37. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
38. Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
39. Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *Knowledge Discovery and Data Mining*, 1996, vol. 96, pp. 226–231.
40. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.
41. Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
42. Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L. Schmid, and Gaudenz Danuser, "Robust single-particle tracking in live-cell time-lapse sequences," *Nature Methods*, vol. 5, no. 8, pp. 695–702, Aug. 2008.
43. Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer, "Automatic differentiation in pytorch," in *NIPS-W*, 2017.
44. Laurens van der Maaten and Geoffrey Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.
45. Christopher S McGinnis, David M Patterson, Juliane Winkler, Daniel N Conrad, Marco Y Hein, Vasudha Srivastava, Jennifer L Hu, Lyndsay M Murrow, Jonathan S Weissman, Zena Werb, et al., "Multi-seq: sample multiplexing for single-cell rna sequencing using lipid-tagged indices," *Nature methods*, vol. 16, no. 7, pp. 619, 2019.
46. Christoph Hafemeister and Rahul Satija, "Normalization and variance stabilization of single-cell rna-seq data using regularized negative binomial regression," *Genome biology*, vol. 20, no. 1, pp. 1–15, 2019.
47. Diederik P. Kingma and Max Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.
48. Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.

## Methods

**QLIPP imaging.** We acquired all data using a Leica DMI-8 inverted widefield microscope, with 20x objective magnification at 0.55NA (air), and 0.4NA condensor on a Hamamatsu Flash-4 LT camera (6.5 um pixels). The cells were held at a constant 37°C, 5% CO2 using the Okolab stage-top incubator (H101-K-Frame). Each of the five polarization states were acquired, in sequence, with 50 ms camera exposure, for each of 5 z-planes for a given field of view. For the unperturbed microglia time series, we acquired 52 time points and 27 fields of view (9 per well) over 24 hours at 27 minute time intervals. For the perturbed microglia time series, we acquired 159 time points, 9 fields of view (4 were used for analysis), over 24 hours at 9 minute time intervals.

Birefringence is an optical property of matter that describes a different refractive index for different orientation axes of polarized light. The differential phase shift caused by these refractive indices, and the orientation axes, allow us to decompose birefringence into two properties: retardance and orientation (respectively). Each of these properties can be represented in terms of a combination of the Stokes parameters, which are Cartesian projections of a vector from spherical coordinates that describes the full polarization state of light. With the application of Mueller matrices that are tuned to our instrument setup, we can translate the Stokes representation into image intensities and vice versa. Therefore, given a set of intensity images of the specimen with known polarization states, we can use the inverse model to compute the sample's corresponding Stokes parameters and then physical properties of retardance and orientation. We estimate the background level of polarization with a 2D polynomial fit. Details for label-free hardware calibration, background correction and reconstruction can be found in previous work (16).

We used the open-source microscope control software Micro-Manager 1.4.22 (`https://micro-manager.org/`) for all image acquisition. The Micro-Manager plugin OpenPolScope (`https://openpolscope.org/`) performs the liquid-crystal-compensator (LCC) calibration by first finding voltages to achieve extinction ($I_{\text{Ext}}$ , intensity minimum). A pre-defined "Swing" voltage (0.03) is applied to induce slight polarization ellipticity along one axis ($I_{0\text{deg}}$). This "Swing" is empirically determined based on the sample, in order to produce maximum polarization contrast. Then, using the brent-optimizer minimization procedure, we find three more voltage states centered on $I_{0\text{deg}}$ that sample other orientations ($I_{45\text{deg}}$, $I_{90\text{deg}}$, $I_{135\text{deg}}$).

Following the reconstruction algorithm described previously (16) and documented on github (`https://github.com/mehta-lab/reconstruct-order`), phase reconstructions used the above hardware parameters plus the total variation regularizer with parameters: `rho=1`, `itr=50`, `absorption=1.0e-3`, `phase=1.0e-5`. Retardance reconstructions used default parameters plus "local fit" background correction, and retardance scaling of 1e4.

**Culture of primary microglia.** De-identified tissue samples were collected with previous patient consent in strict observance of the legal and institutional ethical regulations. Protocols were approved by the Human Gamete, Embryo, and Stem Cell Research Committee (institutional review board) at the University of California, San Francisco.

Primary human microglia were obtained from second trimester (gestational week 18-22) cortical brain tissue using magnetic-activated cell sorting. Tissue samples were dissected in artificial cerebrospinal fluid containing 125 mM NaCl, 2.5 mM KCl, 1mM $MgCl_2$, 1 mM $CaCl_2$, and 1.25 mM $NaH_2PO_4$. Tissue was cut into 1 $mm^3$ pieces, and the tissue was enzymatically digested using 0.25% trypsin (reconstituted from 2.5% trypsin, ThermoFisher 15090046) with addition of 0.5 mg/ml DNase (Sigma Aldrich, DN25) for 20 minutes at 37°C, mechanically dissociated and then passed through a 40 um mesh cell strainer (Corning 352340). The resulting cell suspension (100-150 million cells) was centrifuged for 5 minutes at $300 \times g$ and washed twice with $Ca^{2+}/Mg^{2+}$ - free phosphate buffered solution with addition of 0.5 mg/ml DNase to prevent cell clumping. Cells were re-suspended in 900 ul of MACS buffer (PBS with 0.5% BSA) with addition of 0.5 mg/ml DNase and incubated with 100 ul of the CD11b magnetic beads (Milteniy Biotec, 130-049-601) for 15 minutes following the manufacturer's instructions. After the magnetic beads incubation, cells were washed with 20 ml of PBS, span down at $300 \times g$, resuspended in 0.5 ml of MACS buffer and loaded on a MACS LS column (Milteniy Biotec, 130-042-401). Cells on the column were washed three time with 3 ml of MACS buffer, the column was removed from the magnet, and remaining microglia cells were eluted in 5 ml of microglia culture media. Microglia culture media (50 ml) consisted of: 33 ml of phenol red-free basal media Eagle's (BME), 12 ml Hank's buffered solution, 1 ml B27 (ThermoFisher, A3582801), 0.5 ml N2 (ThermoFisher, 17502048), 2 ml of 33% glucose, 0.5 ml GlutaMax (ThermoFisher, 35050061), and 0.5 ml penicillin/streptomycin (ThermoFisher, 15240062). Purified microglia cells were span down at $300 \times g$ and plated at $100 \times 10^3$ cells per well in microglia media supplemented with 100 ng/ml of rhIL34 (Peprotech, 200-34), 2 ng/ml TGFb2 (Peprotech, 100-35b) and 1x CD lipid concentrate (Lifetech, 11905031) to promote microglia cell survival in the monoculture. Prior to plating, glass-bottom 24 well plates (Cellvis, P24-1.5H-N) were coated with 0.1 mg/ml poly-d-lysine (Sigma-Aldrich, P7280) for 2 hours at room temperature followed by three double-distilled water washes and additionally incubated with laminin and fibronectin in PBS for 3 hours at 37°C. Cell culture media was changed twice a week, and was changed 24 hours prior to the start of the time lapse experiment to allow cells to re-equilibrate.

Time lapse imaging experiment was started on day 6 of culture and continued for 24 hrs in environmental control chamber (5%CO2, 37°C and relative humidity of 70%).

**dynamorph Pipeline.** *dynamorph* is composed of a collection of machine learning/deep learning tools that operate on imaging data and automatically generate morphodynamic summary of target instances. A brief pipeline is described in main text section "Automatic morphodynamic state discovery by *dynamorph*", here we elaborate technical details of each component in the same order. Note that two components of the pipeline required training in advance to state discovery applications: semantic segmentation model and self-supervised encoding model. In this work we dedicated a separate set of experimental data on control microglia (from different individuals) for model training.

***Semantic segmentation.*** A single three-class U-Net classifier (36) was deployed for the cell semantic segmentation, whose training data was derived from a combination of human annotations and machine learning (random forest) prediction. In the preparation of training data, we utilized interface from ilastik and manually annotated microglia, contaminating non-microglia cells (neurons, radial glia) and background/experiment artifacts in 30 frames selected from 13 fields of view. Manual annotation took in total approximately 8 hours, generating identifications for 1252 microglia and 312 non-microglia cells. The annotations were only partial, in which most of the background were not annotated (Supplementary Figure 2B). Most edges of the cell were not annotated or only roughly marked. Then the internal semantic segmentation model (random forest-based) of ilastik was trained and tuned to discriminate foreground (microglia and contaminating cells) and background, providing full segmentation mask of background/experimental artifacts. 50 frames of such background-only masks (foreground pixels were left unannotated) were visualized then inspected, serving as additional training inputs.

All training frames (2048 × 2048 pixels): human annotated frames and ilastik prediction frames, were combined and cropped into small patches (256 × 256 pixels). Two strategies were utilized: a frame could either be cut into patch tiles by a sliding window with no overlap, or by random sampling centers and rotation angles to extract patches from the context. The second strategy served partially as a data augmentation technique and was designed to increase the diversity of cell orientation and position. No further augmentation was performed as all frames were imaged with the same scale and brightness. Due to the sparsity of labels, we excluded patches with little or none annotations.

U-Net was adapted from `https://github.com/qubvel/segmentation_models`, and a pretrained backbone of resnet34 (37) was used to initialize weights. The model was trained end-to-end with standard cross entropy loss and Adam optimizer (38). During training, only partial labels corresponding to human annotations or random forest background predictions were used, weights elsewhere were set to 0. Weight scaling was also employed to balance loss for different classes. No hyperparameter search was performed due to limited amount of data and lack of validation metrics. Final model was evaluated visually based on criterions including cleanliness of cell edges, prediction consistency and accuracy on microglia/non-microglia/ambiguous cells. Note that *dynamorph* pipeline was robust against the output noise of segmentation as it only provided rough intermediate estimates of cell positions and shapes. We thus didn't validate and optimize this step exhaustively to ease the application-time usage.

Qualitatively, we found that ilastik, with a random forest core, performed worse in distinguishing cell types, often generating chimeric segmentation masks of both microglia and non-microglia cells. We speculated that this is due to the lack of global feature extractions in the method as cell type is highly dependent on the overall shapes and large-scale features such as protrusions. On the other hand, U-Net, when trained solely with human annotations, was not capable of delineating clean edges of cells, potentially due to class imbalance as well as lack of samples having clear edges annotated (annotators only marked the cell bodies). Prediction mask would hence frequently exceed the boundaries of a cell, posing difficulties to the following instance separation step. Based upon the above observations, we took a combination approach by augmenting the annotation step using an random forest classifier, generating a set of well-balanced "semi-annotated" training data with samples of clean foreground-background edges. The resulting U-Net generated pixel masks with less chimeric cells as well as clean boundaries. See comparisons in Supplementary Figure 3.

To generate segmentation for a full video, we first separated the video into static frames. For each frame, the full field of view (2048 × 2048 pixels) was divided into patches (256 × 256 pixels) following the same sliding window/tiling strategy, model was then applied on individual patches, results of which were tiled to generate the full prediction mask for this frame. To avoid edge effect, 20 repetitions with different offsets were performed, all 20 prediction masks were aligned and averaged to formulate the final prediction.

***Instance separation.*** We applied clustering on the pixels from segmentation masks to isolate and extract masks of each individual cell. In the procedure, all pixels predicted as foreground (microglia and non-microglia) by U-Net were extracted, then clustered based on their 2D coordinates in the frame. We used DBSCAN (39) to detect core points of each cell as well as to exclude outlier points coming from prediction artifacts. Note that this is feasible when cells are cultured in a relatively sparse environment. In fields with densely populated cells, clustering would not be able to separate boundaries between overlapping/contacting cells, instead more robust end-to-end instance segmentation models (40, 41) would have better results.

Implementation of DBSCAN from scikit-learn was applied, the parameters of which were tuned to separate adjacent cells with small amount of contact. In practice, since most of the prediction masks were clean, we searched the two main parameters (`eps` and `min_samples`) to maximize detected number of cells whose size were above a minimum threshold. The final model employed a maximum neighbor distance (`eps`) of 10 pixels (3.25 um) and a min neighborhood size (`min_samples`) of 250 for core points. Each set of core points and their reachable points were regarded as the mask for an individual cell. We further applied a post-processing step to exclude masks with extreme sizes (>12000 pixels or <500 pixels), serving as a filter for prediction/experimental artifact and overlapping/contacting cells that cannot be separated.

The identity of each cell was then determined based on the average classification score over its mask. We simply averaged the pixel-wise probabilities of both classes (microglia or non-microglia) across the whole cell mask. The resulting ratio of probabilities represented the proportion/likelihood the U-Net regards this specific cell as microglia or non-microglia. Cells with microglia proportions larger than 0.9 were regarded as true microglia, and vice versa for non-microglia. The rest would be regarded as ambiguous cells. Only microglia from training set were used for the self-supervised model training described below. These identities also helped in determining validity of trajectories in the following tracking procedure.

***Tracking.*** Within a full video, instance separation was performed on each static frame independently, results of which were used in this step to connect cells across time and form trajectories. The tracking procedure followed the work of (42), in which the core architecture is a linear assignment problem (LAP). In every two adjacent frames, each frame would have a list of cells generated from the instance separation step, along with their sizes and center positions calculated on the segmentation masks. Then a matching problem was set up between the two lists of cells, with a pairwise cost matrix calculated based on position displacements and shape differences between pairs of cells:

$$cost_{ij} = (r_{ij}^{shape} + \frac{1}{r_{ij}^{shape}}) \times {d_{ij}}^2$$

$$r_{ij}^{shape} = \frac{s_i}{s_j}$$

$d_{ij}$ : distance between cell $i$ and cell $j$ from two adjacent frames

$s_i$ : size of cell $i$

The cost favored cell pairs that were spatially close and morphologically (size) similar. In practice, we also enforced thresholds (maximum displacement 100 pixels, maximum size difference 2 folds) to refine results. Resulting matching pairs along all frames in a video were connected, formulating a set of single cell trajectories that span multiple frames. We further adopted the gap closing strategy in (42) to connect short trajectories that were separated due to imaging/segmentation issues. A similar LAP set up with only squared distance cost was employed.

The set of trajectories obtained through the above procedure would contain both microglia and non-microglia cells. To select only microglia trajectories, the identities of cells in static frames (proportions of microglia segmentation provided by UNet) calculated in the previous step were pooled and filtered. Only trajectories with over 95% of all frames being labeled as microglia (proportion $> 0.9$) were kept. The final collection of cell trajectories would be used to guide concatenation of morphology descriptors as well as to calculate motion descriptors in the downstream clustering analysis. In total, 5731 microglia trajectories were collected, 3715 of which were extracted from the test dataset.

***Self-supervised encoding with VQ-VAE.*** A variational autoencoder variant: VQ-VAE (29) was applied in this work to summarize morphology of cells through a encoding-decoding process. The model was trained on microglia cell patches extracted from the training set and directly applied to the test set data to generate morphological descriptors for cells under different treatment.

We took all individual microglia cells in static frames from the instance separation step and cropped patches of size $256 \times 256$ pixels around the cell centers. If the cell appeared on the border of a field of view, the out-of-border area would be filled with median background values of phase and retardance. Meanwhile, since we focused on analysis of single microglia, we further masked out other cells appearing in the patch area and filled with median background values. Sample patches before the masking can be found in Figure 2, samples after masking can be found in Supplementary Figure 4 and Supplementary Figure 5.

VQ-VAE was implemented in pytorch (43), using two deep convolutional networks as encoder and decoder respectively. Input patches were first down-scaled to $128 \times 128$ and normalized on both channels, then VAE encoder further down-sampled the image by 64 folds. Resulting latent vector for each cell had shape of 16, in which the last dimension was channel. Then consistent with the standard setup of VQ-VAE, we regularized latent space by forcing a discretization step on the channel dimension, in which each vector at a given position out of the $16 \times 16$ grid was matched with 64 embedding vectors and the closest embedding was inserted back to the position and passed to the decoder. A matching loss was enforced in the later phase of training that minimizes frame-to-frame differences between latent vectors of the same cell along its trajectory. A combination of these two losses was used to train the model in an end-to-end fashion.

$$\begin{aligned} L_{\text{VQ-VAE}}(x) = \log p(x|z_q(x)) + \\ \|\text{sg}[z_e(x)] - e_k\|_2^2 + \\ \beta \, \|z_e(x) - \text{sg}[e_k]\|_2^2, \\ \text{where} \quad k = \text{argmin}_j \|z_e(x) - e_j\|_2^2, \\ z_q(x) = e_k \end{aligned}$$

$$\begin{aligned} L_{\text{matching}}(x^{(t)}) = \sum_{i=1,-1} \left\| z_e(x^{(t)}) - z_e(x^{(t+i)}) \right\|_2^2 + \\ 0.1 \sum_{i=0}^{l} \left\| z_e(x^{(t)}) - z_e(x^{(i)}) \right\|_2^2 \end{aligned}$$

$x$ : input cell image
$z_e(x)$ : encoder output
$z_q(x)$ : decoder input
$e$ : closest embedding vector to $z_e(x)$
sg : stop gradient operation
$x^{(t)}$ : $t$-th frame of a cell trajectory
$l$ : length of the cell trajectory

where $\beta = 0.25$ controls the amount of commitment loss regularized during training. In the training procedure, we applied only $L_{\text{VQ-VAE}}$ in the initial phases. When model achieved a loss plateau and relatively high perplexity (indicating how many of the 64 embedding vectors were used), we added in $L_{\text{matching}}$ gradually till a final ratio of 1:0.005 for VQ-VAE loss/matching loss.

Size of the latent space used in this study was consistent with the compression fold in its original work (29). Note that we didn't push further to a much smaller latent space (i.e. <100D) due to the concern that smaller latent vector will consist of more conceptualized and higher-level representations of images, which would cast difficulties in interpretation and quantification.

Note that matching loss was only applied within mini-batch to avoid extensive calculation. During training we enforced a sample order that aligns most of the neighboring/same-trajectory cell pairs sequentially, so with a high chance a cell patch's next or previous frame would exist in the same mini-batch.

***PCA of latent vectors.*** PCA was fitted on $z_e(x)$ for all microglia cells in the training set, and applied to all training and test cells. Implementation from scikit-learn was used, and we extracted PCs that explain the top 50% of all variances: 48 components were extracted, explained variance of which is visualized in Supplementary Figure 10A.

We also tried fitting and transforming $z_q(x)$ (data not shown), which generated very similar results and same interpretations for the top PCs. Distance metrics defined with either $z_e(x)$ or $z_q(x)$ showed similar results (Supplementary Figure 11) as well. But since $z_e(x)$ contained more information before the quantization, we employed it throughout this work.

Other dimensionality reduction methods including tSNE (44) and UMAP (32) were also tested and showed no advantage over PCA as no clustering patterns was detected.

***Clustering of trajectory feature vectors.*** In this work, to further unravel the morphodynamic features, as well as to link transcriptomic profiles of microglia, a downstream clustering was performed on cell trajectories.

Feature vector of each cell was composed of morphology descriptor and motion descriptor: top 48 PCs derived above were used as morphology descriptor; position displacements along the trajectory were averaged, and log mean displacement was used as motion descriptor. The log mean displacement values were scaled to keep variance comparable with the top 2 PCs.

Then a Gaussian Mixture Model (GMM) with 2 components was applied to model different sub-populations of microglia. Since input data were from different treatment group, we assumed different mixture weights for different subset (treatment condition). Given the feature vector $x_i$ and treatment group $y_i$ of a cell, it was assumed to be sampled from:

$$p(x_i, y_i|\Theta) = \sum_{k=1}^{2} \alpha_k^{(y_i)} p_k(x_i|\theta_k)$$

where $\Theta = \{\alpha_1^{(1)}, \alpha_2^{(1)} ..., \alpha_1^{(4)}, \alpha_2^{(4)}, \theta_1, \theta_2\}$ were parameters of the mixture model: $\theta_k$ is the center of $k$-th component, $\alpha_k^{(j)}$ is the weight of $k$-th component in subset $j$. In the test set 4 conditions were used so 4 sets of component weights were fitted with samples from the corresponding groups.

Expectation-Maximization (EM) algorithm was employed to iteratively evaluate the parameters of GMM. E-step evaluates the membership weight of each cell based on the current set of parameters:

$$w_{ik} = \frac{\alpha_k^{(y_i)} p_k(x_i|\theta_k)}{\sum_{k=1}^{2} \alpha_k^{(y_i)} p_k(x_i|\theta_k)}$$

$$\text{where} \quad p_k(x_i|\theta_k) \propto e^{-\frac{1}{2}(\frac{x_i - \theta_k}{s})^2}$$

M-step recalculates the parameters based on new membership weights:

$$\alpha_k^{(j)} = \frac{\sum_{i \in \{i|y_i=j\}} w_{ik}}{\sum_{i \in \{i|y_i=j\}} 1}$$

$$\theta_k = \frac{\sum_i w_{ik} x_i}{\sum_i w_{ik}}$$

Two steps above were repeated till parameters converged. In the procedure we fixed the variance of mixture components to be the same as variance of feature vectors to increase stability. GMM with more than 2 components were tested but interpretability of each component became worse so we limit to the two-state model of microglia.

**scRNA Sequencing.** Cells from three conditions (control, IFN beta and GBM supernatant) were dissociated using 0.25% trypsin solution, labeled with Multiseq barcodes (45) and processed for single cell mRNA sequencing using the 10x Genomics Chromium v3 Platform. CellRanger version 3 was used to generate the count matrix of cells by genes. Cells with more than 20% mitochondrial abundance or less than 500 UMI were removed. Doublets were removed during the demultiplexing of Multiseq barcodes.

SCTransform (46) was applied to the raw count matrix followed by PCA on the residuals. The top 30 principal components were used to find neighbors for Leiden clustering (35) and the UMAP (32) projection.

## Code Availability

Open source python software for reconstruction of label-free optical properties is available at `https://github.`

`com/mehta-lab/reconstruct-order` and for analyzing cell states is available at `https://github.com/czbiohub/dynamorph`.

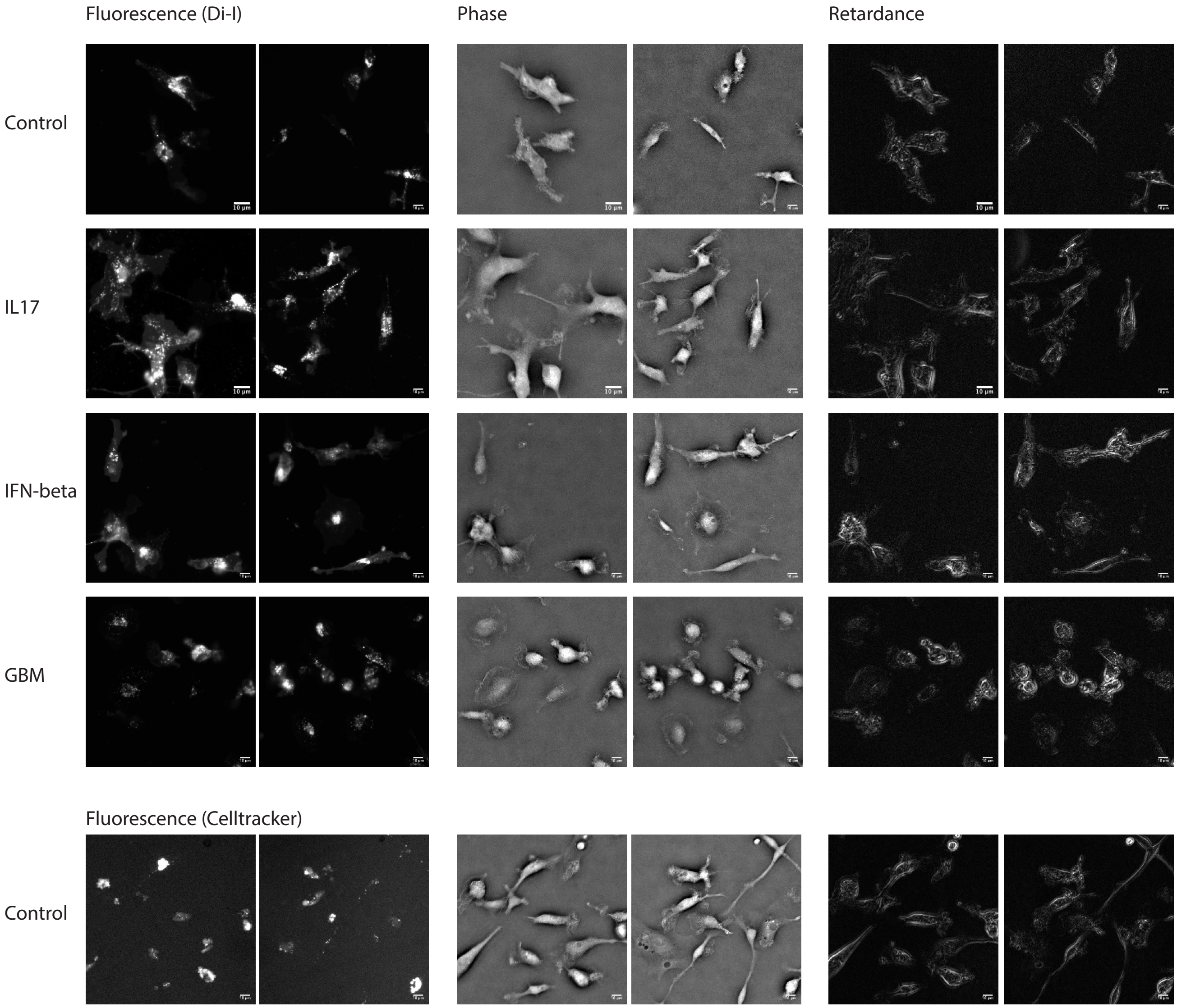
## Acknowledgements

We thank Syuan-Ming Guo for discussions around machine learning and dimensionality reduction. We thank Greg Huber for discussions around the physical models of cell states. J.Z. is supported by NSF CCF 1763191, NIH R21 MD012867-01, NIH P30AG059307, NIH U01MH098953, and grants from the Silicon Valley Foundation and the Chan-Zuckerberg Initiative. B. B. C., L-H. Y., and S.B.M. are supported by the Chan-Zuckerberg Biohub's intramural program. This collaboration was supported by the Chan-Zuckerberg Biohub's intercampus award to T. N. and J. Z.

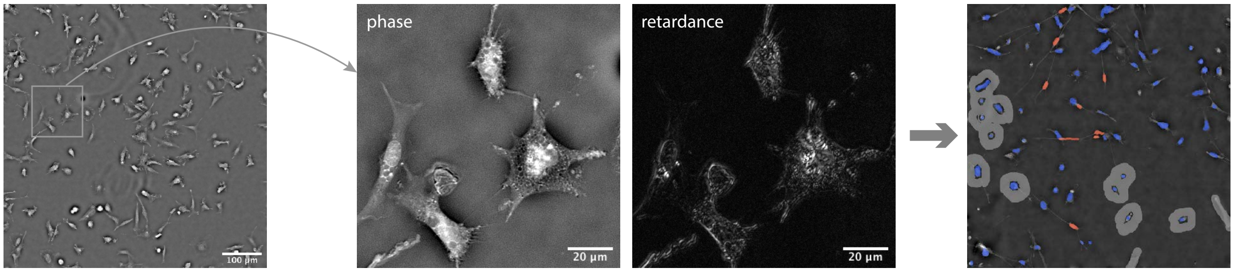
## Competing Interests Statement

Authors declare no competing interests.

## Supplementary figures

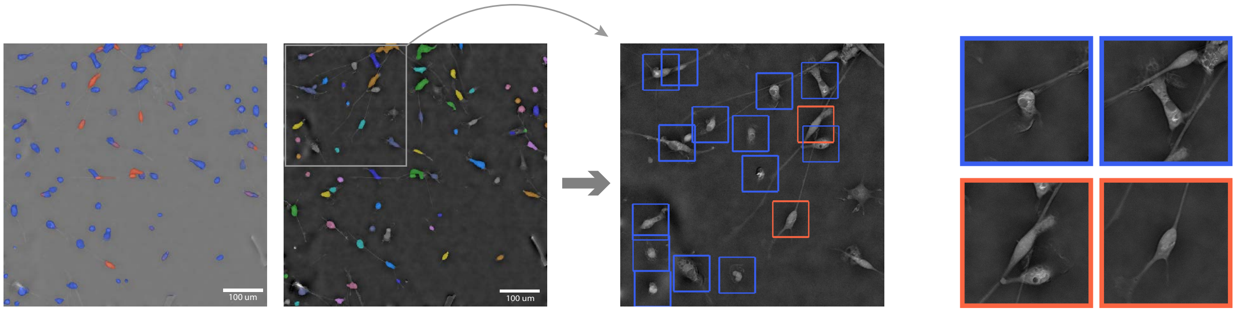


**Supplementary Figure 1. Comparison of microglia morphology as visualized with vital dyes (Di-I and CellTracker CM-DiI) and label-free measurements (phase and retardance):** Two examples from each condition are shown. Di-I and CellTracker are a lipophilic carbocyanine derivatives that internalize through different protocols, but should penetrate the membrane and persist in the cytoplasm enough to outline cell contours. Both dyes under all perturbation conditions show similar internalization and sequestration of the dyes. All scale bars are 10 um.

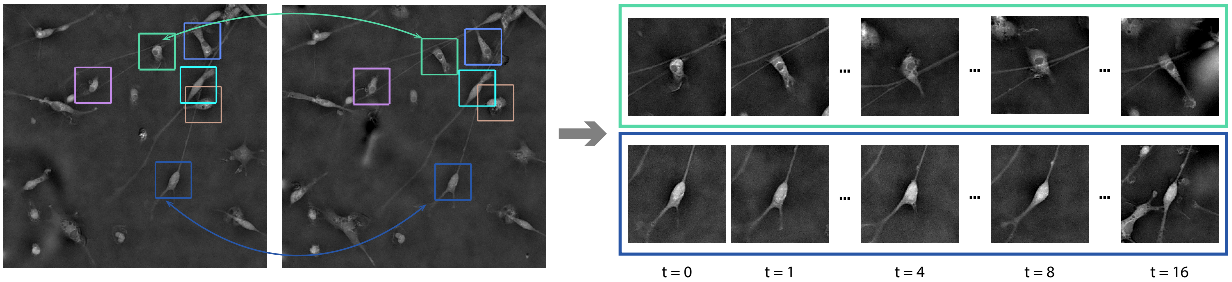


**A.** Acquire and reconstruct phase, retardance channels

**B.** Annotations



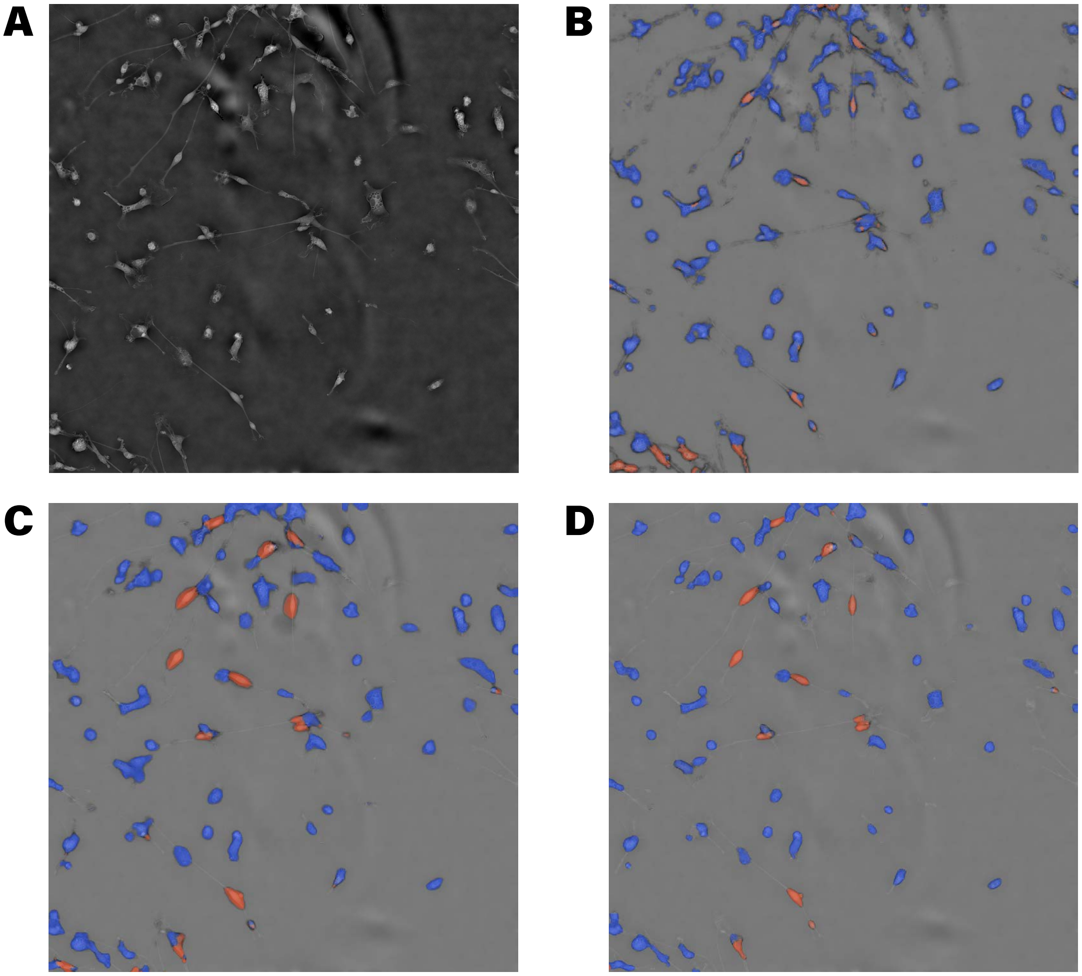
**C.** Semantic segmentation **D.** Instance segmentation **E.** Microglia (blue) and non-microglia (red) marked in boxes



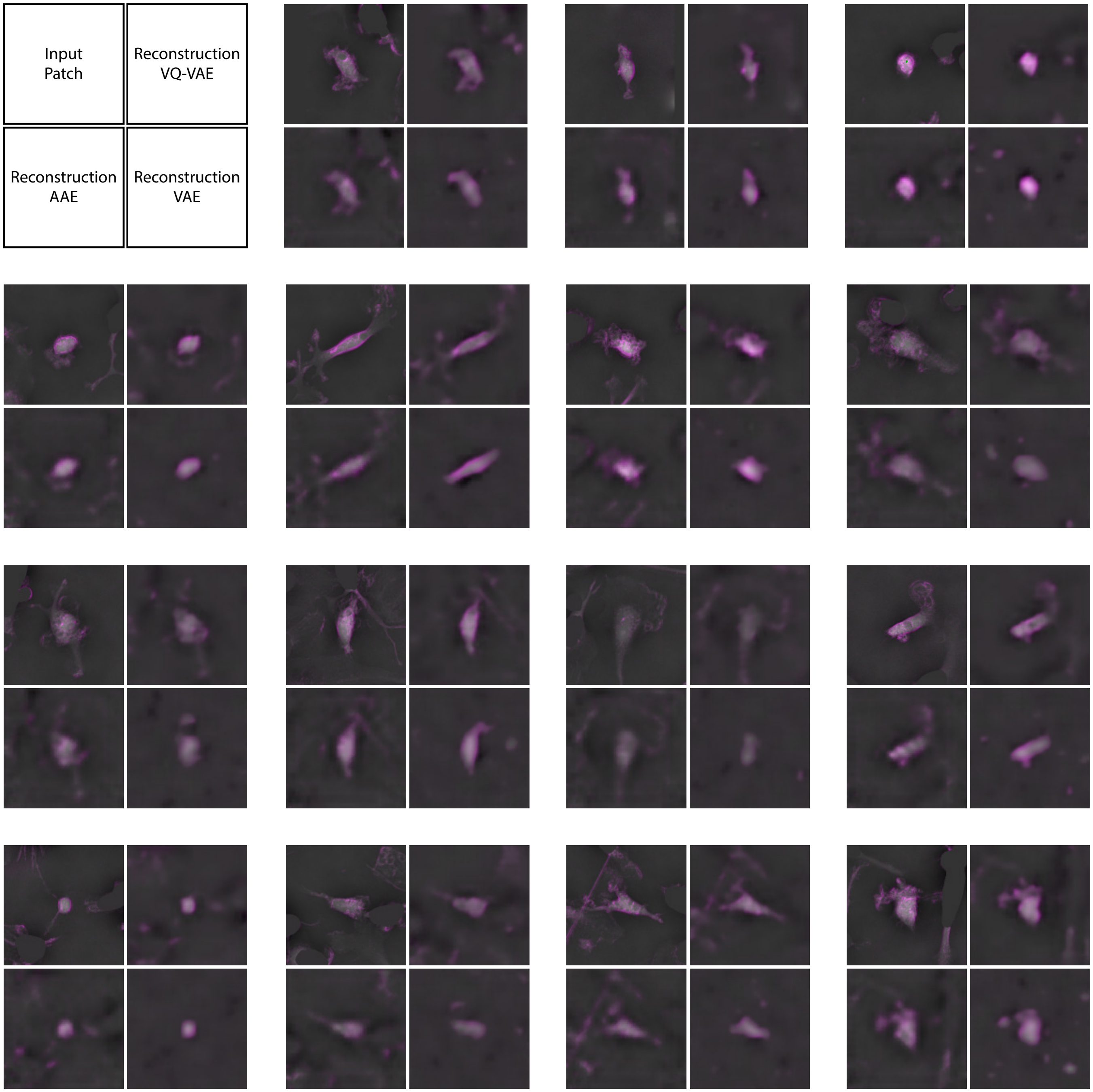
**F.** Frame matching: (left) t = 0, (right) t = 1

**G.** Single cell trajectories

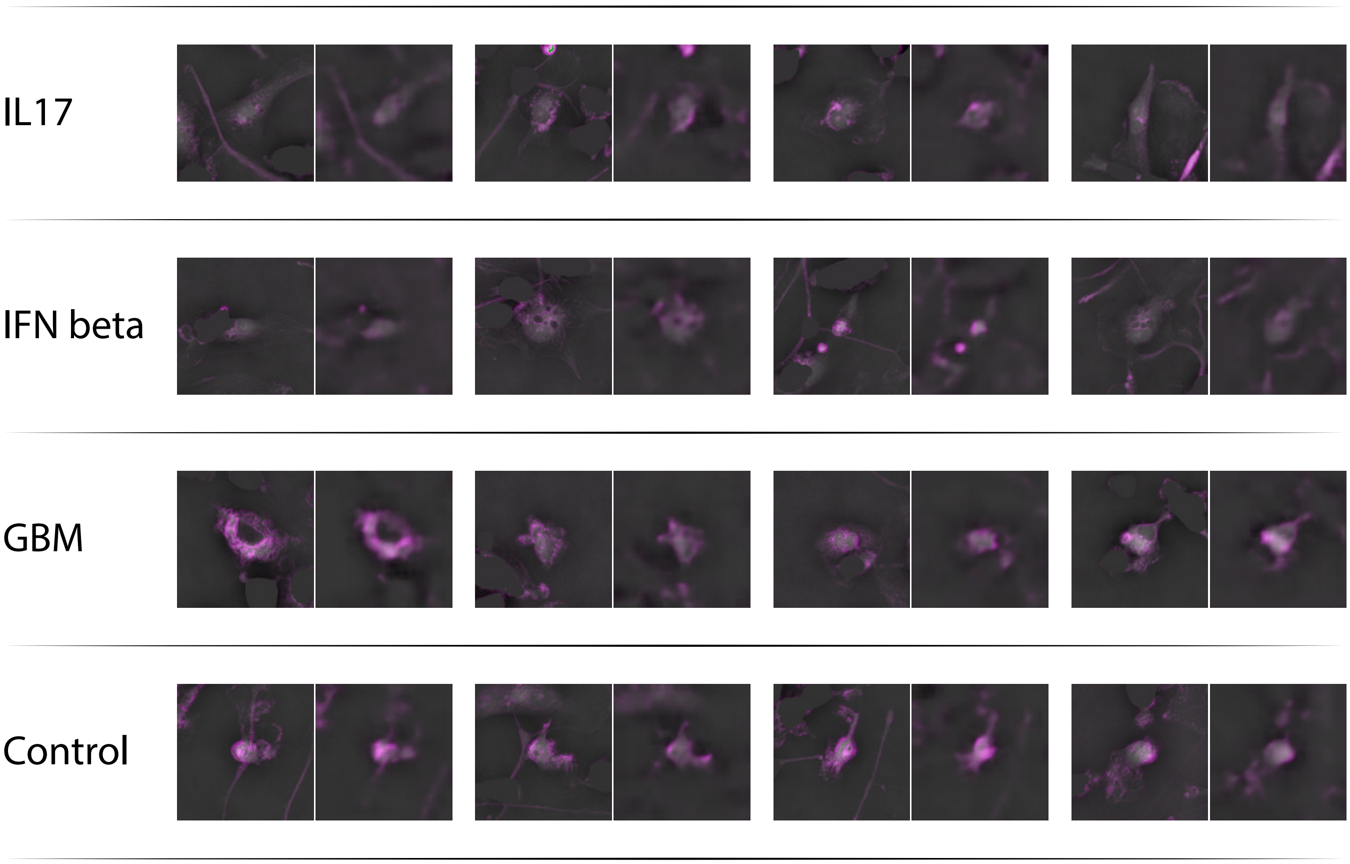
**Supplementary Figure 2. Identification and tracking of microglia with sparse supervision: (A)** Cell morphology and dynamics are visualized without label based on density (phase) and structural anisotropy (retardance). **(B, C)** Sparse human annotations enable microglia segmentation using a combination of conventional machine learning (ilastik/random forest) to separate foreground and background, and deep convolutional networks (U-Net) to separate microglia and contaminating non-microglia cells. **(D, E)** Segmentation masks are further clustered through DBSCAN to form individual cell masks, from which the position and surrounding patch of the cell are extracted. In the encoding pipeline we extracted boxes of $83\mu m \times 83\mu m$ around the cell centers for the patches. Note that stringent filtering was applied on the cell masks to exclude potential artifact and overlapping/contacting cells, leaving only masks/bounding boxes of individual cells. **(F)** Individual cells in adjacent time-points are linked through solving a linear assignment problem. **(G)** Connected components through multiple frames form single cell trajectories. Representative trajectories for microglia (top) and progenitor cells (bottom) are shown.



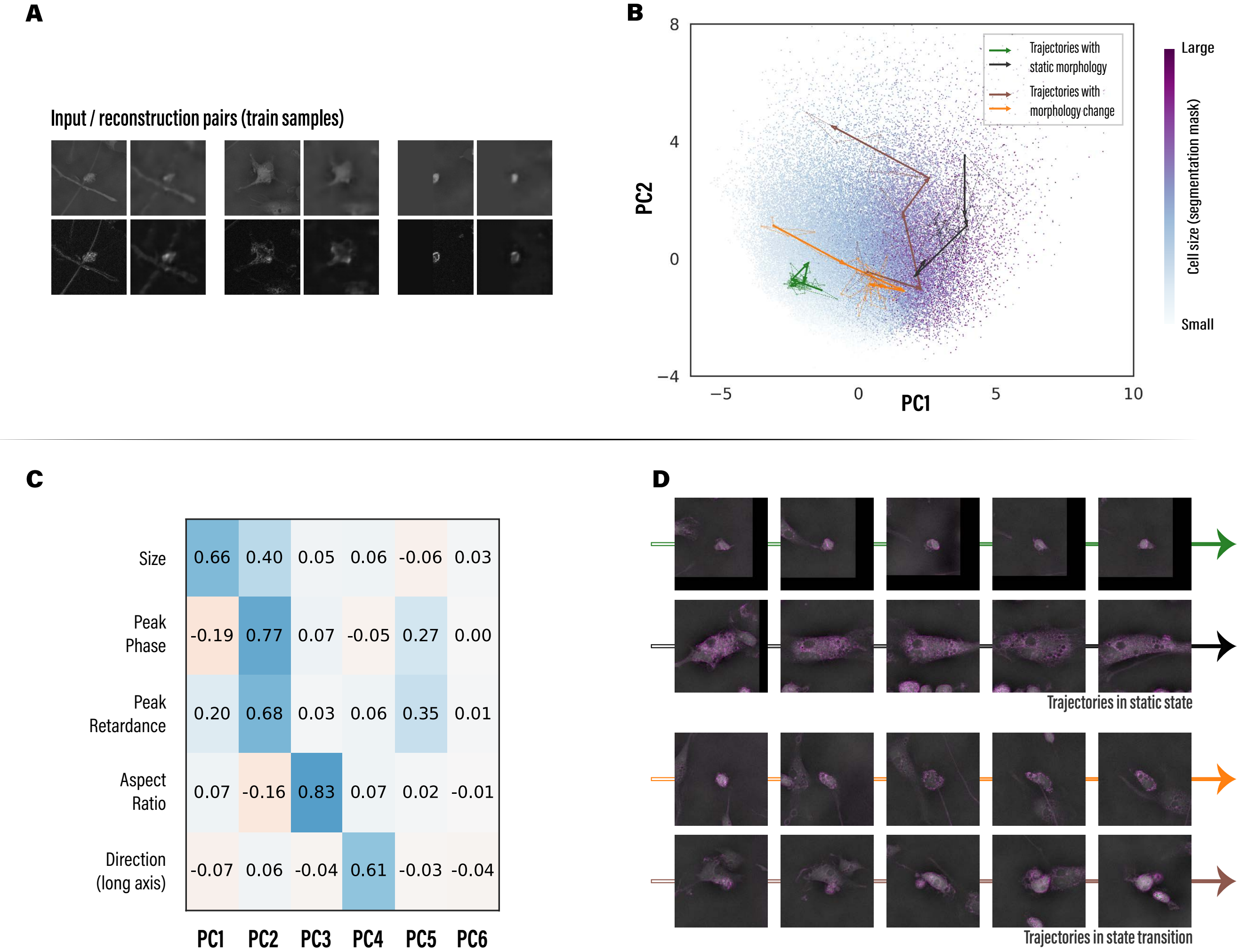
**Supplementary Figure 3. Comparison of different segmentation results: (A)** Raw input of an unannotated slice, phase channel. **(B)** Prediction map from a random forest classifier (ilastik) trained on human annotations. Note that predictions are more noisy and chimeric. **(C)** Prediction map from a U-Net trained on human annotations. Note that most edges exceed the boundaries of cells. **(D)** Prediction map from a U-Net trained on human annotations and random forest generated background annotations.



**Supplementary Figure 4. Sample cell patches (phase channel) randomly selected from the test set and reconstructions from multiple VAE variants:** Reconstruction loss for variational autoencoder (47) (VAE): $0.25 \pm 0.06$, adversarial autoencoder (48) (AAE): $0.19 \pm 0.07$, vector quantised-variational autoencoder (29) (VQ-VAE): $0.18 \pm 0.07$. All variants have latent space of the same size, VQ-VAE is also regularized by the discrete prior. Order of the reconstructions is explained in the upper left corner. Vanilla VAE lost more details than AAE and VQ-VAE, VQ-VAE has overall best reconstruction and image compression fold.

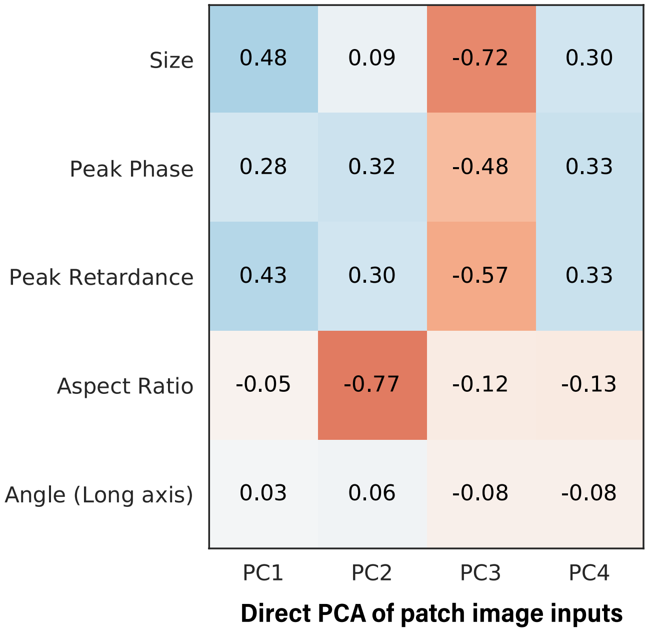


**Supplementary Figure 5. Illustrative cell patches of cells imaged under different perturbations (IL17, IFN beta, GBM, Control) and their reconstructions by VQ-VAE.**

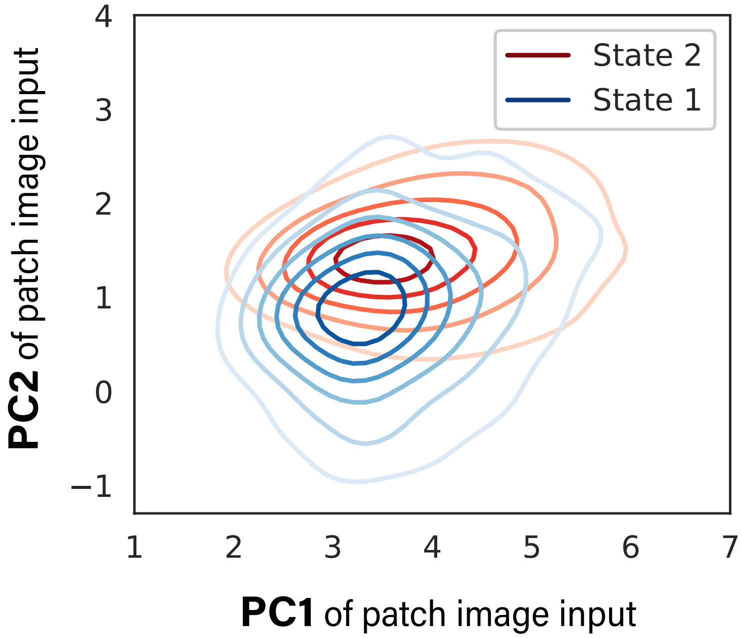


|  | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 |
|---|---|---|---|---|---|---|
| Size | 0.66 | 0.40 | 0.05 | 0.06 | -0.06 | 0.03 |
| Peak Phase | -0.19 | 0.77 | 0.07 | -0.05 | 0.27 | 0.00 |
| Peak Retardance | 0.20 | 0.68 | 0.03 | 0.06 | 0.35 | 0.01 |
| Aspect Ratio | 0.07 | -0.16 | 0.83 | 0.07 | 0.02 | -0.01 |
| Direction (long axis) | -0.07 | 0.06 | -0.04 | 0.61 | -0.03 | -0.04 |

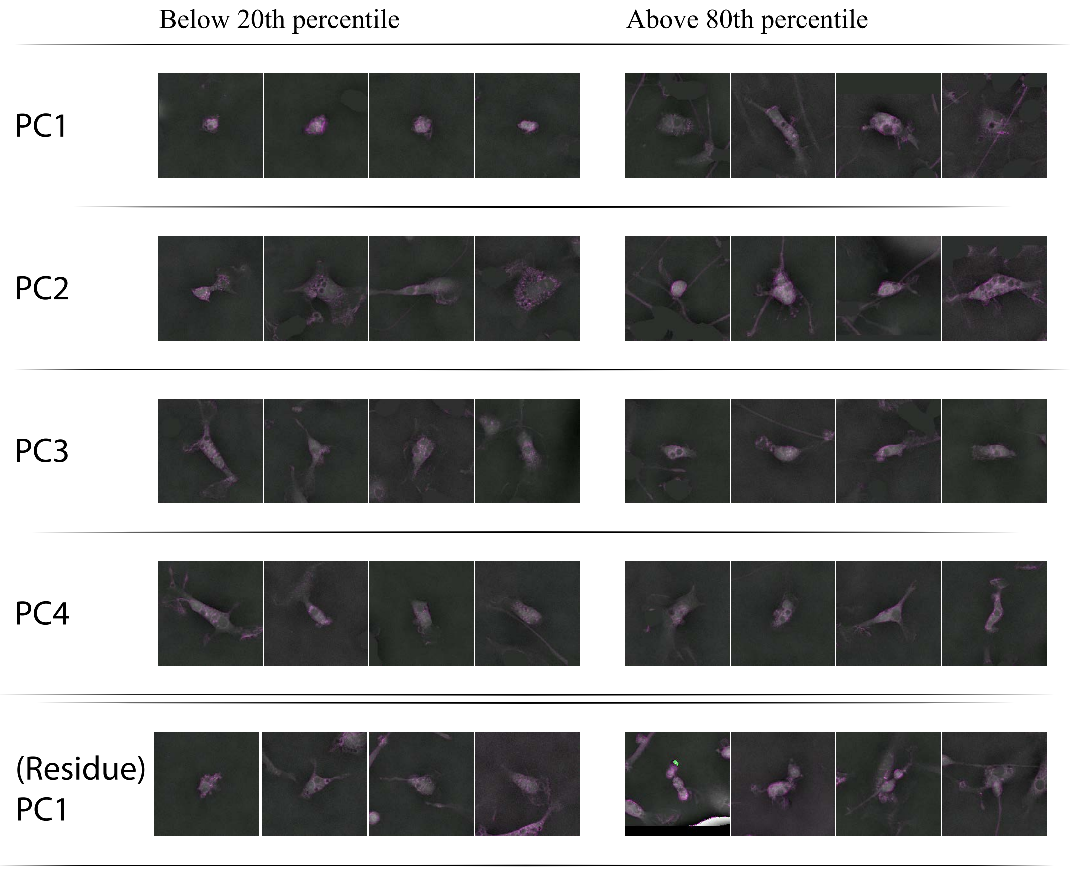
**Supplementary Figure 6. Reconstruction and encodings of samples from the training set: (A)** Example patches from the training set and their reconstructions by VQ-VAE **(B)** Top 2 PCs of all samples from the training set, 4 representative trajectories are plotted. **(C)** Spearman's rank correlation coefficients calculated between top PCs and geometric properties. Note that results are highly similar to test set samples. **(D)** Representative trajectories are visualized.



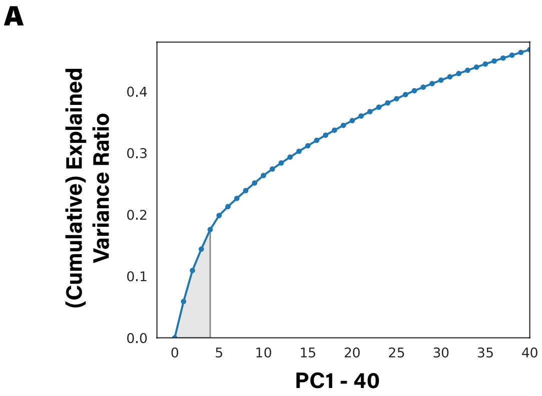
**Supplementary Figure 7. PCA was directly fitted on input images without encoding** from the training dataset and applied to test samples. Resulting PCs were analyzed for correlations to geometric properties. Spearman's rank correlation coefficients were calculated between the same set of geometric properties and top image PCs. Correlations are weaker and more interrelated and thus harder to separate and interpret.



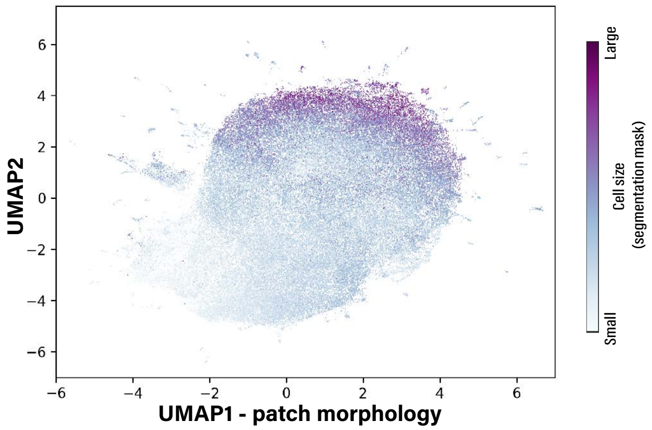
**Supplementary Figure 8. PCs directly calculated on image inputs are used for clustering following the same GMM procedure.** Two GMM states were defined using motion descriptor (speed) and image PCs. KDE plot of trajectories from the two states is illustrated, note that state separation is much worse than GMM states defined with latent vector PCs.



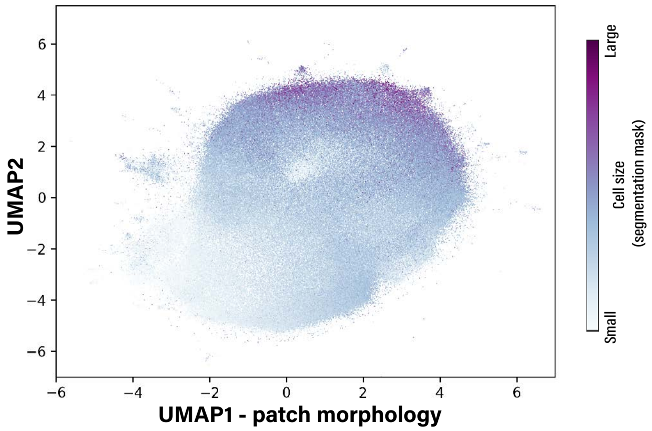
**Supplementary Figure 9. Cell patches randomly sampled from the upper and lower 20 percentiles along each principal component axis (PC1 to 4, residue PC1).** Patches are visualized as combination of phase channel (as grey scale) and retardance channel (as magenta shades). Samples along PC1 show different cell sizes; samples along PC2 show different phase strengths; samples along PC3 and PC4 (controlled for PC1 and PC2 values) show different orientations. We further calculated a residue latent space after excluding all known geometric properties of cells (listed in Figure 2D), samples along the first principal component of the residue space (rPC1) are illustrated in the last row: samples with lower rPC1 values tend to involve less contact/interaction, while higher rPC1 value correlate with complex environments including overlapping cells, cell on the border of imaging view, imaging artifact, etc.



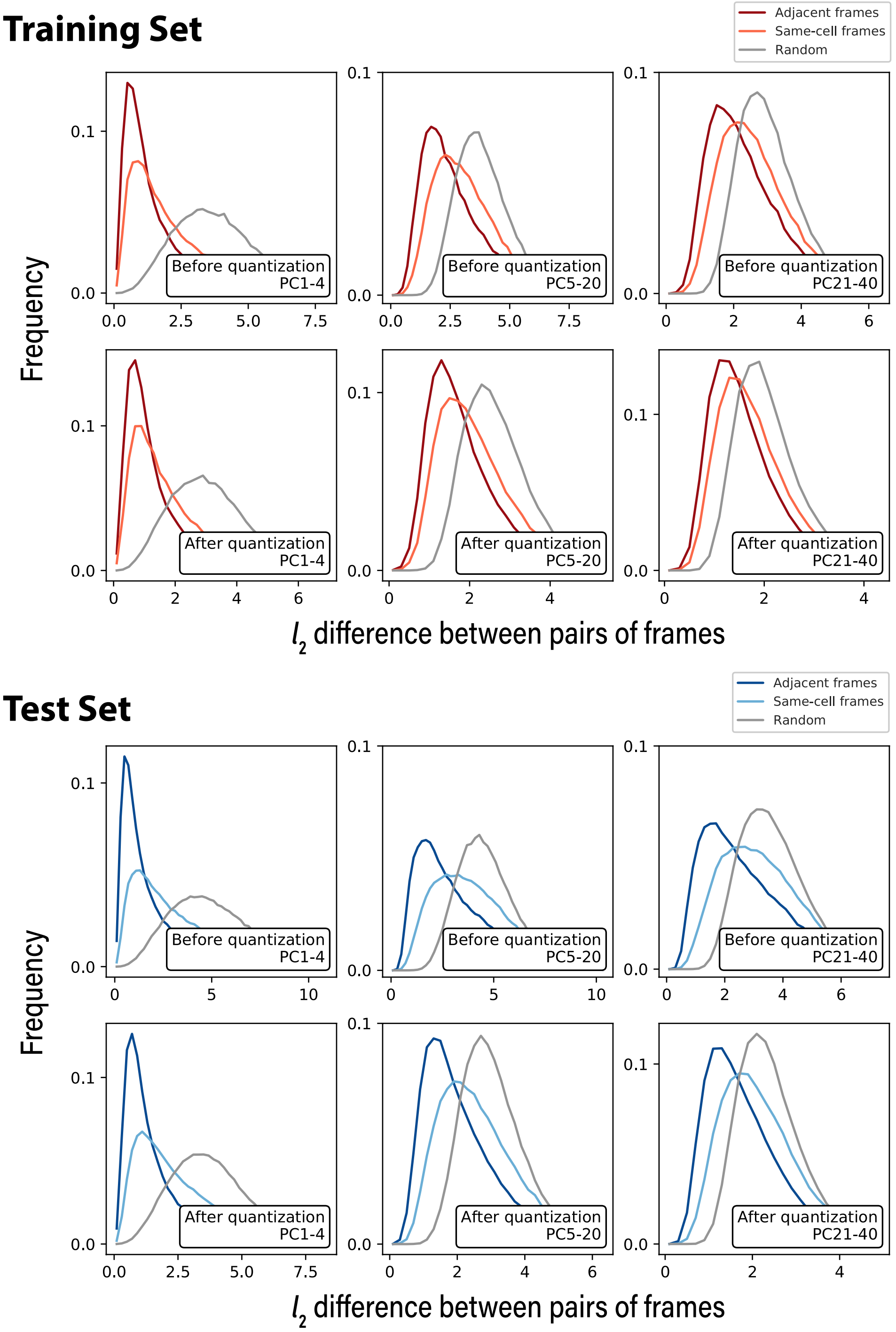
**B** - UMAP reduction of training set



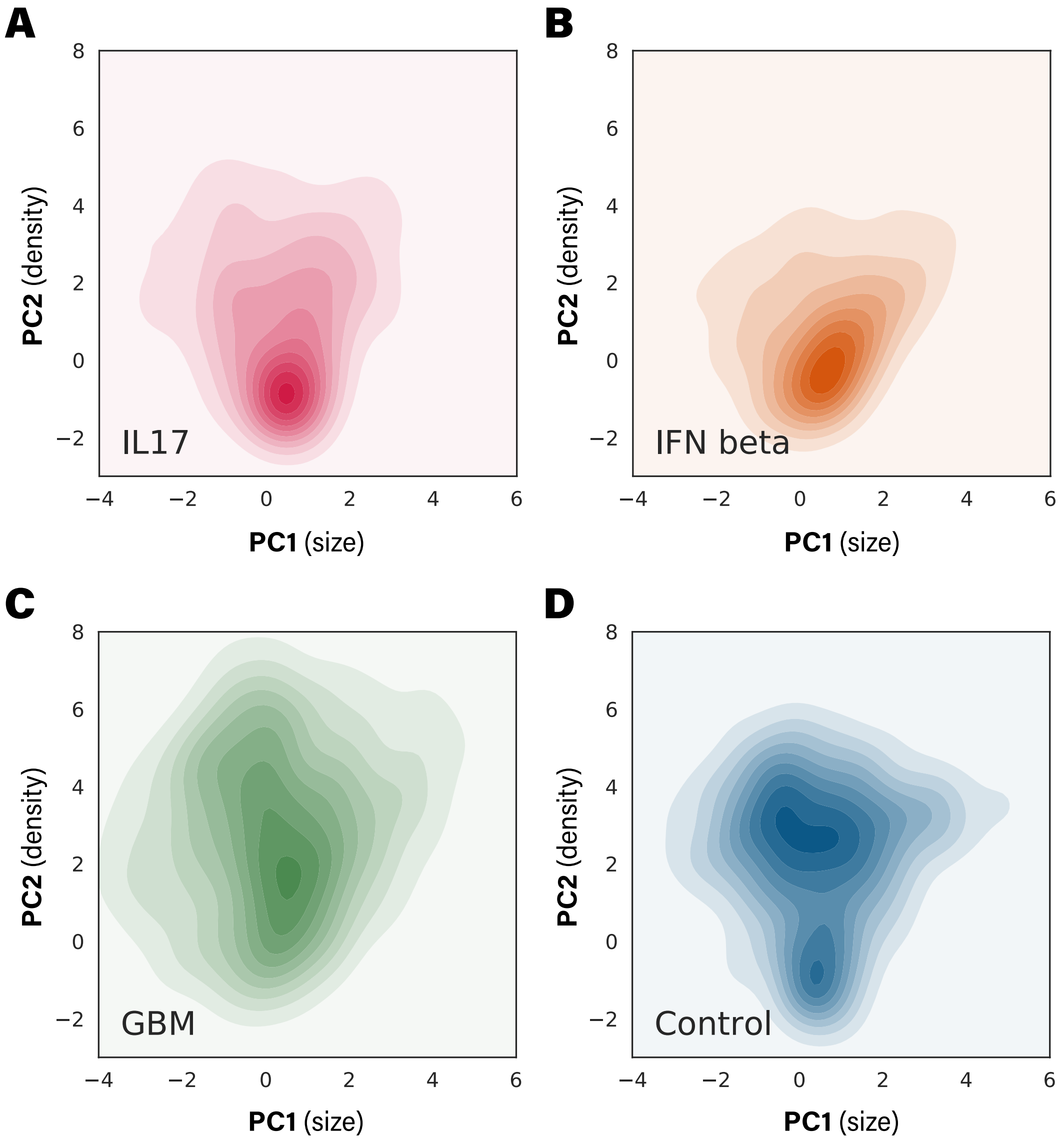
**C** - UMAP reduction of test set



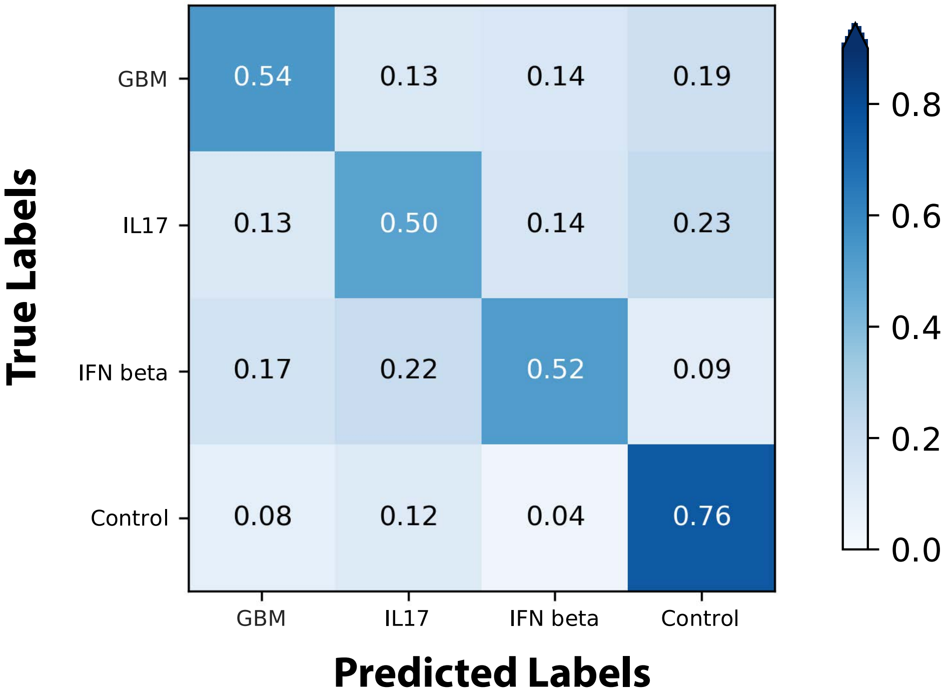
**Supplementary Figure 10. Latent space of morphology is high-dimensional: (A)** Explained variance ratios of the top 40 principal components of the morphology space. Note that analysis in this work focuses only on the first 4 principal components. (shaded area, explained ∼17% of all variance). **(B)** and **(C)** UMAP reduction (2 components) on sample latent vectors from the training set and the test set. No clustering patterns are found after tuning parameters, indicating that the space is highly continuous.



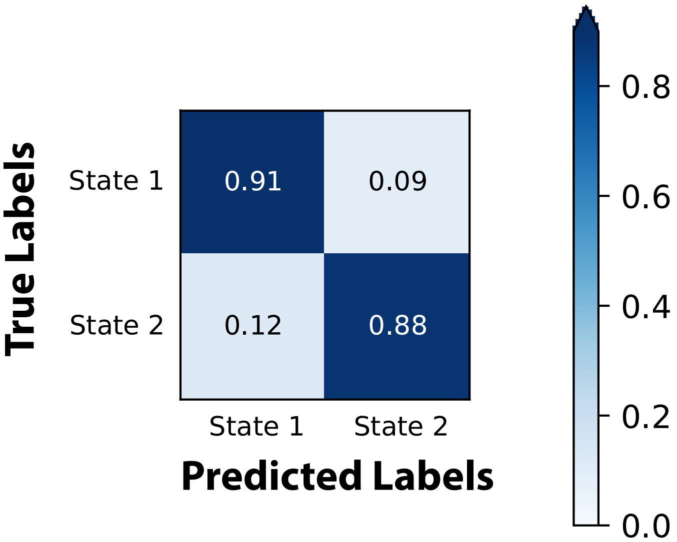
**Supplementary Figure 11. Distributions of latent vector differences between pairs of frames:** On both training set (upper panels) and test set (lower panels), we calculated differences between pairs of frames for three conditions: adjacent frames in trajectories (deep red/blue), non-adjacent pair of frames from the same trajectory (light red/blue) and random pairs of frames selected from all static patches (grey). In all settings distances in the latent space match well with our prior of shape similarities (adjacent > same-cell > random pair). Quantization step in the VQ-VAE doesn't interfere with the distance metric. Also note that the trends are much more significant in top PCs (PC1-4).



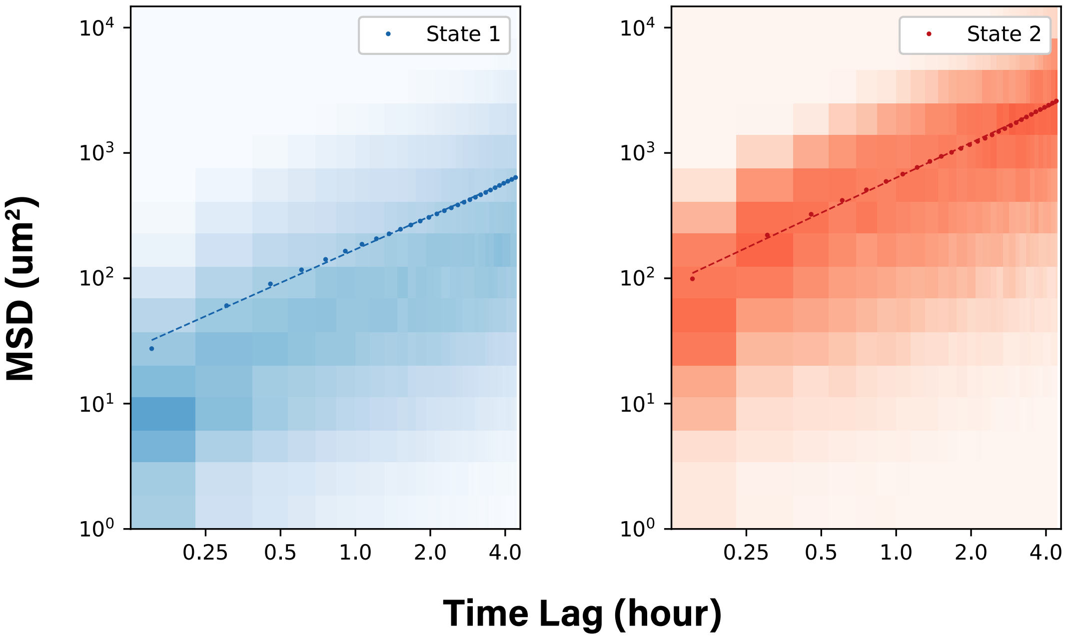
**Supplementary Figure 12. Kernel density estimate (KDE) plots of trajectory-averaged PC1 and PC2 from the test dataset.**



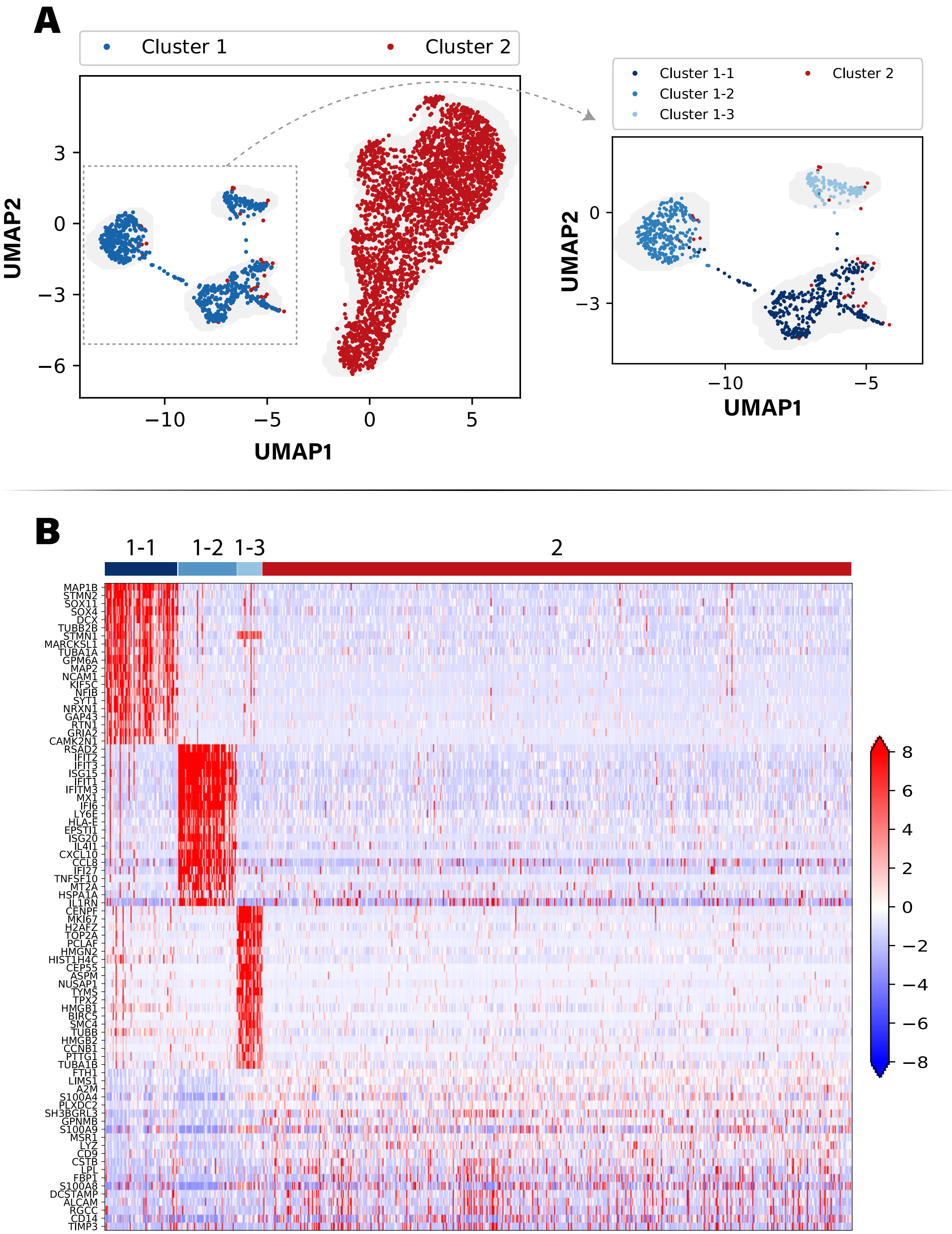
**Supplementary Figure 13. Confusion matrix of classifiers that predict condition/subset from trajectory feature vectors.** We employed gradient boosted trees with default hyper-parameters as the predictor models and trained them under 10-fold cross validation.



**Supplementary Figure 14. Confusion matrix of classifiers that predict GMM state from trajectory feature vectors.** We employed gradient boosted trees with default hyper-parameters as the predictor models and trained them under 10-fold cross validation.



**Supplementary Figure 15. Density map of squared cumulative displacements calculated over trajectories from the two GMM states.** Mean square displacements (MSD) are plotted, log-log fit results of the MSD could be found in Figure 4E



**Supplementary Figure 16. Details of scRNA clustering: (A)** UMAP projection and clustering of microglia cells based on scRNA-seq. Note that cells from all three conditions are included. Note that cluster 1 is further separated into 3 sub-clusters by applying the same method, in total identifying four molecularly distinct clusters. **(B)** Heatmap of the top differentially expressed genes in each cluster. Cluster 2 includes primarily control microglia and microglia treated with GBM supernatant. Cluster 1-1 is characterized by non-neuronal genes. Cluster 1-2 has high expression of interferon response genes. Cluster 1-3 is defined by high expression of cell cycle genes consistent with a dividing population.

## Supplementary tables

| **Latent vector PCs** | **State-1** | **State-2** | **Diff.** |
|---|---|---|---|
| PC1 | 0.44 | 0.17 | 0.27 (0.17SD) |
| PC2 | 0.66 | 2.80 | 2.14 (1.04SD) |
| PC3 | -0.07 | 0.08 | 0.15 (0.16SD) |
| PC4 | -0.13 | 0.00 | 0.13 (0.14SD) |
| Log speed | 3.25 | 4.03 | 0.78 (1.11SD) |
| Proportion (IL-17) | 0.79 | 0.21 | |
| Proportion (IFN beta) | 0.93 | 0.07 | |
| Proportion (GBM) | 0.42 | 0.58 | |
| Proportion (Control) | 0.25 | 0.75 | |

| **Image direct PCs** | **State-1** | **State-2** | **Diff.** |
|---|---|---|---|
| PC1 | 3.68 | 3.98 | 0.30 (0.22SD) |
| PC2 | 0.92 | 1.41 | 0.49 (0.52SD) |
| PC3 | -0.70 | -1.50 | 0.80 (0.64SD) |
| PC4 | 1.30 | 1.64 | 0.34 (0.57SD) |
| Log speed | 3.23 | 3.84 | 0.61 (0.87SD) |
| Proportion (IL-17) | 0.66 | 0.34 | |
| Proportion (IFN beta) | 0.77 | 0.23 | |
| Proportion (GBM) | 0.37 | 0.63 | |
| Proportion (Control) | 0.38 | 0.62 | |

**Supplementary Table 1.** Details of the mixture components detected by GMM using morphodynamic features. “Diff.” column shows the difference in values and standard deviation (SD) between two states. Clusters of latent vector PCs are mainly separated by PC2 and Log speed. Test subsets show different distributions of the two morphodynamic states. Clustering (with the same parameter and procedure) on image direct PCs is less clear, separations between mixture components and test subsets are weaker.

## Supplementary videos

**Video Set 1.** Video clips of representative cell trajectories visualized in Figure 2 and Supplementary Figure 6: we show full fields of view with bounding box and enlarged view of the cell. Note that among the 8 trajectories, 4 are more stable in appearance (corresponding to green and black arrows in Figure 2 and Supplementary Figure 6), 4 undergo significant morphology change (orange and brown arrows).

**Video Set 2.** Video clips of representative cell trajectories from each subset of the test dataset: 4 trajectories are provided for each treatment condition (control, GBM, IL17, IFN beta).

**Video Set 3.** Video clips of representative cell trajectories from each of the two morphodynamic state defined by GMM: 5 trajectories are provided for each state.

**Video Set 4.** Video clips of representative cell trajectories that underwent state transitions.