**Title**

DynaMorph: self-supervised learning of morphodynamic states of live cells

**Permalink**

https://escholarship.org/uc/item/3rk838cw

**Journal**

Molecular Biology of the Cell, 33(6)

**ISSN**

1059-1524

**Authors**

Wu, Zhenqin
Chhun, Bryant B
Popova, Galina
et al.

**Publication Date**

2022-05-15

**DOI**

10.1091/mbc.e21-11-0561

Peer reviewed

# DynaMorph: self-supervised learning of morphodynamic states of live cells

Zhenqin Wu[a,†], Bryant B. Chhun[b,†], Galina Popova[c,†], Syuan-Ming Guo[b], Chang N. Kim[c], Li-Hao Yeh[b], Tomasz Nowakowski[c,‡,*], James Zou[d,‡,*], and Shalin B. Mehta[b,‡,*]

[a]Department of Chemistry, Stanford University, Stanford, CA 94305; [b]Chan Zuckerberg Biohub, San Francisco, CA 94158; [c]Department of Anatomy, University of California, San Francisco, San Francisco, CA 94143; [d]Department of Biomedical Data Science, Stanford University, Stanford, CA 94305

**ABSTRACT** A cell's shape and motion represent fundamental aspects of cell identity and can be highly predictive of function and pathology. However, automated analysis of the morphodynamic states remains challenging for most cell types, especially primary human cells where genetic labeling may not be feasible. To enable automated and quantitative analysis of morphodynamic states, we developed DynaMorph—a computational framework that combines quantitative live cell imaging with self-supervised learning. To demonstrate the robustness and utility of this approach, we used DynaMorph to annotate morphodynamic states observed with label-free measurements of optical density and anisotropy of live microglia isolated from human brain tissue. These cells show complex behavior and have varied responses to disease-relevant perturbations. DynaMorph generates quantitative morphodynamic representations that can be used to compare the effects of the perturbations. Using DynaMorph, we identify distinct morphodynamic states of microglia polarization and detect rare transition events between states. The concepts and the methods presented here can facilitate automated discovery of functional states of diverse cellular systems.

## INTRODUCTION

Organs and tissues of the human body consist of an astonishing diversity of cell types, classically described by their anatomical location, shape, dynamic behavior, gene expression, and protein expression. 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 morphodynamic states can be difficult to identify either visually or with molecular markers. Quantitative analysis of morphodynamic states of human cells with high-throughput methods is a timely area of research.

While recent advances in single-cell genomics have significantly contributed to our understanding of molecular diversity of different cell types, morphological and behavioral properties of cells cannot be predicted from gene expression data alone (Way et al., 2021). Recent studies of morphological states of cells have relied on imaging of fixed cells labeled with a panel of fluorescent markers (Bray et al., 2016), live three-dimensional (3D) imaging of different cell compartments labeled with genetic markers (Driscoll et al., 2019), and phase-contrast imaging of live cells (Pincus and Theriot, 2007; Van Valen et al., 2016; Chan et al., 2020; Zaritsky et al., 2020). The morphological states have been analyzed with low dimensional representations computed with geometric or biophysical models (Rezaie et al., 2002; Pincus and Theriot, 2007; Keren et al., 2008; Marshall, 2011; Zaritsky et al., 2015; Shadkhoo and Mani, 2019), supervised learning of morphological labels (Neumann et al., 2010; Van Valen et al., 2016; Christiansen et al., 2018; Ounkomol et al., 2018; Moen et al., 2019; Guo et al., 2020; He et al., 2020), and, recently, self-supervised learning of latent representations of shape

(Chan et al., 2020; Zaritsky et al., 2020). These analytical approaches have been motivated by the unmet need for quantitative descriptions of specific, complex biological functions, such as motility of single cells (Pincus and Theriot, 2007; Keren et al., 2008; Marshall, 2011; Kimmel et al., 2018; Driscoll et al., 2019), collective cell migration (Zaritsky et al., 2015; Shadkhoo and Mani, 2019), cell cycle (Neumann et al., 2010; Van Valen et al., 2016; Moen et al., 2019), spatial gene expression (He et al., 2020), and spatial protein expression (Ounkomol et al., 2018; Guo et al., 2020; Kobayashi et al., 2021). In addition, data-driven integration of shape and gene expression (Carpenter and Sabatini, 2004; Neumann et al., 2010; Lane et al., 2017; Yang et al., 2021; Gerbin et al., 2021; He et al., 2020) is now enabling rapid analysis of functional roles of genes. Collectively, these studies illustrate that deep learning is poised to enable quantitative and interpretable analysis of complex cell phenotypes that are too high dimensional and too subtle for human vision to discern.

Despite this progress, we currently lack automated analysis methods for high-throughput screening of morphodynamic states of human cells, including their temporal dynamics, due to few key technological limitations:

- Measurements of morphodynamics of live human cells can be challenging because they are difficult to label consistently and without perturbing their native function. While dozens of molecular reporters have been developed to visualize cell structure based on gene expression studies, introducing molecular reporters in primary human cells is frequently impractical.

- Supervised learning from large-scale multidimensional imaging datasets remains challenging because identifying subtle changes in cell behaviors requires large amounts of annotation by biological experts, which is time consuming and hence not scalable.

- While the description of the molecular states of cells in terms of their gene expression profiles is now well established, similar morphodynamic states of cell behavior remain to be established.

Learning the latent representation of cell shapes via self-supervised reconstruction (Chan et al., 2020; Zaritsky et al., 2020; Kobayashi et al., 2021) of images is emerging as a promising method of transforming complex phenotypes into quantitative and interpretable shape modes. These shape modes provide a reduced set of cell shapes from which a large variety of cell shapes observed in the data can be generated. The shape modes that can be learned from the imaging data fundamentally depend on the information captured by the imaging modality. Self-supervision of cell contours with a variational autoencoder (VAE) (Pincus and Theriot, 2007) has enabled analysis of shape modes of motile cells. Self-supervision of qualitative phase-contrast images with an adversarial autoencoder has enabled prediction of the metastatic potential of melanoma cells (Zaritsky et al., 2020). Self-supervision of fluorescence images of labeled proteins with a regularized vector-quantized variational autoencoder (VQ-VAE) has enabled automated clustering of localization patterns of proteins (Kobayashi et al., 2021). These papers suggest that these and similar approaches may enable automated analysis of morphodynamic states of complex cells. In this paper, we demonstrate analysis of the morphodynamic states of human microglia with self-supervised reconstruction of quantitative label-free time-lapse images (Guo et al., 2020) using VQ-VAE regularized by temporal continuity of the cell shape.

Microglia are the resident macrophages of the central nervous system that are involved in brain development and homeostasis, as well as immune responses (Salter and Stevens, 2017). The microglial response to secreted cytokines or viral infections elicits profound changes in the transcriptome and dynamics of cells that are unique to perturbations (Chhatbar et al., 2018; Lively and Schlichter, 2018). Microglia survey brain parenchyma with their highly motile processes and respond to changes in the brain homeostasis by altering cell shape and motility (Nimmerjahn et al., 2005; Bernier et al., 2019), 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 manually curated features (Pincus and Theriot, 2007; Kimmel et al., 2018; Chan et al., 2020; Zaritsky et al., 2020) have been used to describe complex cell dynamics, including random search, persistent migration toward target, changes in cell shape due to interaction with other cells and targets, and endocytosis. Therefore analyzing the dynamic behavior of microglia under different perturbation conditions represents an ideal use-case for methods of interpreting complex cell dynamics.

We acquired reproducible measurements of cellular architecture and dynamics of microglia under immunogenic perturbations (Figure 1A). To avoid the necessity to virally transduce the cells with reporters, as has been done traditionally (Hansen et al., 2010; Noctor et al., 2019), we used quantitative label-free imaging with phase and polarization (QLIPP) (Guo et al., 2020). QLIPP measures phase and retardance, which report physical properties of the optical density and optical anisotropy, respectively, of the cells (Figure 1B). Trajectories of live cells (Figure 1C) were used as an input for DynaMorph, a deep learning-enabled framework that automatically extracts descriptors of shape and dynamics. DynaMorph relies on VQ-VAE to learn a latent representation of cell shape (Figure 1D). Enforcing temporal continuity of the latent shape modes in the neighboring frames leads to a quantitative representation of cell shape. We found that latent shape space regularized by the temporal continuity generalized well to new experiments. We evaluated dependence of the latent shape space on the imaging modality and found that shape modes learned from quantitative measurements were consistent across experiments. In this study, we combine shape and motion descriptors of cells to represent cell behavior and employed mixture modeling to identify different cell states (Figure 1E). The meaning of a "cell state" can vary with physiological and methodological context. In this work, we refer to "morphodynamic states" as a combination of morphological and temporal features. From the trajectory of cells in the latent shape space, we identified transitions among morphodynamic states of single cells. The same approach enabled detection of transitions in the morphodynamic states of cells as a result of immunogenic perturbations.

## RESULTS
### Learning a latent representation of cell shapes that generalizes across experiments
We evaluated fluorescent probes that label membrane and QLIPP for imaging shape and dynamics of primary microglia. We observed that the distribution of membrane probes can vary over time due to the turnover of the membrane. On the other hand, phase and retardance images acquired with QLIPP provided consistent (Supplemental Figure 1 - Supplement 1) readout of the cell shape and dynamics. We acquired label-free time-lapse images from primary human microglia over two sets of experiments. The first set of experiments was done with purified microglia without any perturbation and provided the training data for the model. The second set of experiments was done in the presence of immunogenic perturbations and provided the data used for testing the model and discovery of cell states. In the perturbation experiment (test set), multiple
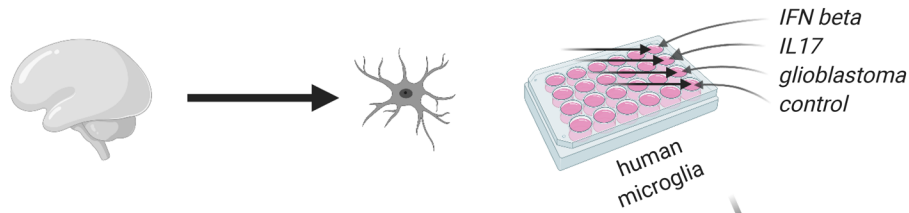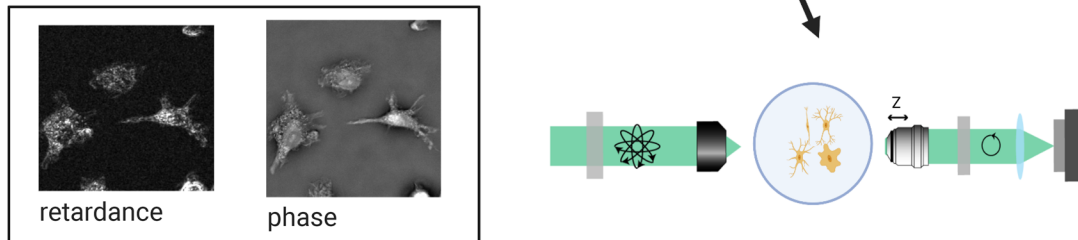
**FIGURE 1:** DynaMorph enables automated discovery of morphodynamic transitions in human microglia: (A) Human microglia are isolated from brain tissue and plated in 24-well plates and perturbed with cytokines of relevance to infection (IFN beta, IL17) or cancer (GBM supernatant), (B) Morphodynamics of perturbed microglia, along with control cells, are imaged using QLIPP imaging, which measures isotropic and anisotropic optical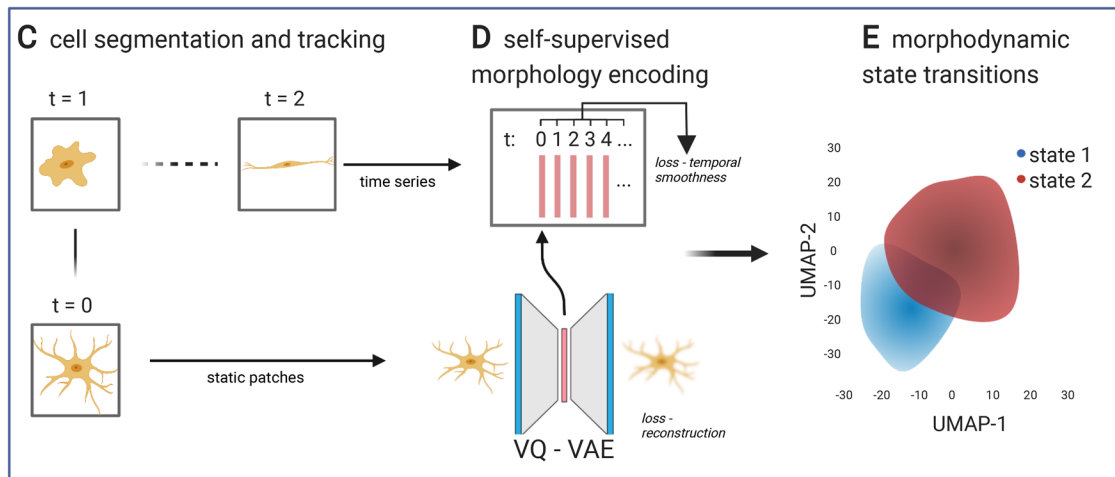 path lengths of cells. (C) Cells were segmented and tracked by in-house developed tools. (D) Generalizable and quantitative shape modes were learned from the thousands of tracked cells using a self-supervised model that reconstructs cell shape. (E) Morphodynamic states and transition among states under each perturbation were revealed via dimensionality reduction algorithms and clustering of most significant features.

treatments relevant to infection or cancer were applied to mimic different inflammatory states in neural tissue. Next, image segmentation methods were applied to extract individual cell positions in the time-lapse data (see *Methods*; Supplemental Figure 1 - Supplement 2). Then single-cell trajectory data were generated through instance segmentation and tracking modules described in Jaqaman *et al.* (2008) (see *Methods*; Supplemental Figure 1 - Supplement 3). While the specific segmentation and tracking methods suited our data, in principle any method that extracts cell position and correlates them across time is sufficient for the DynaMorph pipeline.

We observed significant diversity in shape and behavior of microglia across perturbations (see videos in Supplemental Video Set 1). While microglia are classically described as amoeboid, ramified, mobile, or immobile depending on their cell state, in our videos

they exhibited a variety of complex shapes accompanied by rare state transitions that we attempt to capture in our model. We used self-supervised deep learning to learn quantitative descriptions of complex shapes of microglia.

Patches of individual microglia from the training set were used to train a variant of autoencoder: VQ-VAE (van den Oord and Vinyals, 2017) (Figure 2; Supplemental Figure 2 - Supplement 1). The VQ-VAE was trained to reconstruct phase and retardance channels of the patches. The reconstruction of images serves as an auxiliary task to compute latent vectors for all patches. The latent vectors of all patches constitute the latent (or learned) shape space of microglia, which we abbreviate to *shape space* in the following text.

We tested multiple variants of autoencoder and concluded that VQ-VAE was the best-performing model architecture with the lowest
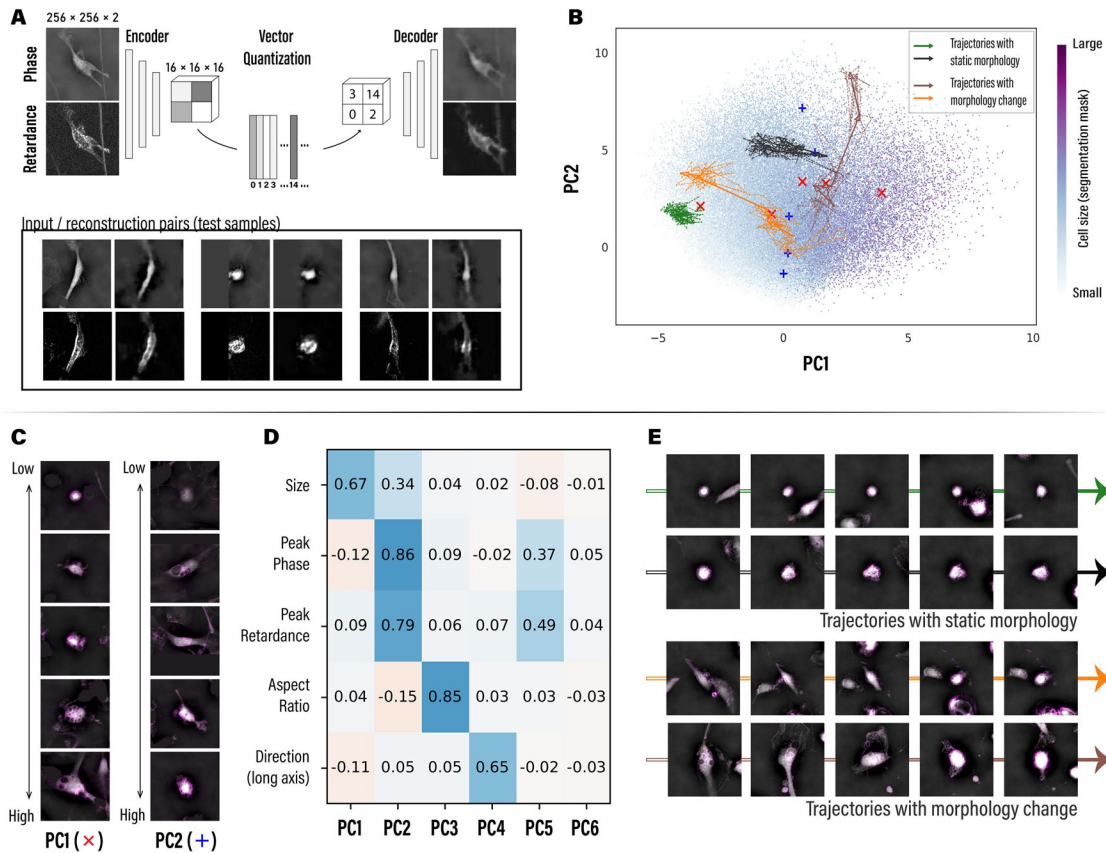
**FIGURE 2:** DynaMorph learns a generalizable and quantitative representation of cell shape via self-supervision: (A) We encoded the shape of microglia in quantized latent vectors and reconstructed the shape from the latent vectors through a VQ-VAE, which is trained in a self-supervised fashion. The 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 Supplemental Figure 2 - Supplement 1 demonstrates generalizability of the model. (B) Shape modes that describe the most significant differences were computed with PCA of the latent vectors. The resulting top two PCS 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. Four representative trajectories: two cells with static shape (green and black) and two cells undergoing changes in shape (orange and brown) are plotted in the PCA space as arrows. (C) The first two shape modes (PC1 and PC2) were interpreted by sampling representative patches along each PC axis (cross markers in panel B). Clear trends of changing size/cell optical 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 optical 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 (Supplemental Figure 2 - Supplement 1C). (E) Representative cells plotted in (B) are visualized. Significant changes in shape could be linked to jumps in the latent space.

reconstruction loss and consistent shape modes (Supplemental Notes; Supplemental Figure 2 - Supplement 2; and Supplemental Figure 2 - Supplement 3) in accordance with results reported for the general image reconstruction task (van den Oord and Vinyals, 2017; Razavi *et al.*, 2019). The self-supervised encoding led to a noise-robust and interpretable learned shape space of cells as discussed in the sections below.

In addition to the regular reconstruction and regularization loss, we imposed a matching loss in the model to minimize frame-to-frame distance between latent vectors of the same cell along its trajectory based on the prior knowledge that cell architecture changes smoothly between consecutive frames. This approach is motivated by the previous work that enforced temporal slowness constraint on the feature representation of adjacent frames (Zou *et al.*, 2012). Similar concepts have been applied to learning similari-

ties among human face images (Schroff *et al.*, 2015). As a result of applying the temporal matching loss, cell trajectories with small changes in cell morphology tend to stay localized in shape space, while trajectories with significant size/optical density/shape changes can be easily identified as trajectories showing large jumps in shape space. We explored how different scales of time matching loss affect reconstruction and embedding quality (see Supplemental Notes and Supplemental Figure 2 - Supplement 10) and chose the model with optimal weight for the following analysis. For full details of the training procedure, see *Methods* and loss curves during training in Supplemental Figure 2 - Supplement 4A.

To evaluate how the shape space depends on the input data, we trained and evaluated models using individual label-free channels (brightfield, phase, and retardance). By comparing reconstruction loss, latent space correlations to peak intensity, aspect ratio and

long axis, and latent vector predictive capability, we found that models trained on single channels perform worse than models trained on complementary channels (see Supplemental Notes and Supplemental Figure 2 - Supplement 4).

Overall, we find that the shape space of the DynaMorph autoencoder substantially compresses the data in the raw imaging—by over 680 times. The model trained on normalized cell patches has reconstruction loss 0.19 ± 0.21 SD (cell/foreground are usually 3–4 SD brighter than background). This was further validated on the test set on which model reconstruction loss is 0.25 ± 0.20 SD with no detectable signs of overfitting.

Comparison of reconstructed shapes from the test set (Figure 2A and Supplemental Figure 2 - Supplement 5) and the training set (Supplemental Figure 2 - Supplement 1A) along with the analysis of the shape space described in the next section show that our self-supervised model trained on training dataset generalized well between independent experiments and can be used to compare cell state changes between control microglia and cells treated with multiple perturbations.

### Learned shape space provides a quantitative description of shape

The learned shape space could provide a quantitative reduction of complex shapes of cells. We sought to evaluate if the shape space provides quantitative metric of shape, i.e., we sought to answer the following questions: are the most significant modes of the shape space interpretable? Does the distance in the shape space capture shape similarity?

*Interpretation of the learned shape space.* We use principal component analysis (PCA) to extract the top shape modes and visualize the shape space. Distributions of the first two principal components (PCs) are displayed in Figure 2B, in which each individual dot represents an encoded cell patch, and the color indicates the cell size measured from its segmentation mask. A horizontal gradient of cell size is observed, which suggests quantitative relationships between top PCs and cell properties. Supplemental Figure 2 - Supplement 1B shows similar results for the cells from the training set.

To further unravel these relationships, we sought to link the top PCs with heuristic cell geometric properties. We first sampled cell patches along PC axes (Figure 2C and Supplemental Figure 2 - Supplement 6) and interpreted the cell properties visually. We were able to identify significant correlations in each PC: PC1 relates with cell size; PC2 relates with peak phase and retardance, which measure cell optical density; PC3 and PC4 relate with cell orientations. Quantification using Spearman's rank correlation coefficients (Figure 2D) is consistent with our visual interpretation of each PC. Thus the first two PCs of the shape space correlate strongly, but not exclusively, with cell size and cell optical density.

The correlation of the first PC with cell size intriguingly parallels the analysis of cell masks of other immune cell types in Chan *et al.* (2020). They also show that the most significant shape mode of three different motile cell types is cell size. In addition, the second most significant shape mode in our data relates with cell optical density as measured with phase and retardance, which comports with the result in Zaritsky *et al.* (2020) that light scattering is a significant descriptor of cell behavior. Thus our results show that the shape space learned by the VQ-VAE model indeed captures dominant shape modes that have been reported in distinct cell types. This correspondence suggests that cell size and cell optical density may be highly conserved descriptors of cell states due to mechanisms that are yet to be understood.

Intriguingly, while 80% of the shape variance of ameboid cells analyzed in Chan *et al.* (2020) is captured by five shape modes, we found that the first four shape modes (or PCs) of microglia account for less than 20% of all variances (Supplemental Figure 2 - Supplement 7A). The high variance of the shape space of microglia can be due to more complex shapes of microglia, such as diversity of protrusions, subcellular structures and variations in cell optical density, location of nuclei in migrating cells, and so on. As we mentioned above, the inclusion of several imaging channels (brightfield, phase, and retardance) increases the performance of the model possibly by increasing the diversity of morphological information encoded in our input data. UMAP (McInnes *et al.*, 2018) projections of the shape vectors of cells show no apparent clustering patterns (Supplemental Figure 2 - Supplement 7, B and C), suggesting that microglia shape undergoes continuous change and hence no discrete states could be defined based on static frames.

To validate that the learned shape space is more informative, we applied PCA to the input images and followed the same analysis procedure. We found that the top PCs computed from the input images are weaker and are much less correlated (Supplemental Figure 2 - Supplement 8A) with cell morphology and consequently are harder to interpret. The PCs of input images could not be separated by clustering analysis (Supplemental Figure 2 - Supplement 8B and Supplemental Table 1). We also tried directly using heuristic morphological descriptors in downstream analysis (see Supplemental Notes and Supplemental Figure 3 - Supplement 3), but results are worse than learned shape space. These results illustrate that the shape space learned with VQ-VAE indeed provides an interpretable and quantitative representation of the complex shape of microglia that cannot be derived from simple preidentified shape parameters alone.

*Trajectories in the learned shape space characterize dynamics.* To evaluate how the shape space regularized with temporal matching models of cellular dynamics, we generated single-cell trajectories of latent representations by concatenating representations of static frames. As a control, we compared the distance between the latent representations of temporally adjacent and nonadjacent frames (both before and after quantization of latent vectors). We found that the distance between temporally adjacent frames is substantially less than nonadjacent frames and random baseline, indicating that our shape space captures the continuity of cellular dynamics (Supplemental Figure 2 - Supplement 9).

We further compared the distance distribution between latent representations generated with and without matching loss. Results show that the reconstruction task already reduces the distances between adjacent frames, as compared with random pairs, while including the matching loss further regularized the shape space such that it reliably represents the temporal evolution of the shape (Supplemental Figure 2 - Supplement 9).

Visualization of individual cell trajectories in the shape space shows that the majority of cells had stable shape modes 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 top shape modes are relatively constant, suggesting that their morphodynamic states are stable throughout the experiment. However, we also observed a number of "leap" events in the shape space, which coincide with cells undergoing significant transitions in shape. Representative examples (marked in orange and brown) are also shown in Figure 2E in which transitions in size or optical density could be observed. Full video clips of these example cell trajectories can be found in Supplemental Video Set 2.

## DynaMorph discerns between pro- and anti-inflammatory microglia states

Microglia can respond to pro- and anti-inflammatory stimuli by changing their shape, dynamics, and gene expression. While a range of small molecules, including cytokines and chemokines and their individual effect on microglia, has been well described, cell conditions with both pro- and anti-inflammatory signaling molecules in the microenvironment present a more challenging task. One example is a brain tumor microenvironment, such as glioblastoma (GBM), where a wide array of signaling molecules can simultaneously promote proinflammatory states of microglia and anti-inflammatory states. Characterizing the nature as well as heterogeneity of microglia states under defined perturbation conditions could enable automated analysis of complex responses of microglia to disease states and could ultimately serve as a benchmark for evaluating the effects of pharmacological perturbations in a high-throughput manner.

We treated cells with a proinflammatory cytokine interleukin 17-A (IL17; 100 ng/ml), which has been implicated in several brain diseases (Kawanokuchi *et al.*, 2008; Choi *et al.*, 2016; Yu *et al.*, 2016), and interferon beta (IFN beta; 1000 U/ml), an antiviral and anti-inflammatory cytokine (Blank and Prinz, 2017; Mudò *et al.*, 2019). In parallel, we treated microglia with supernatant from primary human GBM cells. Cell 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 Supplemental Video Set 3. As already shown in Figure 2, reconstruction of patches and interpretation of shape space are reliable across perturbations.

We first explored distribution of shape modes and speed of cells under different perturbations. For each cell, we collected and averaged its PC1 and PC2 values along the trajectory (Figure 3A). Distributions of these trajectory-averaged PC1 and PC2 values are shown for each treatment in Figure 3B and Supplemental Figure 3 - Supplement 1. PC1 of shape showed a similar distribution across all four conditions, suggesting that the size distribution of the cells is consistent across different conditions. In the PC2 space, which correlates closely with cell optical density, IL17 and IFN beta subsets had similar distributions, which differed from control and GBM subsets. This suggests that microglia under IL-17 and IFN-beta treatment tend to have lower optical density while maintaining a similar size. Speed is plotted separately (Figure 3B) and against two top shape modes as kernel density estimation (KDE) plots (Figure 3, C and D). Microglia display a broad range of motility under different conditions with the fast-moving population (control subset) traveling over two times faster than the slow-moving population under an anti-inflammatory IFN beta treatment. Note that this measurement is based on centroids of cells and hence could be confounded by shape changes, though the scale of shape changes is much smaller than typical cell movement.

To better understand the interplay between shape and dynamic behavior, we constructed trajectory feature vectors (TFVs) by first concatenating the PCs of the shape space with the instantaneous speed of the cell (between two consecutive frames) and then averaging them along the trajectory (Figure 3A). These vectors are then clustered to distinguish morphodynamic states of the cell population (Figure 1F) and to analyze how cells transition between these states.

Interestingly, while microglia treated with a single cytokine (IL17 or IFN beta) displayed unimodal speed distribution (Figure 3B), GBM supernatant induced a much wider range of speeds with clearly distinguishable individual peaks around 16, 32, and 60 µm/h.

This finding suggests that several intermixed populations of microglia may exist under the GBM cell microenvironment, further highlighting the importance of single-cell tracking in understanding cell behavior under complex cell signaling milieu. Given the striking differences in cell speed, we plotted changes in top shape modes and average cell speed (Figure 3, C and D). Our data reveal that microglia under proinflammatory stimuli had a greater proportion of faster and denser cells, while under the IFN beta treatment the cell distribution was shifted toward slower moving cells. Control microglia had a distribution resembling the trend observed in IL17 treatment, suggesting that in vitro culture conditions induce microglia activation as has been previously noted (Gosselin *et al.*, 2017). In contrast, cells treated with the GBM supernatant had a shape distribution that simultaneously resembled both control and IL17 and IFN treatment, suggesting that different cell populations within this condition would adopt either a pro- or an anti-inflammatory morphodynamic cell profile rather than being a homogeneous population of an intermediate response.

The separation was further revealed in a discriminative study that predicts treatment/condition purely based on TFVs (Supplemental Figure 3 - Supplement 2; see Supplemental Notes for details). In the results from gradient boosted decision tree models (Friedman, 2001), confusion between control subset and IFN beta subset is especially small, while confusion between IL-17 and GBM is larger.

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 induces a state characterized by reduced optical density and migration speed. In contrast, exposure to GBM supernatant, which likely contains proinflammatory and anti-inflammatory cytokines, induces a broader distribution that might consist of multiple morphodynamic states.

## Identification of morphodynamic states and transitions between them

To further probe how learned shape space and motion of microglia under different perturbations can reveal morphodynamic states, we employed mixture modeling of the TFVs of test cell patches to locate and quantify the underlying cell states. We found that TFVs formed a continuous spectrum rather than discrete clusters. Therefore instead of using clustering algorithms such as hierarchical or K-means clustering, we employed a Gaussian mixture model (GMM) to identify morphodynamic states in the test population, which could be further used to describe distribution of morphodynamic states across perturbations.

We averaged TFVs over 2-h-long segments to reduce noise in the mixture modeling. In addition, we used a hidden Markov model (HMM) (see *Methods* for details) to evaluate how using the TFVs of individual frames affects state detection. Details of the hidden states and the proportion of trajectory assignments computed with GMM and HMM models are shown in Supplemental Tables 1 and 2. The GMM-based detection of morphodynamic states of 2-h-long segments was almost as sensitive as HMM-based detection of morphodynamic states of individual frames, likely because cells in our experiments had reached steady state dynamics after the perturbations. Therefore we report clustering results obtained with the simpler GMM model.

The model revealed two components in the combined test population (see *Methods* for details). We refer to these components as state-1 and state-2, respectively, in the following text. Morphodynamic distributions of the two states are described in Figure 4F. Qualitatively, based on cell appearance and features of the two states, we describe state-1 (blue) cells by their large size, low-optical
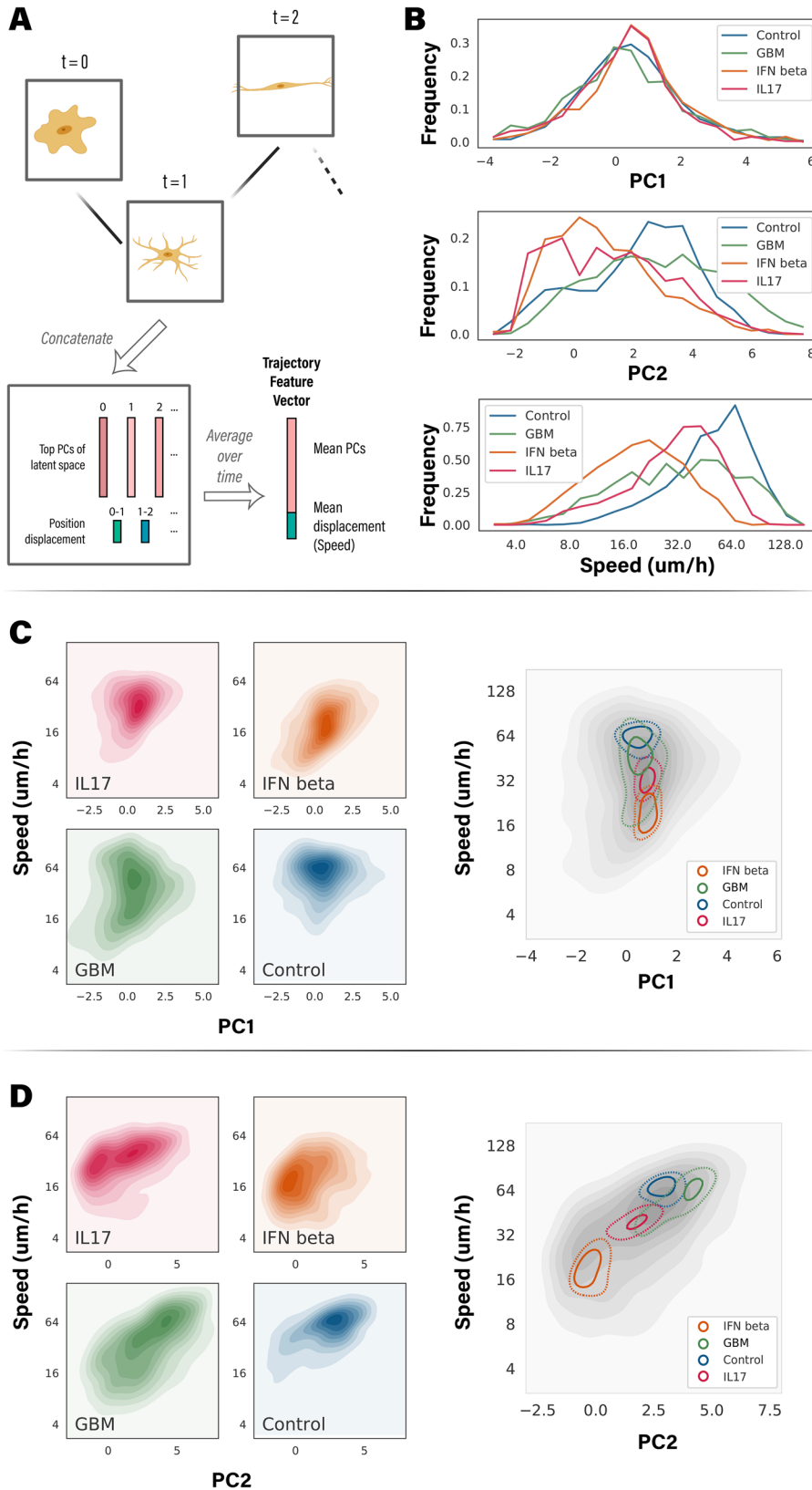
density, and slow movement and state-2 (red) cells by their higher optical density and fast movement. The two states differ the most on PC2 and mean speed and have slight differences along PC1 as well. Based on the previous observations that microglia adopt an amoeboid shape on exposure to anti-inflammatory stimuli, we speculate that state-2 represents activated (faster moving, denser, and round-shaped) microglia, while state-1 represents homeostatic (ramified and slower moving) microglia.

Compared with intercondition separation of cells, interstate separation is much stronger, suggesting that the states identified by GMM are superior representations of morphodynamic modes of microglia. Representative samples from the two states are illustrated in Figure 4B with their trajectories plotted as colored lines. Representative cell trajectories can also be found in Supplemental Video Set 4.

A more quantitative analysis of cell motion showed 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 subset, as anticipated. We also consider persistence of cell direction in the two states, which represents an important feature of cell migration, noting how the direction of the movement of state-1 cell in Figure 4B aligned with the direction of its cell body. We computed angles between movement and the long axis of the cell body in each frame and computed vector sums over these angles. Resulting trajectory-summed directions are visualized in Figure 4H. Distributions of the directions show a dominant peak at 0° (Figure 4D), indicating a very high preference for cell body-aligned motion. Notably, state-2 cells are dense and round with less pronounced alignment between

mean displacement, i.e., mean speed. The geometric properties of cell size and density correlated highly with PC1 and PC2, respectively. Trajectories under IFN beta treatment show significant differences from trajectories under control treatment in cell optical density and speed. (C) and (D) We analyzed changes in average cell size (PC1) and average cell speed, as well as average cell optical density (PC2) and speed by plotting their joint distributions. Left panels show optical density plots of distributions under different perturbations. Right panels show the combined distributions, with apex contours under each perturbations marked. Separations between IFN and control conditions observed in (B) were confirmed.

**FIGURE 3:** DynaMorph reveals differences in morphodynamics of microglia under different perturbations. (A) To compare how the morphodynamics of microglia change in response to different perturbations, we computed trajectory-averaged features as follows: PCs of the shape space 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 TFVs under multiple perturbations by plotting probability densities of trajectory averaged PC1, PC2 and
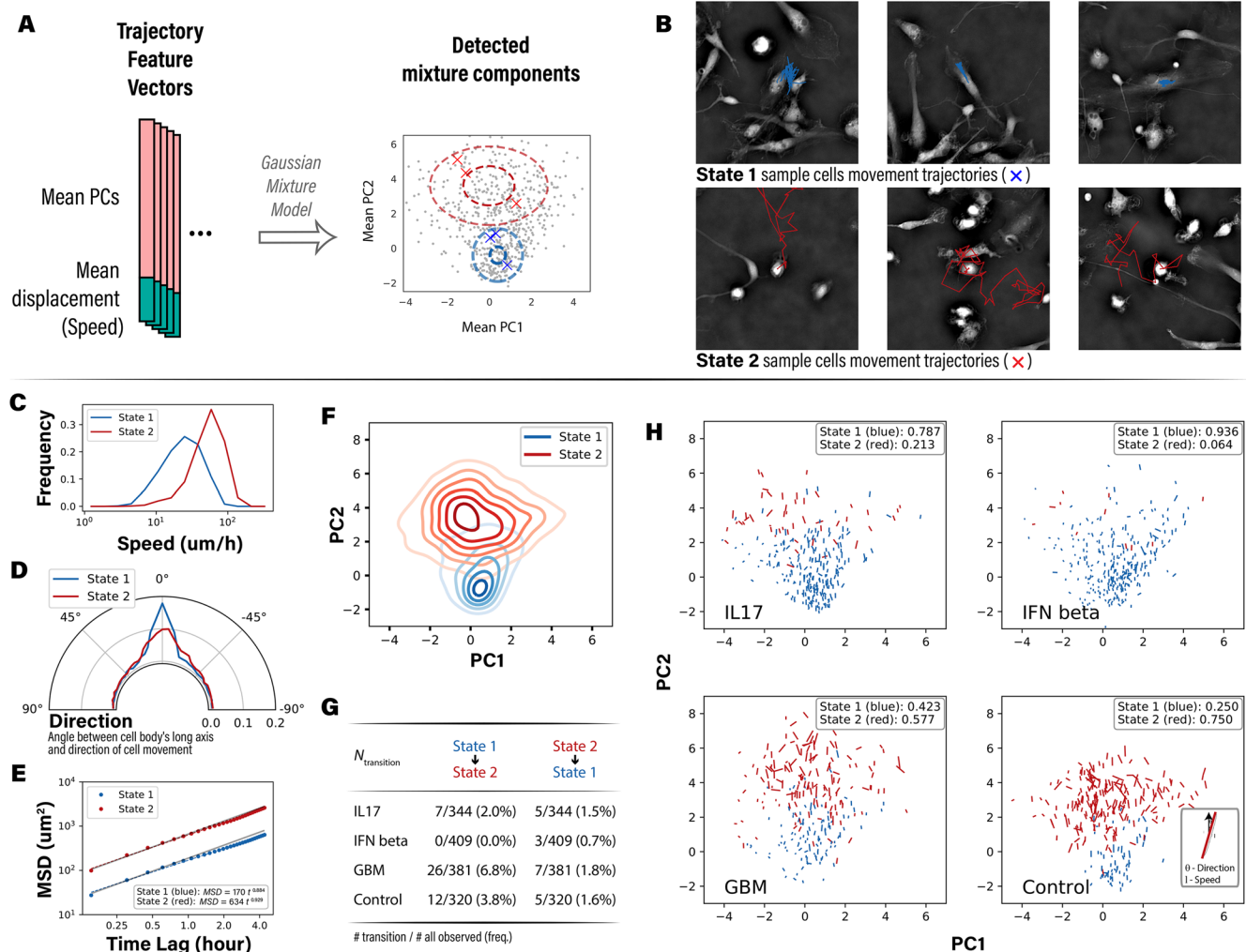
FIGURE 4: Cell trajectories are clustered into two states based on the morphodynamic features. (A) We identified two states of the trajectories by pooling their TFVs and clustering them with a GMM. B Representative trajectories from the two states (corresponding to cross markers in panel A) are shown: 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 (grey shades indicate standard random motion) over the timescale 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 optical density relative to cells in state-2, as seen from the KDE of PC1 (cell size) and PC2 (cell optical 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 h 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 is noted in the insets.

cell movement direction and cell axis. To quantify whether the cellular motion was diffusive, constrained, or persistent, we analyzed the mean-squared displacement (MSD) of the two states. Figure 4E and Supplemental Figure 4 - Supplement 1 plot the log-log fit of MSD over time lag. Both states have similar slopes (~0.9) that are close to 1, indicating a near-random motion, as expected from the isotropic perturbations experienced by cells. State-2 cells have larger intercept, indicating a larger diffusion constant and hence faster migration. Taken together, the distinct morphodynamic properties exhibited by two cell states revealed by GMM suggest the

two-state model is a good description of the cell populations across perturbation conditions.

Among the different perturbation conditions in the test population, cells are widely distributed between two states depending on the treatment (Figure 4H and Supplemental Table 1); a majority of the control subset cells are in state-2 (~75%), while almost all IFN beta-treated cells are in state-1 (>90%); the GBM subset falls between the two extremes and is almost equally split between the two states. Note that due to the huge separation between control and IFN beta subsets, many of the cells in IL-17 subset are labeled as

state-1. We treat this conflicting classification as a side effect of the simplified two-state model, as feature distribution in Figure 3, C and D clearly shows the distinction between IFN beta and IL-17 subsets in cell shape and motion. The treatment of microglia with GBM supernatant led to a distribution that included nearly equal distribution across both states, suggesting a complex mixture of both pro- and anti-inflammatory cytokines.

Furthermore, the quantitative definition of states allowed us to quantitatively characterize state transition events such as the ones shown in Figure 2E (orange and brown trajectories). Given the shape and motion changes throughout the imaging period, we separated each trajectory into segments (2 h each) and applied the unsupervised GMM to estimate posterior probabilities of cell states for each segment. Cells that have different states assigned to different segments would be of great interest to the study of dynamics and behavior of broad cell types under different conditions. 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 the most frequent cell transitions were observed among cells treated with GBM supernatant. One possible explanation for this imbalance is that IFN-treated cells represent a single polarization axis, while a heterogeneous cell signaling milieu derived from cancer cells provides conflicting pro- and anti-inflammatory signals, instructing cells to transition between the states. 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 within the chosen time frame. This imbalance between the rates of state transitions correlates with the higher state-2/state-1 ratio in GBM and the control environment and may explain the longitudinal accumulation of cells in a more activated state under these culture conditions. Representative state-transition cells can be found in Supplemental Video Set 5.

We further compared the detected morphodynamic states with scRNA measurements of the same cell populations. Interestingly, the separation of cells in state-1 and state-2 from control and the IFN group parallels the clusters identified with cell transcriptome (see Supplemental Notes; Supplemental Figure 4 - Supplement 2; and Supplemental Figure 4 - Supplement 3 for details), suggesting that correlative analysis of gene expression and morphodynamics can reveal molecular programs underlying these phenotypes. In our preliminary analysis, scRNAseq revealed a greater degree of granularity in each of the cell populations, such as cluster 1 of the scRNA-seq separating into three additional subclusters. Cluster 1-2 was defined by high expression of IFN response genes, and predictably, this cluster was primarily derived from the cells treated with IFN beta. IFN exposure induces morphological changes associated with increased cell perimeter, which reports ramification of microglia membrane (Aw *et al.*, 2020). It was also shown that infections with neurotropic viruses, leading to IFN response, also lead to decreased velocity and distance traveled for cultured microglia cells (Fekete *et al.*, 2018). These observations are in direct agreement with the higher proportion of cells in state-1, characterized by lower cell velocity. Interestingly, scRNAseq analysis also identified a population of cells with a high expression of cell cycle genes (Cluster 1-3), which would also be predicted to have a slower speed and potentially larger cell body. These results point to the fact that different molecular states may be underlying very similar morphodynamic states.

**DynaMorph can be easily extended to other cell types**
Building on the successful application of DynaMorph to the discovery of microglia cell states, we sought to evaluate the generalizabil-

ity of our model by applying DynaMorph to nonmicroglia cells that were present in our cultures. Specifically, neural progenitors can be easily identified by their elongated cell bodies with radially extending processes, different from the amoeboid or ramified shape of microglia.

VQ-VAE trained with the control experiment cells generalizes well to the progenitor cells from the perturbation experiment, with a mean VAE reconstruction loss of 0.25 SD over patches containing progenitor cells, close to the reconstruction performance of microglia. Then through the same trajectory building and encoding pipeline, we extracted the TFVs for 125 identified progenitor cells. Sample cells along with their movement are illustrated in Figure 5A; corresponding cell trajectories can be found in Supplemental Video Set 6.

Figure 5B shows the first two shape modes and motion features of progenitor cells. These results indicate that progenitor cells are smaller (lower PC1) and less dense (lower PC2) than microglia. Progenitor cells have a distinct cell motion as compared with microglia. Aligned with its polarized shape, a progenitor cell's movement usually follows the direction of its protrusions, as illustrated in Figure 5A. This results in a much stronger preference for directional motion as compared with microglia and, consistent with the fact that these are not immune cells, progenitor cells did not substantially alter their behavior when subjected to immunogenic perturbations.

In Figure 5C we plot the average of shape features over trajectories for microglia and progenitor cells, which demonstrates a clear difference in shape between the two cell types. Finally, to directly compare morphodynamic states of these two cell types, we used the TFVs which contain motion and shape information to calculate pairwise Euclidean distances. Figure 5D illustrates the distribution of distances between progenitors or between progenitor and microglia. Results reaffirm that microglia are in general different from progenitor cells, and state-1 microglia are more similar to progenitor cells. This result agrees with the visual evidence that state-1 microglia have a similar elongated cell shape and a higher tendency for directional movement (Figure 4B). Together, these results suggest that our model performs equally well on the highly motile immune cells and less dynamic progenitor population and can be used to distinguish between different cell populations while maintaining the level of granularity to identify different states within the same cell type.

## DISCUSSION
In this work we present DynaMorph, a deep learning framework for the dimensionality reduction and for the quantitative analysis of live cell imaging data. We demonstrate utility of the framework by analyzing morphodynamic states from two complementary physical properties of cells—their optical density and anisotropy. By combining novel self-supervised learning with tracking and mixture modeling, the workflow allows us to analyze cell dynamics and behavior at both the individual cell level and the population level.

Our work formalizes an analytical approach for a data-driven discovery of morphodynamic cell states based on quantitative shape and motion descriptors. A cell state can be rigorously described in terms of measurements of the cells, and recent developments in measurement techniques, including imaging modalities and single-cell genomics, keep adding to the growing list of the features that can be measured. Time-lapse imaging is high dimensional and therefore admits multiple definitions of a cell state. All components in the pipeline can be adapted based on the goals of analyses and characteristics of the imaging system. Here we specifically demonstrate the utility of this method to discover and
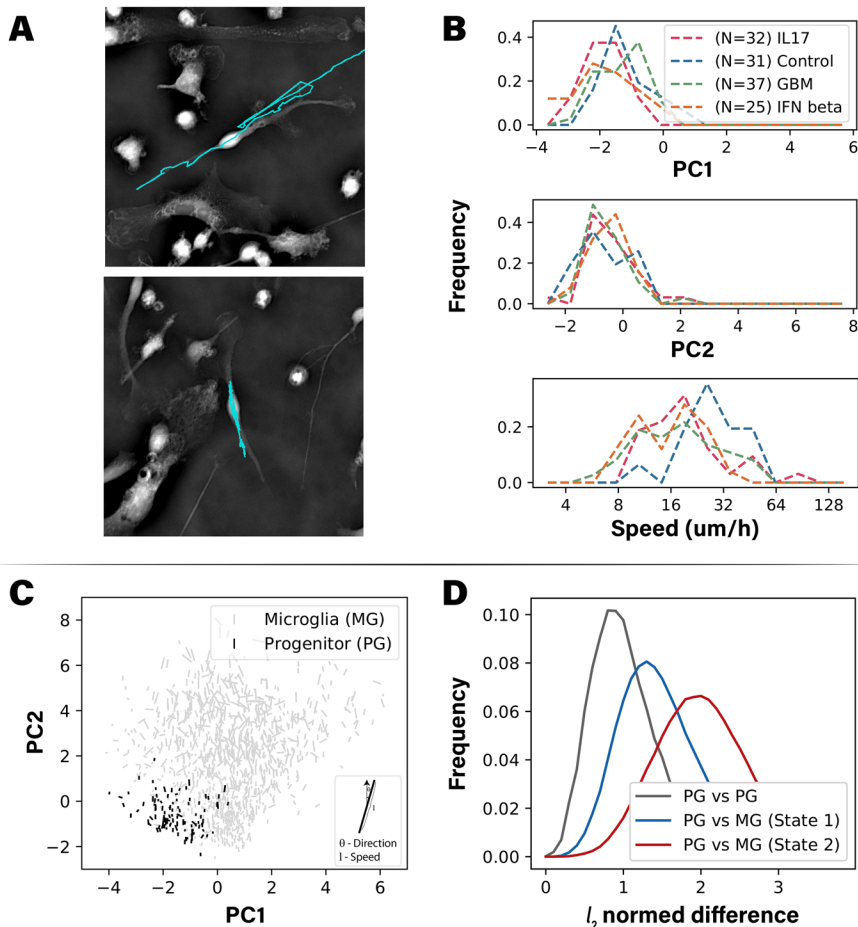
**FIGURE 5:** Morphodynamic state of neural progenitor cells in the test dataset. (A) Example trajectories of progenitor cells are shown; cell motions are marked as cyan lines. Note that the elongated cell bodies align with cell motions. (B) Probability densities of top PCs and cell speeds. Shape of progenitors is robust to perturbations. (C) Scatter plots comparing shape, speed, and orientation of progenitor cells and microglia (similar to Figure 4H). The two cell types occupied different regions of the shape space. (D) $l_2$ normed differences of TFVs between progenitor cells and microglia in two states. Note that progenitor cells are closer to microglia that are in state-1.

cell anisotropy can be complemented with molecular measurements of metabolic states, cell cycle, or proteins of relevance to specific cell types. Furthermore, while we demonstrated this proof-of-principle on microglia and neural progenitor cells, this approach can be used to track cell transitions during cell differentiation and lineage specification as well as broad cellular functions like viral infections and response and apoptosis.

The following are areas of improvement in DynaMorph:

1. Segmentation and tracking accuracy are currently limited to sparse cell populations. The denser environment significantly increases the difficulty of recognizing individual cells. Methods that incorporate more human supervision or models based on geometric representations of shapes can better extract and analyze cell behavior in a dense environment.

2. We found good agreement between gene expression changes and morphodynamics in response to different perturbations at the resolution of cell population. However, a direct correlation between the single-cell morphodynamic behavior and its transcriptomic profile is still an area of future work. The transitions in gene expression and morphodynamics driven by the perturbations can be elucidated by paired measurements of gene expression and morphodynamics per cell. Correlative single-cell measurements of morphodynamic states and single-cell transcriptomes, such as Patch-Seq or laser microdissection, are needed to decisively link morphodynamics and molecular programs underlying these phenotypes.

classify morphodynamic states of human primary microglia cultured in vitro and transitions between these states on exposure to polarizing, disease-relevant environments. We show that the GBM microenvironment differently affects microglia cells by inducing a wide range of states consistent with both pro- and anti-inflammatory responses. While it is not known what factors predispose individual microglia cells to adopt one or another state under complex cell environments such as brain tumors, 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. Once generalizable models are trained and evaluated as we describe, label-free imaging and DynaMorph can enable online inference of changes in the morphodynamic states in cell population with single-cell resolution, opening new avenues for single-cell analyses.

In addition to the application presented in this work, DynaMorph can be used with other cell types and other quantitative measurements of dynamic architecture to delineate morphodynamic states. For example, the physical measurements of cell optical density and

3. The biological meaning of the morphodynamic states of the microglia we have identified from an average of feature vector across trajectories (Figure 4) is confounded by two possible causes: a) the cell populations subjected to the same perturbations exhibit heterogeneous dynamics, and b) the dynamic behavior in the latent shape space does not always align with the existing vocabulary of cell behavior. We found that HMM is marginally better at detecting heterogeneity in behavior in cells that had reached steady state after perturbation. However, imaging the evolution of dynamics after perturbation at high time resolution and optimizing the HMM can enable a more complete analysis of biological heterogeneity. How to ascribe a biological interpretation to the learned representation is an open question for the field. A mathematically rigorous and experimentally generalizable representation can be considered a quantitative vocabulary with which we describe the dynamics of cells.

Taken together, DynaMorph is a versatile and quantitative framework for the automated discovery of morphodynamic states of cells

from live imaging data. Extraction of shape and motility descriptors can be coupled with a broad spectrum of downstream tasks, for example, anomaly detection, unsupervised state discovery, and supervised phenotype classification. Further integration with proteomic and transcriptomic assays can enable a more complete understanding of cell types and cell states. Identifying morphodynamic states in the context of disease could help elucidate the natural and induced changes in cell behavior, potentially informing novel therapeutics or diagnostics.

## METHODS
Request a protocol through *Bio-protocol*.

### Label-free brightfield, phase, and retardance imaging
We acquired all data using a Leica DMI-8 inverted widefield microscope, with 20× objective magnification at 0.55 NA (air), and 0.4 NA condenser on a Hamamatsu Flash-4 LT camera (6.5-µm 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 was acquired, in sequence, with 50-ms camera exposure for each of five z-planes for a given field of view. For the unperturbed microglia time series (training set), we acquired 52 time points and 27 fields of view (9 per well) over 24 h at 27-min time intervals. For the perturbed microglia time series (test set), we acquired 159 time points and 9 fields of view (4 were used for analysis) over 24 h at 9-min 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, allows 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 from there compute 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 (Guo *et al.*, 2020).

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 calibration by first finding voltages to achieve extinction ($I_{Ext}$, intensity minimum). A predefined "Swing" voltage (0.03) is applied to induce slight polarization ellipticity along one axis ($I_{0deg}$). 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 $I_{0deg}$ that sample other orientations ($I_{45deg}$, $I_{90deg}$, $I_{135deg}$).

Following the reconstruction algorithm described previously (Guo *et al.*, 2020) 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 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 with CD11b magnetic beads. Tissue samples were dissected in artificial cerebrospinal fluid containing 125 mM NaCl, 2.5 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, and 1.25 mM $NaH_2PO_4$. The tissue was cut into 1-mm³ pieces, and the tissue was enzymatically digested using 0.25% trypsin (reconstituted from 2.5% trypsin, ThermoFisher 15090046) with the addition of 0.5 mg/ml DNase (Sigma Aldrich, DN25) for 20 min at 37°C, mechanically dissociated, and then passed through a 40-µm mesh cell strainer (Corning 352340). The resulting cell suspension (100–150 million cells) was centrifuged for 5 min at 300 × *g* and washed twice with $Ca^{2+}/Mg^{2+}$-free phosphate-buffered solution (PBS) with the addition of 0.5 mg/ml DNase to prevent cell clumping. Cells were resuspended in 900 µl of MACS buffer (PBS with 0.5% bovine serum albumin) with the addition of 0.5 mg/ml DNase and incubated with 100 µl of the CD11b magnetic beads (Miltenyi Biotec, 130-049-601) for 15 min following the manufacturer's instructions. After the magnetic bead incubation, cells were washed with 20 ml of PBS, spun down at 300 × *g*, resuspended in 0.5 ml of MACS buffer, and loaded on a MACS LS column (Miltenyi Biotec, 130-042401). Cells on the column were washed three times with 3 ml of MACS buffer, the column was removed from the magnet, and the 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, 12 ml HBSS, 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 spun down at 300 × *g* and plated at 100 ×10³ cells per well in microglia media supplemented with 100 ng/ml of rhIL34 (Peprotech, 200-34), 2 ng/ml TGFb2 (Peprotech, 100-35b) and 1 × 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 h at room temperature followed by three double-distilled water washes and additionally incubated with laminin and fibronectin in PBS for 3 h at 37°C. Cell culture media were changed twice a week and were changed 24 h prior to the start of the time-lapse experiment to allow cells to re-equilibrate. Cytokine and IFN treatments were performed 24 h before imaging. Final dilutions of 1000 U/ml of human recombinant IFN beta (Peprotech, 300-02BC), 100 ng/ml of human recombinant IL17A (R&D Systems, 7955-IL-010/CF), or supernatant from primary human GBM cell cultures (diluted with fresh microglia culture media at 1:1 ratio) were used 24 h prior to imaging.

The time-lapse imaging experiment was started on day 6 of culture and continued for 24 h in an 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 a morphodynamic summary of target instances. Here we elaborate technical details of the preprocessing steps and DynaMorph pipeline. Note that two components of the pipeline required

training in advance to state discovery applications: the semantic segmentation model during preprocessing and the self-supervised encoding model. In this work we dedicated a separate set of experimental data on control microglia for model training and conducted a majority of the analysis on perturbed microglia.

In order to identify and track microglia, DynaMorph first uses segmentation algorithms to extract individual cells from the imaging data. Without detailed annotation of the data, we built the segmentation method based on the deep learning semantic segmentation model U-Net with weak supervision from a human expert. It should be noted that a variety of similar deep learning-based segmentation tools are proposed in the recent literature, most of which will be fully applicable in our pipeline. For the application of DynaMorph, the selection of segmentation tool is interchangeable as long as the segmentation results provide an unbiased selection of target cell population.

The input data for both semantic segmentation and VQ-VAE models are 2D images of computed phase and retardance that measure integrated optical density and anisotropy across the depth of the cell. The raw collected data are 3D (five z-slices acquired in multiple polarization channels). The 2D phase is computed from the full stack of brightfield images via deconvolution. The retardance is computed from an average of the intensities across the five z-slices. Subsequent model training is more tractable with 2D data instead of 3D while capturing the cell architecture across the depth.

*Semantic segmentation.* A single three-class U-Net classifier (Ronneberger *et al.*, 2015) was deployed for the cell semantic segmentation whose training data were derived from a combination of human annotations and augmented background annotations. In the preparation of training data, we utilized the interface from ilastik (Berg *et al.*, 2019) and manually annotated microglia, contaminating nonmicroglia cells (mostly progenitor cells also containing neurons, radial glia), and background/experiment artifacts in 30 frames selected from 13 fields of view. Manual annotation took in total approximately 8 h generating identifications for 1252 microglia and 312 nonmicroglia cells. The annotations were only partial in which most of the background was not annotated (Supplemental Figure 1 - Supplement 3B). To compensate (see Supplemental Figure 1 - Supplement 2 for comparison), we utilized the internal segmentation module (random forest-based) of ilastik to separate foreground (microglia and contaminating cells) and background (under human supervision), which generates supplementary background annotations that could be used as data augmentation. Fifty frames of ilastik-generated background annotations were inspected and used as additional training inputs.

All training frames (2048 × 2048 pixels): human annotated frames and augmented 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.

U-Net was adapted from https://github.com/qubvel/segmentation_models, and a pretrained backbone of resnet34 (He *et al.*, 2016) was used to initialize weights. The model was trained end to end with standard cross entropy loss and Adam optimizer (Kingma and Ba, 2014). During training, weight scaling was employed to balance loss for different classes. Little hyperparameter search was performed due to the limited amount of annotations available for segmentation. The final segmentation model was evaluated visually based on criteria including quality of cell segmentation mask, prediction accuracy on microglia/contaminating cells, and robustness against imaging artifact. It should be noted that *DynaMorph* uses segmentation results to extract single-cell patches for which task segmentation accuracy is not a major concern. Thus we did not exhaustively validate and optimize segmentation performances to ease application-time usage. This step is also interchangeable with any off-the-shelf segmentation method/package that could provide an unbiased selection of target cells.

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 sliding window/tiling strategy; the model was then applied on individual patches, the results of which were tiled to generate the full prediction mask for this frame. To avoid the edge effect, 20 repetitions with different offsets were performed, and all 20 prediction masks were aligned and averaged to generate the final prediction.

*Instance segmentation.* 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 progenitor cells) by U-Net were extracted and then clustered based on their 2D coordinates in the frame. We used DBSCAN (Ester *et al.*, 1996) 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 (Redmon *et al.*, 2016; He *et al.*, 2017) would have better results.

Implementation from scikit-learn was applied, the parameters of which were tuned to separate adjacent cells with a small amount of contact. Masks with extreme sizes (>12,000 pixels or <500 pixels) were excluded to avoid 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 pixelwise probabilities of both classes (microglia and progenitor cells) across the whole cell mask. The resulting ratio of probabilities represented the proportion/likelihood that this specific cell is microglia or progenitor cells. Cells with microglia proportions larger than 0.9 were regarded as true microglia and vice versa for progenitor cells. The rest would be regarded as ambiguous cells. These identities helped in determining the validity of trajectories in the following tracking procedure.

*Tracking.* Within a full video, instance separation was performed on each static frame independently, the results of which were used in this step to connect cells across time and form trajectories. The tracking procedure followed the classic work of Jaqaman *et al.* (2008) in which the core architecture is a linear assignment problem (LAP). The instance separation step generated a list of cells along with their sizes and center positions were calculated from the predicted segmentation masks. Then a matching problem was set up between two lists of cells from adjacent time frames with a pairwise cost matrix calculated based on position displacements and shape differences between pairs of cells,

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

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

$d_{ij}$: 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 had a similar size. 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 Jaqaman *et al.* (2008) to connect short trajectories that were separated due to imaging/segmentation issues. A similar LAP setup with only squared distance cost was employed.

Identities of the trajectories obtained through the above procedure will be calculated by averaging static frame identities. Only trajectories with over 95% of all frames being labeled as microglia (proportion >0.9) were kept for the main analysis. In total, 5731 microglia trajectories were collected, 3715 of which were extracted from the test dataset. Meanwhile, trajectories with over 50% of the frames labeled as progenitor cells formed another small collection of 125 cells, which were used in the extension analysis.

***Self-supervised encoding with VQ-VAE.*** A variational autoencoder variant: VQ-VAE (van den Oord and Vinyals, 2017) was applied in this work to summarize the shape of cells through an encoding-decoding process. The model was trained on microglia cell patches (phase and retardance channels) extracted from the training set and directly applied to the test set data to generate shape 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 cells, we masked out other cells appearing in the patch area with median background values. Sample patches before the masking can be found in Figure 2. Samples after masking can be found in Supplemental Figure 2 - Supplement 2 and Supplemental Figure 2 - Supplement 5.

VQ-VAE was implemented in pytorch (Paszke *et al.*, 2017) using two deep convolutional networks as encoder and decoder, respectively. Input patches were first downscaled to $128 \times 128$ and normalized on both channels. The VAE encoder further downsampled the image by 64 folds, resulting in a latent vector for each cell with shape of 16, with channel as the last dimension. 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.

$$L_{\text{VQ-VAE}}(x) = \log p\left(x \mid z_q(x)\right) + \left\| sg\left[z_e(x)\right] - e_k \right\|_2^2 + \beta \left\| z_e(x) - sg\left[e_k\right] \right\|_2^2$$

where $k = \text{argmin}_j \left\| z_e(x) - e_j \right\|_2^2$, $z_q(x) = e_k$

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

$x$: 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 phase. After 100 epochs (or when the model could generate reasonable reconstructions), we added in $L_{\text{matching}}$ with a weight ratio of 0.0005 and trained the model until convergence.

The size of the latent space used in this study was consistent with the compression fold in its original work (van den Oord and Vinyals, 2017). Note that we did not push further to a much smaller latent space (i.e., <100 D) due to the concern that a 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 the 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, the explained variance of which is visualized in Supplemental Figure 2 - Supplement 7A.

We also tried fitting and transforming $z_q(x)$ (unpublished data) which generated very similar results and the same interpretations for the top PCs. Distance metrics defined with either $z_e(x)$ or $z_q(x)$ showed similar results (Supplemental Figure 2 - Supplement 9) 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 (Maaten and Hinton, 2008) and UMAP (McInnes *et al.*, 2018) were also tested and showed no advantage over PCA as no clustering patterns were detected.

***Gaussian mixture modeling of TFVs.*** To further unravel the morphodynamic features as well as to link transcriptomic profiles of microglia, clustering was performed on cell trajectories. We extracted 2-h-long segments from the trajectories and computed average shape and motion descriptors over segments to compute TFVs. The top 48 PCs derived from PCA of latent vectors were used as the shape descriptors. Distance traveled per frame along the trajectory was averaged and its log was used as the motion descriptor. The log mean distance values of segments were scaled to keep their variance comparable with the top two PCs of latent shape vectors.

We then used GMM to discover discrete states of TFVs. GMM enables interpretation of morphodynamic differences across perturbation as changes in the weight of each mixture component because GMM allows assignment of different weight distributions per mixture component for individual perturbations. We chose to use a two-component GMM after testing clustering robustness (measured

by Silhouette score) against the number of components. More specifically, we assumed two mixture components (two states) for microglia, and cells are sampled from one of the two components based on a Bernoulli distribution dependent on the perturbation condition.

Given the TFV $x_i$ and perturbation condition $y_i$ of a cell, the marginal probability of occurrence of $x_i$ is

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

where $\Theta = \left\{ \alpha_1^{(1)}, \alpha_2^{(1)}, \ldots, \alpha_1^{(4)}, \alpha_2^{(4)}, \theta_1, \theta_2 \right\}$ were parameters of the mixture model: $\theta_k$ is the center of $k$-th component, and $\alpha_k^{(j)}$ is the weight of $k$-th component in condition $j$. In the two-state model, $\alpha_1^{(j)}$ and $\alpha_2^{(j)}$ (summed to 1) formed the Bernoulli distribution for condition $j$. In the test set, four conditions were used so four sets of component weights were fitted with samples from the corresponding groups.

The expectation-maximization algorithm was employed to iteratively evaluate the parameters of GMM.

The E-step evaluates the membership weight of each cell based on the current set of parameters,

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

where $p_k(x_i | \theta_k) \propto \exp\left( -\frac{1}{2} \left( \frac{x_i - \theta_k}{s} \right)^2 \right)$

The M-step recalculates the parameters based on new membership weights,

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

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

The two steps above were repeated until parameters converged. In the procedure we fixed the variance of mixture components to be the same as the variance of TFVs to increase stability.

**HMM of TFVs.** Instead of using the smooth-then-cluster approach described above, we explored the identification of morphodynamic transitions per frame using a HMM.

Different from the GMM approach in which each trajectory is represented by a vector of the same size (TFV), HMM directly models the hidden state (cell state) of each time frame of the cell trajectory. To enable the comparison with GMM, we parameterized the HMM to be similar to GMM (2-state, same mean, and covariance priors).

The two states identified by HMM have similar morphodynamic properties as GMM: state-1 has lower PC2 and slower movements; state-2 has opposite values. We assigned a trajectory to state-1 (or state-2) if over 90% of all frames are predicted as the corresponding hidden state. Transition trajectories are identified based on opposing hidden states of the first and last 10 frames. Besides, there are 5 ~ 15% of all trajectories (in different perturbation conditions) that could not be assigned due to the fluctuating hidden state predictions along the sequence. We think this is due to the sensitivity of HMM to noise in imaging/tracking/embedding along each frame. Features of state-1 and state-2 trajectories are similar to states identified by GMM. The frequency of transition events identified by HMM is systematically 1 ~ 2% higher than GMM-based detection. Further tuning of HMM parameters may identify more states. Some smoothing measures could be preapplied (e.g., aggregate neighbor frames) to reduce the influence of noise.

## Predicting treatment conditions with TFVs

We explore the possibility of predicting perturbation conditions purely based on microglia's morphodynamic response using TFVs. In this task, TFVs are used to predict the perturbation condition in which the cell is cultured. A 10-fold cross-validation strategy is used, and a gradient boosted tree is selected for the prediction task after comparing multiple machine learning models. During the evaluation, the model (with default parameters) is trained with 9 folds of cell trajectories from the perturbation experiment and evaluated on the remaining fold (randomly split). In total 10 models are trained, and the results are aggregated to generate a confusion matrix reported in Supplemental Figure 3 - Supplement 2.

Among all four perturbation conditions, the model has the best accuracy in control subset cells, while the other three disease-relevant conditions are more intertwined. Confusion between the control subset and the IFN beta subset is the lowest, indicating a large morphodynamic difference between these two groups. This is also reflected by the GMM clustering outcome as the two subsets are composed of cells of different states. Both GBM and IL17 subsets have ~20% cells being predicted to come from the control group, which hints that these could be the proportion of microglia not affected by the stimuli.

The overall accuracy of this prediction task is 0.58 when using shape features from VQ-VAE with phase and retardance channel inputs, which could be compared with other model setups reported in Supplemental Figure 2 - Supplement 4D; Supplemental Figure 2 - Supplement 3D; and Supplemental Figure 3 - Supplement 3B. As this task evaluates how embedding differentiates between perturbations, the accuracy score could partially reflect the embedding quality.

## scRNA sequencing

Cells from three conditions (control, IFN beta and GBM supernatant) were dissociated using 0.25% trypsin solution, labeled with Multiseq barcodes (McGinnis *et al.*, 2019), and processed for single-cell mRNA sequencing using the 10× 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 (Hafemeister and Satija, 2019) was applied to the raw count matrix followed by PCA on the residuals. The top 30 PCs were used to find neighbors for Leiden clustering (Traag *et al.*, 2019) and the UMAP (McInnes *et al.*, 2018) projection.

## CODE AND DATA 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/mehta-lab/dynamorph/. The refined version of label-free acquisition and analysis software is available as napari plugins at https://github.com/mehta-lab/recOrder.

## ACKNOWLEDGMENTS

## REFERENCES

Aw E, Zhang Y, Carroll M (2020). Microglial responses to peripheral type 1 interferon. J Neuroinflammation 17, 1–13.

Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C, Schiegg M, Ales J, Beier T, Rudy M, et al. (2019). Ilastik: Interactive machine learning for (bio)image analysis. Nat Methods 16, 1226–1232, doi: https://doi.org/10.1038/s41592-019-0582-9.

Bernier L-P, Bohlen CJ, York EM, Choi HB, Kamyabi A, Dissing-Olesen L, Hefendehl JK, Collins HY, Stevens B, Barres BA, et al. (2019). Nanoscale surveillance of the brain by microglia via camp-regulated filopodia. Cell Rep 27, 2895–2908.

Blank T, Prinz M (2017). Type i interferon pathway in cns homeostasis and neurological disorders. Glia 65, 1397–1406.

Bray M-A, Singh S, Han H, Davis CT, Borgeson B, Hartland C, Kost-Alimova M, Gustafsdottir SM, Gibson CC, Carpenter AE (2016). Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. Nat Protoc 11, 1757–1774, https://doi.org/10.1038/nprot.2016.105.

Carpenter AE, Sabatini DM (2004). Systematic genome-wide screens of gene function. Nat Rev Genet 5, 11–22. https://doi.org/10.1038/nrg1248.

Chan CK, Hadjitheodorou A, Tsai TY-C, Theriot JA (2020). Quantitative comparison of principal component analysis and unsupervised deep learning using variational autoencoders for shape analysis of motile cells. bioRxiv https://doi.org/10.1101/2020.06.26.174474.

Chhatbar C, Detje CN, Grabski E, Borst K, Spanier J, Ghita L, Elliott DA, Jordao MJC, Mueller N, Sutton J, et al. (2018). Type i interferon receptor signaling of neurons and astrocytes regulates microglia activation during viral encephalitis. Cell Rep 25, 118–129.

Choi GB, Yim YS, Wong H, Kim S, Kim H, Kim SV, Hoeffer CA, Littman DR, Huh JR (2016). The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring. Science 351, 933–939.

Christiansen EM, Yang SJ, Ando DM, Javaherian A, Skibinski G, Lipnick S, Mount E, O'Neil A, Shah K, Lee AK, et al. (2018). In silico labeling: Predicting fluorescent labels in unlabeled images. Cell 173, 792–803. e19, https://doi.org/10.1016/j.cell.2018.03.040.

Driscoll MK, Welf ES, Jamieson AR, Dean KM, Isogai T, Fiolka R, Danuser G (2019). Robust and automated detection of subcellular morphological motifs in 3D microscopy images. Nat Methods 16, 1037–1044. https://doi.org/10.1038/s41592-019-0539-z.

Ester M, Kriegel H-P, Sander J, Xu X (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Knowledge Discovery and Data Mining 96, Portland, OR: AAAI, KDD 96, 226–231. https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf.

Fekete R, Cserép C, Lénárt N, Tóth K, Orsolits B, Martinecz B, Méhes E, Szabó B, Németh V, Gönci B, et al. (2018). Microglia control the spread of neurotropic virus infection via p2y12 signalling and recruit monocytes through p2y12-independent mechanisms. Acta Neuropathol 136, 461–482.

Friedman JH (2001). Greedy function approximation: a gradient boosting machine. Ann Stat 1189–1232.

Gerbin KA, Grancharova T, Donovan-Maiye R, Hendershott MC, Brown J, Dinh SQ, Gehring JL, Hirano M, Johnson GR, Nath A, et al. (2021). Cell states beyond transcriptomics: Integrating structural organization and gene expression in hiPSC-derived cardiomyocytes. Cell Systems 12, 670–687.e10. https://doi.org/10.1016/j.cels.2021.05.001.

Gosselin D, Skola D, Coufal NG, Holtman IR, Schlachetzki JC, Sajti E, Jaeger BN, O'Connor C, Fitzpatrick C, Pasillas MP, et al. (2017). An environment-dependent transcriptional network specifies human microglia identity. Science 356, eaal3222.

Guo S-M, Yeh L-H, Folkesson J, Ivanov IE, Krishnan AP, Keefe MG, Hashemi E, Shin D, Chhun BB, Cho NH, et al. (2020). Revealing architectural order with quantitative label-free imaging and deep learning. eLife 9, e55502. https://doi.org/10.7554/eLife.55502.

Hafemeister C, Satija R (2019). Normalization and variance stabilization of single-cell rna-seq data using regularized negative binomial regression. Genome Biol 20, 1–15.

Hansen DV, Lui JH, Parker PR, Kriegstein AR (2010). Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature 464, 554–561.

He B, Bergenstråhle L, Stenbeck L, Abid D, Andersson A, Borg Å, Maaskola J, Lundeberg J, Zou J (2020). Integrating spatial gene expression and breast tumour morphology via deep learning. Nature Biomedical Engineering 1–8, https://doi.org/10.1038/s41551-020-0578-x.

He K, Zhang X, Ren S, Sun J (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.

He K, Gkioxari G, Dollár P, Girshick R (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision, 2961–2969.

Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, Danuser G (2008). Robust single-particle tracking in live-cell time-lapse sequences. Nat Methods 5, 695–702. ISSN 1548-7105. doi: https://doi.org/10.1038/nmeth.1237.

Kawanokuchi J, Shimizu K, Nitta A, Yamada K, Mizuno T, Takeuchi H, Suzumura A (2008). Production and functions of il-17 in microglia. J Neuroimmunol 194, 54–61.

Keren K, Pincus Z, Allen GM, Barnhart EL, Marriott G, Mogilner A, Theriot JA (2008). Mechanism of shape determination in motile cells. Nature 453(7194), 475–480. https://doi.org/10.1038/nature06952.

Kimmel JC, Chang AY, Brack AS, Marshall WF (2018). Inferring cell state by quantitative motility analysis reveals a dynamic state system and broken detailed balance. PLoS Comput Biol 14, e1005927. https://doi.org/10.1371/journal.pcbi.1005927.

Kingma DP, Ba J (2014). Adam: A method for stochastic optimization. In 3rd International Conference on Learning Representations, San Diego, CA: ICLR 2015.

Kobayashi H, Cheveralls KC, Leonetti MD, Royer LA (2021). Self-supervised deep-learning encodes high-resolution features of protein subcellular localization. bioRxiv, 2021.03.29.437595, https://doi.org/10.1101/2021.03.29.437595.

Lane K, Van Valen D, DeFelice MM, Macklin DN, Kudo T, Jaimovich A, Carr A, Meyer T, Pe'er D, Boutet SC, Covert MW (2017). Measuring signaling and RNA-seq in the same cell links gene expression to dynamic patterns of NF-κB activation. Cell Syst 4, 458–469.e5. https://doi.org/10.1016/j.cels.2017.03.010.

Lively S, Schlichter LC (2018). Microglia responses to pro-inflammatory stimuli (lps, ifnγ+ tnfα) and reprogramming by resolving cytokines (il-4, il-10). Front Cell Neurosci 12, 215.

Maaten LVD, Hinton G (2008). Visualizing data using t-sne. J Mach Learn Res 9(Nov), 2579–2605.

Marshall WF (2011). Origins of cellular geometry. BMC Biol 9, 57. https://doi.org/10.1186/1741-7007-9-57.

McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V, Hu JL, Murrow LM, Weissman JS, Werb Z, et al. (2019). Multi-seq: sample multiplexing for single-cell rna sequencing using lipid-tagged indices. Nat Methods 16, 619.

McInnes L, Healy J, Melville J (2018). Umap: Uniform manifold approximation and projection for dimension reduction. ArXiv preprint arxiv:1802.03426.

Moen E, Borba E, Miller G, Schwartz M, Bannon D, Koe N, Camplisson I, Kyme D, Pavelchek C, Price T, et al. (2019). Accurate cell tracking and lineage construction in live-cell imaging experiments with deep learning. bioRxiv 803205, https://doi.org/10.1101/803205.

Mudò G, Frinchi M, Nuzzo D, Scaduto P, Plescia F, Massenti MF, Di Carlo M, Cannizzaro C, Cassata G, Cicero L, et al. (2019). Anti-inflammatory and cognitive effects of interferon-β1a (ifnβ1a) in a rat model of alzheimer's disease. J Neuroinflammation 16, 1–16.

Neumann B, Walter T, Hériché J-K, Bulkescher J, Erfle H, Conrad C, Rogers P, Poser I, Held M, Liebel U, et al. (2010). Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature 464(7289), 721–727. https://doi.org/10.1038/nature08869.

Nimmerjahn A, Kirchhoff F, Helmchen F (2005). Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo. Science 308, 1314–1318.

Noctor SC, Penna E, Shepherd H, Chelson C, Barger N, Martínez-Cerdeño V, Tarantal AF (2019). Periventricular microglial cells interact with dividing precursor cells in the nonhuman primate and rodent prenatal cerebral cortex. J Comp Neurol 527, 1598–1609.

Ounkomol C, Seshamani S, Maleckar MM, Collman F, Johnson GR (2018). Label-free prediction of three-dimensional fluorescence images from

transmitted-light microscopy. Nat Methods 15, 917. https://doi.org/10.1038/s41592-018-0111-2.

Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, Lin Z, Desmaison A, Antiga L, Lerer A (2017). Automatic differentiation in pytorch. In NIPS-W.

Pincus Z, Theriot JA (2007). Comparison of quantitative methods for cell-shape analysis. J Microsc 227, 140–156. https://doi.org/10.1111/j.1365-2818.2007.01799.x.

Razavi A, van den Oord A, Vinyals O (2019). Generating diverse high-fidelity images with VQ-VAE-2. In Advances in Neural Information Processing Systems, ed. H Wallach, H Larochelle, A Beygelzimer, F Alché-Buc, E Fox, and R Garnett, Vol. 32. Curran Associates, Inc., 14866–14876.

Redmon J, Divvala S, Girshick R, Farhadi A (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, 779–788.

Rezaie P, Trillo-Pazos G, Greenwood J, Everall IP, Male DK (2002). Motility and ramification of human fetal microglia in culture: An investigation using time-lapse video microscopy and image analysis. Exp Cell Res 274, 68–82. https://doi.org/10.1006/excr.2001.5431.

Ronneberger O, Fischer P, Brox T (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, 234–241. Springer.

Salter MW, Stevens B (2017). Microglia emerge as central players in brain disease. Nat Med 23, 1018.

Schroff F, Kalenichenko D, Philbin J (2015). Facenet: A unified embedding for face recognition and clustering. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 815–823. https://doi.org/10.1109/CVPR.2015.7298682

Shadkhoo S, Mani M (2019). The role of intracellular interactions in the collective polarization of tissues and its interplay with cellular geometry. PLoS Comput Biol 15, e1007454. https://doi.org/10.1371/journal.pcbi.1007454.

Traag VA, Waltman L, van Eck NJ (2019). From louvain to leiden: guaranteeing well-connected communities. Sci Rep 9, 1–12.

van den Oord A, Vinyals O (2017). Neural discrete representation learning. In Advances in Neural Information Processing Systems, 6306–6315.

Van Valen DA, Kudo T, Lane KM, Macklin DN, Quach NT, DeFelice MM, Maayan I, Tanouchi Y, Ashley EA, Covert MW (2016). Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. PLoS Comput Biol 12, e1005177. https://doi.org/10.1371/journal.pcbi.1005177.

Way GP, Natoli T, Adeboye A, Litichevskiy L, Yang A, Lu X, Caicedo JC, Cimini BA, Karhohs K, Logan DJ, et al. (2021). Morphology and gene expression profiling provide complementary information for mapping cell state. bioRxiv. https://www.biorxiv.org/content/early/2021/10/22/2021.10.21.465335.

Yang KD, Belyaeva A, Venkatachalapathy S, Damodaran K, Radhakrishnan A, Katcoff A, Shivashankar GV, Uhler C (2021). Multi-domain translation between single-cell imaging and sequencing data using autoencoders. Nat Commun 12, 31. https://doi.org/10.1038/s41467-020-20249-2.

Yu A, Duan H, Zhang T, Pan Y, Kou Z, Zhang X, Lu Y, Wang S, Yang Z (2016). Il-17a promotes microglial activation and neuroinflammation in mouse models of intracerebral haemorrhage. Mol Immunol 73, 151–157.

Zaritsky A, Welf ES, Tseng Y-Y, Rabadán MA, Serra-Picamal X, Trepat X, Danuser G (2015). Seeds of locally aligned motion and stress coordinate a collective cell migration. Biophys J 109, 2492–2500. https://doi.org/10.1016/j.bpj.2015.11.001.

Zaritsky A, Jamieson AR, Welf ES, Nevarez A, Cillay J, Eskiocak U, Cantarel BL, Danuser G (2020). Interpretable deep learning of label-free live cell images uncovers functional hallmarks of highly-metastatic melanoma. https://doi.org/10.1101/2020.05.15.096628.

Zou W, Zhu S, Yu K, Ng AY (2012). Deep learning of invariant features via simulated fixations in video. In: Advances in Neural Information Processing Systems, Lake Tahoe, CA: Curran Associates, Inc. 3203–3211. https://proceedings.neurips.cc/paper/2012/file/13d7dc096493e1f77fb4ccf3eaf79df1-Paper.pdf.