Electronic Supplementary Material (ESI) for Integrative Biology

Supplementary Information: Appendix

Endothelial cell phenotypic behaviors cluster into dynamic transition programs modulated by angiogenic and angiostatic cytokines

Tharathorn Rimchala[a], Roger D. Kamm[a,b], Douglas A. Lauffenburger[a,c,1]

[a]Department of Biological Engineering,
Massachusetts Institute of Technology
77 Massachusetts Ave.,
Cambridge, MA 02139.

[b]Department of Mechanical Engineering,
Massachusetts Institute of Technology
77 Massachusetts Ave.,
Cambridge, MA 02139.

[c]Department of Biology,
Massachusetts Institute of Technology
77 Massachusetts Ave.,
Cambridge, MA 02139.

[1]To whom correspondence should be addressed. Email: lauffen@mit.edu, Tel: (617) 252-1629, Fax: (617) 258-0204.

1 Supplementary Methods

1.1 Image processing and contour tracking algorithm

In contour tracking of the live cell images, the raw images are converted to from 16-bit to 8-bit to enhance the processing speed. To even out the cell-to-cell variations in brightness, the images are minimally enhanced using median filtering and entropy filtering routines in MATLAB R2011bs Image processing toolbox. The first image in the time-series is initialized either by 1. using intensity of the image or 2. using non-informative mask. In most cases, the two methods of contour initialization yields reasonably good agreement (determined by the square difference in the contour point coordinates). In rare case when the two methods do not yield similar contours, the image is enhanced and the contour is initialized using the intensity thresholded mask of the enhanced image. For subsequent images, the detected contour of the immediate prior image is use to initialize contour in the contour finding routine. Given that in most of the time series images cells do not drastically change shape and size over one imaging interval, the detected contour from the prior image often serves as an appropriate initial contour. In the traditional naive active contour algorithm, the objective function is defined by the gradient in image intensity alone. For single cell tracking application, however, the active contours performance can be significantly improved by using the gradient in intensity and the detected contour of the neighboring images in the time series to bias the contours objective function. The image preprocessing subroutines and subroutine for parsing the contour mask as initial contour for the level set optimization of image in the subsequent time step were developed by the authors. The level set active contour method is proposed and developed by Tony F. Chan and Luminita A. Vese ([1]). The level set routine was adopted from Yue Wus contribution on publicly accessible MathWorks file exchange website (http://www.mathworks.com/matlabcentral/fileexchange/23445) and slightly modify the level sets objective function so the algorithm works well for our images.

1.2 Semi supervised sessile vs. motile state classification

In our dataset, there are more than 500,000 contour instances that need to classified into either sessile or motile state. To meet the challenge of this classification task, we take a semi-supervised learning approach. First, we generate *state-labeled data* by sampling 10 non-overlapped trajectories training sets (about 2% of the total trajectories) and clustering them using agglomerative hierarchical clustering algorithm based on Euclidean separation of the contour instances in feature space. Based on their cluster assignment, these individual instances are labeled S or M. We then use this state-labeled training set to train 50 base classifiers (decision stumps) using AdaBoost algorithm ([2, 3]). To evaluate the classifier performance, we perform K fold cross validation (with K varying from 2 to 15) by subdividing the labeled data into K smaller chunks and use the first 1/K fraction to train the based classifier. We then evaluate the error rate of the ensemble classifier on the other K-1/K fraction of the labeled data (SI Appendix Fig S3). We find that most K-fold cross validated ensemble classifiers classify about 8% of the data incorrectly. More important, the test error rates are comparable to the corresponding training error rates of the ensemble classifiers, suggesting that a small fraction of the labeled data cannot be corrected classified with the ensemble classifier. The cross validated ensemble classifier is used to classify the rest of the contour instances and the classification results were visualized against the contour traces to ensure that the classification result follow definitions of sessile and migratory states. The subroutines for sampling the trajectories and for generating the labeled data is developed by the authors. The adaptive boosting algorithm is implemented in MATLAB by by Dirk-Jan Kroon and is publicly available through his MathWorks file exchange page (http://www.mathworks.com/matlabcentral/fileexchange/27813-classic-adaboost-classifier). The agglomerative hierarchical clustering routine is available through MATLAB R2011as Bioinformatics Toolbox. The PCA routine used in data analysis in this manuscript is developed by Laurens van der Maaten and is publicly downloadable as part of the toolbox for Dimensionality Reduction from the following website: (http://homepage.tudelft.nl/19j49/Matlab_Toolbox_for_Dimensionality_Reduction.html).

1.3 Angiogenesis sprouting assay in a high throughput microfluidic device (HTD)

PDMS device preparation

High throughput microfluidic photoresist pattarned silicon wafer mold was designed in house and custom-ordered from the Stanford University Microfluidic Foundry. The microfluidic system consisting of PDMS (polydimethylsiloxane; Silgard Dow Chemical, MI; Cat.No. 184) was prepared on SU-8 2050 photoresist-patterned wafers (MicroChem, MA) using a standard soft lithography process described previously ([4, 5]). The fabricated PDMS channel and the microscopy grade cover slip used to seal the channel were sterilized and dried at 80°C overnight. Subsequently, they were plasma treated (Harrick, CA) in air, and bonded together to form a closed microfludic channel. After the plasma bonding, all microfluidic channels were coated with 1 mg/mL poly-D-lysine hydrobromide (Sigma-Aldrich St. Loius, MO; Cat.No. P7886) and incubated for at least 4 hours at 37°C in a humidified environment. The device was then washed thoroughly with sterile water and dried at 80°C overnight to allow the PDMS surface to return to its native hydrophobicity - a crucial surface property in confining the extracellular matrix within a specified region.

Extracellular matrix casting and cell seeding

Microfluidic device that have been bonded, sterilized and surface treated were brought to room temperature prior to gel injection. Type I rat tail collagen is diluted to 2.0 mg/mL concentration and calibrated to pH 7.4 as in the on-gel sprouting assay. While at 4°C, the collagen gel solution was carefully injected into the microfluidic gel region through a gel filling port using a standard 200 μL micropipette tip. The collagen gel was allowed to solidify at 37°C in a humidified chamber for at least one hour. After gel solidification, 37°C cell culture medium was flown into the device on both sides of the gel through the medium ports. The gel was incubated with the cell culture medium for at least one hour before cell seeding. At the cell seeding time, hMVECs and HUVECs cell suspensions were diluted to the instant monolayer seeding density, flown into the channel, and allowed to adhered for at least one hour prior to additional medium filling.

Inflammatory cytokine treatment

After at least 24 hour of seeding in cell culture medium (EGM2MV; Lonza NJ Cat.No. CC-3202), hMVEC culture were switched to conditioned medium contain- ing specified concentrations of recombinant human VEGF and PF4 (Peprotech NJ; Cat.No. 100-20 and 300-16 respectively). Conditioned media were refreshed every 24 hours onward.

Angiogenic sprout visualization and quantification

Sprouting endothelial cells in HTD were visualized under a phase contrast microscope every 24 hours after seeding. Images were taken and analyzed using an image processing MATLAB script developed in house. At the end point of the assay, 3D images of DAPI and Alexa-568 Phalloidin (Molecular Probes, Eugene, OR; Cat.No. A12380) stained samples (in ongel and HTD setups) were imaged using a laser scanning microscopes (Zeiss LSM510 and Olympus FV1000).

1.4 Hierarchical clustering of single cell state trajectories and phenotypic cluster evaluation

The likelihood function of single cell state trajectories serves as an objective function for inferring the maximum likelihood estimates and Bayesian inference of the phenotypic state transition rates. The two sets of parameters that determine the likelihood function which in turn determine the parameter estimates are:{a} $f_{ss'}$ the trajectory length normalized frequencies of the transition from s to s' states; and {b} $\sum t_s$ the total waiting time in a particular state s for that trajectory. To investigate the similarities or differences in the likelihood functions of all the trajectories, we compute $f_{ss'}$ and $\sum t_s$ of each trajectories and use them as

classification features.

We perform hierarchical clustering of the single cell trajectories using an agglomerative clustering routine clustergram, which is available through MATLAB R2011bs Bioinformatics Toolbox. For all the cytokine conditions investigated in this study, we found that the clustergram routine yields clustering pattern that follow the phenotypic behavior of single cells within the clusters (SI Appendix Fig S5), which we refer to as *'phenotypic program based'* clustering pattern (grouping).

In clustering analysis, the main criteria used to evaluate the goodness of the clustering result are: 1. compactness, 2. separation, and 3. partition fuzziness. Better clustering results is characterized by higher level of compactness (cluster members should be as close as possible), higher level of separation (distinct clusters should be separated as widely as possible), and lower level of fuzziness. To validate the phenotypic program based grouping, we compute the mean intra-cluster spread, mean inter-group distances, and mean classification entropies as scalar metrics of compactness, separation and fuzziness of the clustering results respectively. In our study, intra-cluster spread is the pairwise distance of all data points within a cluster, inter-cluster separation is the pairwise Euclidean distance between two cluster centers, and classification entropy is the degree of uncertainty in cluster membership defined in an information theoretic sense. The classification entropy is Shannons information entropy in which the probability of the uncertain random variable is cluster membership.

We show that the phenotypic program based clustering is more compact, better separated, and better partitioned than the condition based clustering (SI Appendix Fig S7). As an alternative to hard clustering in which each data point belongs to exactly one cluster, cluster partition can be *'fuzzy'*. Fuzzy partition allows each data point to be assigned to different clusters with varying degree of cluster membership – the degree to which a data point associates with a particular cluster. By comparing the cluster membership of all the data point assigned different clusters in the condition based grouping and phenotypic program based group, we show that the phenotypic based grouping allows more distinct cluster assignment (SI Appendix Fig S7), suggesting that phenotypic based grouping is better way of clustering the data than the condition based one.

1.5 Pairwise statistical comparisons by Kolmogorov Smirnov test

Pairwise comparison were performed most extensively in two tasks: 1. comparing the condition- based rate MLEs ($\lambda^{(\mathrm{cond})}$) across cytokine conditions (Fig S5ab) and 2. comparing the cluster weights MLE across cytokine conditions (Fig S12a-c). In comparing $\lambda^{(\mathrm{cond})}$, 1000 bootstrapped samples of 50 single cell MLE trajectories were drawn from the pool of trajectories within each condition. Maximum likelihood of $\lambda^{(\mathrm{cond})}$) were computed from the sampled trajectories to form $\lambda^{(\mathrm{cond})}$ distributions. In comparing the cluster weights, 1000 bootstrapped samples of 50 trajectories were drawn from the trajectories in each condition. The trajectories were assigned to one of the five state transition dynamic clusters based on the relative Mahalonobis distances of the trajectories to all the cluster centers. The cluster weights are computed for each bootstrapped sample to form the distribution of cluster weights. In both of the pairwise statistical comparison task, both the $\lambda^{(\mathrm{cond})}$ distributions and cluster weight distributions across conditions are compared using MLE Kolmogorov-Smirnov test with the significance level of 0.05.

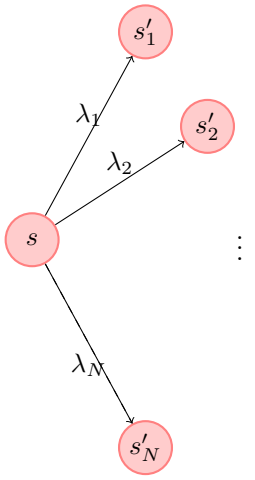
2 Supplementary Modeling Approaches

2.1 Modeling single cell state trajectories as continuous time Markov chains (CTMCs)

In this work, we model individual cell as a decision making entity called Markov agent that transition among a finite number of phenotypic states. As we follow individual agent over time, we can trace out a sequence of states through which the agent traverses as well as the corresponding waiting times before each transition. We refer to the observed sequence as a cell's state trajectory. In choosing a stochastic model to describe the state transition of angiogenic endothelial cells, we showed that state trajectories satisfy the two Markov criteria can be modeled as a continuous time Markov chain: 1. memorylessness and 2. conditional independence properties (SI Appendix Fig S4).

A continuous time Markov chain (CTMC) is defined by the following descriptors: (1) a finite state set $\mathbb{S}$, (2) initial (marginal) state probabilities, (3) transition probabilities, and (4) state waiting time parameter. In the case of angiogenic endothelial cells, the appropriate set of phenotypic states are sessile (S), proliferative (P), migratory (M), and apoptotic (A). In the following section, we construct the likelihood expression of a single cell state trajectories from which the state transition rate parameters can be optimized.

Likelihood of one transition



To construct an analytical expression for the likelihood function, we first derive the probability of an occurrence of a state transition. Consider a one step transition from s to a finite number of state reachable from s'_n shown below. The transition to state $s' \neq s$ happens at an exponentially distributed random time with rate parameter $\mu = \lambda_1 + \ldots + \lambda_N$. At the transition time, the new state s' is chosen with the probabilitiy

$$p_{s'} = \frac{\lambda_{s'}}{\lambda_1 + \ldots + \lambda_N} = \frac{\lambda_{s'}}{\mu}.$$

Given the transition rate parameter set $\Lambda = \{\lambda_s\}$ and the waiting time parameter μ and assuming that the process is in state s initially, the likelihood of the observing a transition ss' is given by

$$\begin{aligned} \ell(s_{k+1} = s' | s_k = s; T = \tau; \Lambda_s) &= Pr(\text{dwelling in } s_i \text{ for } t_i) \times \\ &\quad Pr(\text{transitioning from } s \text{ to } s') \\ &= e^{-\mu_s t_i} \times \frac{\lambda_{ss'}}{\mu_s}. \end{aligned}$$

Likelihood of one state trajectory

As the next step, consider an experimentally observed single cell state trajectory as a sequence of state transitions. Let $\mathcal{U} = (s_0, t_0, s_1, t_1, \ldots, s_{k-1}, t_{k-1}, s_k)$ denotes the set of random variables describing a CTMC of single cell state trajectory up to time t and let $\ell_{ss'}(t)$ represents the likelihood of ss' type transition at time t. Under the CTMC assumption, individual transition are independent of one another and the likelihood of a state trajectory $\mathcal{U}$ is simply the product of individual transition in the trajectory. As such, the likelihood of a particular state trajectory with η transitions is given by:

$$\ell(\mathcal{U}_t | s_o, \Lambda) = P_{s_o} \prod_{i=1}^{\eta} \ell_{s_i s_{i+1}}(t_i),$$

where P_{s_o} is the initial probability of finding the process in state s_o initially.

We can further simplify the likelihood expression as follows. Let $\eta_{ss'}$ be the total number ss' type transitions in a trajectory and let $\mathbb{H}$ be the set of all transition types. Since the transitions in a trajectory are independence, one can factorize the above likelihood expression based on the transition types. The resulting likelihood expression is given by

$$\begin{aligned}\ell(\mathcal{U}_t|s_o,\Lambda) &= P_{s_o}\prod_{ss'\in\mathbb{H}}\Big(\prod_{h_{ss'}=1}^{\eta_{ss'}}\ell(ss',\tau_{h_{ss'}})\Big)\\ &= P_{s_o}\prod_{ss'\in\mathbb{H}}\left(\frac{\lambda_{ss'}}{\mu_s}\right)^{\eta_{ss'}}\exp(-\mu_s\sum_{h_{ss'}=1}^{\eta_{ss'}}\tau_{h_{ss'}}).\end{aligned}$$

To obtain the generalized likelihood expression for the entire observed population, one assumes independence of state transition among cells in the population, in which case the joint likelihood is simply the product of the likelihood of all trajectories within the population.

Estimation of the transition rate parameters

To obtain the state transition rate estimates from the data, we rely on two parameter estimation techniques: Maximum likelihood estimation, and Bayesian estimation. Both of these estimation methods find parameter values (in MLE case) or posterior rate distribution of the parameter (in BE case) that are most consistent with the observation as described by the likelihood distribution.

Maximum Likelihood Estimation

Consider a set of state trajectories $\mathbb{U}=\{\mathcal{U}=(s_o,s_1,t_1,\ldots s_{\eta-1},t_{\eta-1},s_\eta)\}$. To find the parameter set that is most consistent with the observed trajectories, we seek to optimize the above likelihood function for a collection of state trajectories subject to the following constraints:

$$\begin{aligned}\sum_{s'}\lambda_{ss'} &= \mu_s, \qquad \forall s \qquad \text{and}\\ \lambda_{ss'} &\geq 0, \qquad \forall ss'\in\mathbb{H}.\end{aligned}$$

Since it is more convenient to optimize the logarithm of likelihood, we set up the optimization in term of log likelihood using the Lagrange's method:

$$\begin{aligned}\underset{(\lambda_{ss'})\in\Lambda}{\operatorname{argmax}}\log(\ell(\Lambda)) &= \underset{(\lambda_{ss'})\in\Lambda}{\operatorname{argmax}}\,\mathcal{L}\\ &= \underset{(\lambda_{ss'})\in\Lambda}{\operatorname{argmax}}\left[\log(P_{s_o})+\sum_{ss'\in\mathbb{H}}\left(\eta_{ss'}\log\Big(\frac{\lambda_{ss'}}{\mu_s}\Big)-\mu_s\sum_{h_{ss'}=1}^{\eta_{ss'}}t_{h_{ss'}}\right)-\sum_s\left(\zeta_s(\sum_{ss'}\lambda_{ss'}-\mu_s)\right)\right],\end{aligned}$$

where $\zeta_{ss'}$ are the Lagrange's multipliers. For each of the rate parameter $\lambda_{ss'}$, we take the derivatives of the log likelihood with respect to $\lambda_{ss'},\mu_s,and\zeta_{ss'}$ and set them to zero. The resulting system of equations take the form:

$$\begin{aligned}\frac{\partial}{\partial\lambda_{ss'}} &= 0 = \frac{\eta_{ss'}}{\lambda_{ss'}}-\zeta_s \quad ,\\ \frac{\partial}{\partial\mu_s} &= 0 = -\frac{\eta_{ss'}}{\mu_s}-\sum_{h_{ss'}=1}^{\eta_{ss'}}t_{h_{ss'}}+\zeta_s \quad ,\\ \frac{\partial}{\partial\zeta_s} &= 0 = \sum_s\lambda_{ss'}-\mu_s.\end{aligned}$$

Assuming that $\lambda_{ss'} > 0$, we rearrange the above expression and to obtain the maximum likelihood estimates of the parameters:

$$\begin{aligned} \lambda_{ss'}^{\mathrm{MLE}} &= \frac{\eta_{ss'}}{\sum_{h_{ss'}} t_{h_{ss'}}} \left(1 - \frac{\eta_{ss'}}{\sum_s \eta_{ss'}}\right) \quad , \\ \mu_s^{\mathrm{MLE}} &= \frac{\sum_{s'} \eta_{ss'}}{\sum_{h_{ss'}} t_{h_{ss'}}} \left(1 - \frac{\eta_{ss'}}{\sum_s \eta_{ss'}}\right) \quad , \\ \zeta_s &= \frac{\sum_{h_{ss'}} t_{h_{ss'}}}{1 - \frac{\eta_{ss'}}{\sum_{s'} \eta_{ss'}}}. \end{aligned}$$

The above maximum likelihood estimators can be easily applied to multiple trajectories (i.e. a subpopulation of multiple cells) by extending the summation of log likelihood over all trajectories $\mathbb{U} = \{\mathcal{U}\}$.

Bayesian estimation

To estimate the posterior distribution of the rate parameter we rely on the Bayes' theorem which posits that the posterior distribution of the parameters given the evidence (observed data) equals the likelihood of the observed data given the parameters weighted by the evidence (marginal probability of the parameter), i.e.

$$\begin{aligned} P(\Lambda|\mathbb{U} = \{\mathcal{U}\}) &= \frac{P(\mathbb{U}|\Lambda) \times P(\Lambda)}{P(\mathbb{U})} \\ , &= \frac{P(\mathbb{U}|\Lambda) \times P(\Lambda)}{\int_\Lambda P(\mathbb{U}|\Lambda) \times P(\Lambda)}. \end{aligned}$$

2.2 Evaluation of the continuous time Markov chain criteria and application for modeling phenotypic state transition data

Phenotypic state trajectories of a single cell can be represented a sequence of time-indexed random variables. To determine if these trajectories can be represented by a Continuous time Markov chain, we evaluate whether they satisfy the essential properties of a continuous time Markov process: 1. Exponential waiting time (memorylessness) and 2. Markov property (conditional independence).

2.2.1 Waiting time distribution

We acquire the waiting time distribution by computing the dwell time within a particular state phenotypic state from all the single cell state trajectories. We observe that the waiting time distribution of instances in state S and M fit relatively well to exponential waiting time distribution with the goodness of fit of 0.73 for instances S state and 0.97 for instances in M state. We are not able to obtain a reliable fit of the waiting time distribution of instances in P, partially due to the small number of instances. For the A state instances, the notion of waiting time (before transition) does not apply because trajectories terminate after transitions to A state.

2.2.2 Conditional independence assumption (Markov property)

We first investigate the conditional independent assumption within smallest fragments of state trajectories containing consecutive state transitions. These are three state fragments of state trajectories. To determine whether our data satisfy the Markov property, we compare the likelihood distributions of the three state fragment data predicted by either the model without the conditional independence assumption (full dependence model) and without the conditional independence assumption (conditional independence model). The

full dependence model predicts that these likelihood of a three state fragment $S_1S_2S_3$ is given by

$$P(S_3|S_2,S_1) \quad = \quad \frac{N(S_1,S_2,S_3)}{N(S_2,S_3)},$$

where N are the occurrences of specified fragments in the data. On the other hand, the conditional independent model predicts that

$$P(S_3|S_2,S_1) \quad = \quad P(S_3|S_2)\times P(S_2|S_1) \quad = \quad \frac{N(S_2,S_3)}{N(S_2)}\times\frac{N(S_1,S_2)}{N(S_1)}.$$

We can estimate the likelihood distribution from the occurrence of these fragments in the data. If the conditional independent model predicts a statistically similar likelihood distribution to the full dependence model, then we conclude that the conditional independence assumption well approximate the observed state transitions within the single cell state trajectories data set and that our single cell state trajectories follow the conditional independent assumption. To measure the differences between the likelihood distributions predicted by the full dependence and the conditional independence models, we compare the symmetric Jensen-Shannon divergence (JSD) between the two distributions against JSD of computationally generated single cell state trajectories (background data generated from a full independence model in which the three state fragments are fully independent and there is no inherent state transition patterns). Since the background data set confers no dependence among the subsequent phenotypic transitions, the likelihood distributions of this background dataset is consistent to both full dependence and conditional independence models ($P(S_1|S_2,S_3) = P(S_3|S_2)\times P(S_2|S_1)$ in both cases). We show that the average JSD of the data is comparable to or smaller than of the background (independent transition data), suggesting that the two models yields statistically the same likelihood distribution (SI Appendix Fig S4 c-d).

2.3 Switch model parameter optimization and model selection

To model the switch-like sprouting in response to VEGF and PF4, sprout density data were typically fitted to a four-parameter Hill equation or hyperbolic Tangent switching equation with respective to one cytokine. For the four parameter Hill equation, the predicted sprout response with changing one cytokine is given by

$$\begin{aligned} F_v^{\mathrm{Hill}}(V|P) &= a_{0,V} + a_{1,V}\Big(\frac{V^{h_v}}{a_{2,V}^{h_v}+V^{h_v}}\Big); \\ F_p^{\mathrm{Hill}}(P|V) &= a_{0,P} + a_{1,P}\Big(\frac{1}{a_{2,P}^{h_p}+P^{h_p}}\Big), \end{aligned}$$

where a_0 denotes basal response, a_1 is a lump parameter representing the effective maximal strength of the response, a_2 is a lump parameter representing effective binding/signal propagation coefficient, which determines the zero cross point of the response function. The parameter h represents the Hill coefficient which determines the sharpness of the switch-like response. For the hyperbolic Tangent switching equation, the predicted response due to one cytokine is given by

$$\begin{aligned} F_v^{\mathrm{Tanh}}(V|P) &= b_{0,V} + b_{1,V}(\mathrm{Tanh}(b_{2,V}V + b_{3,V})); \\ F_p^{\mathrm{Tanh}}(P|V) &= b_{0,P} + b_{1,P}(\mathrm{Tanh}(b_{2,P}V + b_{3,P})), \end{aligned}$$

where b_0 denotes the basal response, b_1 - a lump parameter representing the effective maximal strength of the response, b_2 - a lump parameter representing the sharpness of the response, and b_3 - a lump parameter representing effective binding/signal propagation coefficient controlling the zero crossing point of the response. To model the combined effect of the two cytokines, we consider variants of the switch like models

in which the combined effects of the two cytokines multiplicative and additive.

$$
\begin{aligned}
F^{\text{Hill}}_{(+)}(V,P) &= \left[\alpha_{0,V} + \alpha_{1,V}\left(\frac{V^{h_v}}{\alpha_{2,V}^{h_v} + V^{h_v}}\right)\right] + \left[\alpha_{0,P} + \alpha_{1,P}\left(\frac{1}{\alpha_{2,V}^{h_p} + P^{h_p}}\right)\right] \\
&= \alpha_0 + \alpha_1\left(\frac{V^{h_v}}{\alpha_2^{h_v} + V^{h_v}}\right) + \alpha_3\left(\frac{1}{\alpha_4^{h_p} + P^{h_p}}\right); \\
F^{\text{Hill}}_{(\times)}(V,P) &= \alpha_0 + \alpha_1\left(\frac{V^{h_v}}{\alpha_2^{h_v} + V^{h_v}}\right) \times \left(\frac{1}{\alpha_4^{h_p} + P^{h_p}}\right); \\
F^{\text{Tanh}}_{(+)}(V,P) &= \beta_0 + \beta_1\Big(\text{Tanh}\big(\beta_2 V + \beta_3\big)\Big) + \beta_4(\text{Tanh}(\beta_5 P + \beta_6)\Big); \\
F^{\text{Tanh}}_{(\times)}(V,P) &= \beta_0 + \beta_1\Big(\text{Tanh}\big(\beta_2 V + \beta_3\big)\Big) \times \Big(\text{Tanh}\big(\beta_4 P + \beta_5\big)\Big).
\end{aligned}
$$

For each these model variants, we employ MATLAB's `nlinfit` function which finds optimal model parameters using the Levenberg-Marquardt algorithm (LMA) to minimize the least square error between the switch model prediction and the observed sprout density. We assess the performance of difference model variants by 1. the normality of residual (Fig S6b and d) and 2. the forward and reverse Kullbeck-Liebler divergence ($D_{KL}(F^*||F_o)$ and $D_{KL}(F_o||F^*)$)between the observed sprout density and the model prediction (Fig S6e). These measures represent the differences between the predicted and the observed sprout density distributions.

2.4 Comparing the objective functions for the condition-based vs. the cluster-based phenotypic transition rate estimates

In this section, we examine the difference in the objective functions used to derive the *condition based* and the *cluster based rate estimates*. Starting with the likelihood expression derived in section 2, for *condition based estimates*, we optimize the likelihood function over the set of trajectories within one experimental treatment condition. Alternatively, for *cluster based estimates*, we derive the maximum likelihood values after clustering the trajectories.

Given a set of experimentally observed state trajectories collected under a set $\mathbb{C}$ of $\mathcal{N}_c$ cytokine conditions, let's assume that a set $\mathbb{K}$ of $\mathcal{N}_k$ clusters are detected, where $\mathbb{K}$ is the set of all clusters and $\mathbb{C}$ is the set of all conditions. Let $\rho_{ss'}^{c,k}$ denotes the total number ss' type transitions observed in the single cell state trajectories of cells under condition c and assigned to cluster k. (These subpopulations may be distinct in state transition dynamics as consistent with the diverse population model for the sake of model comparison.) Then, the log likelihood of observing just the trajectories within cluster k under condition c is given by

$$
\mathcal{L}^{(c,k)} = \log(\ell^{(c,k)}) = \sum_{\mathcal{U}^{(c,k)}} \log(P_{s_o}) + \eta_{ss'}^{(c,k)} \log\left(\frac{\lambda_{ss'}}{\mu_s}\right) - \mu_s \sum_{h_{ss'}=1}^{\eta_{ss'}^{(c,k)}} t_{h_{ss'}}.
$$

Let $\xi_{ss'}^{(c,k)}(\lambda_{ss'}, \mu_s)$ be the derivative of log-likelihood with respect to $\lambda_{ss'}$ evaluated on the set of trajectories within condition c and condition k. From section 2, the this derivative take the following form:

$$
\xi_{ss'}(\lambda_{ss'}, \mu_s) = \sum_{\text{trajectories}} \eta_{ss'}\left(\frac{1}{\lambda_{ss'}} - \frac{1}{\mu_s}\right) - \sum_{\text{trajectories}} \sum_{h_{ss'}=1}^{\eta_{ss'}} t_{h_{ss'}}.
$$

Under the *uniform population model*, we optimize the log likelihood on the set each condition separate such that the derivative of likelihood for the subsets of trajectories within each condition n_c satisfy the

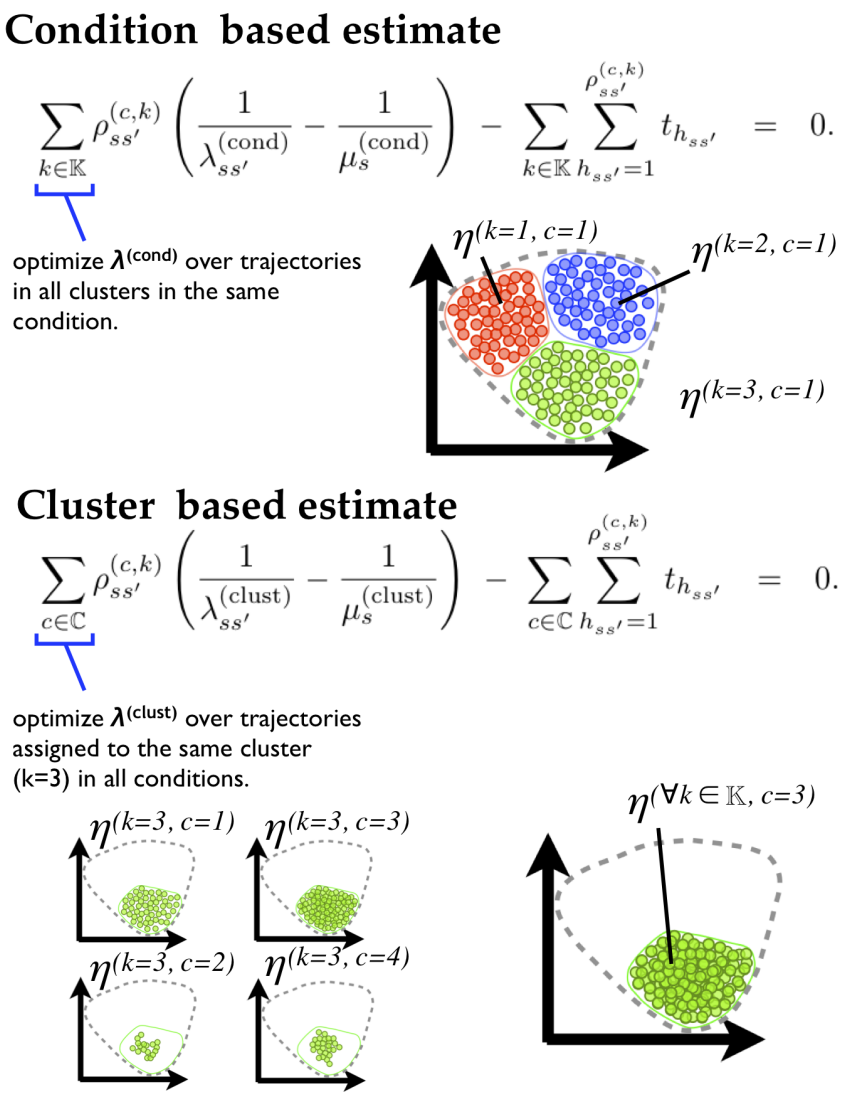


Fig SM1: Condition based and cluster based estimates are computed over different sets of single cell trajectories. Condition based estimates are optimized over single cell trajectories taken from the same cytokine conditions, while cluster based estimates are optimized over trajectories taken from the same cluster.

following optimal condition

$$\sum_{k\in\mathbb{K}} \xi_{ss'}^{(c,k)}(\lambda_{c_{ss'}}^{(n_c)}, \mu_{c_s}^{(n_c)}) = 0 \quad \text{i.e.,}$$

$$\sum_{k\in\mathbb{K}} \rho_{ss'}^{(c,k)} \left(\frac{1}{\lambda_{c_{ss'}}^{(n_c)}} - \frac{1}{\mu_{c_s}^{(n_c)}} \right) - \sum_{k\in\mathbb{K}} \sum_{h_{ss'}=1}^{\rho_{ss'}^{(c,k)}} t_{h_{ss'}} = 0.$$

Alternatively, under the *diverse population model*, the derivative of log-likelihood follows the relation

$$\sum_{c\in\mathbb{C}} \xi_{ss'}^{(c,k)}(\lambda_{k_{ss'}}^{(n_k)}, \mu_{k_s}^{(n_k)}) = 0 \quad \text{i.e.,}$$

$$\sum_{c\in\mathbb{C}} \rho_{ss'}^{(c,k)} \left(\frac{1}{\lambda_{k_{ss'}}^{(n_k)}} - \frac{1}{\mu_{k_s}^{(n_k)}} \right) - \sum_{c\in\mathbb{C}} \sum_{h_{ss'}=1}^{\rho_{ss'}^{(c,k)}} t_{h_{ss'}} = 0.$$

In attempting to relate the condition and cluster based estimates, we introduce $\lambda_{ss'}^{(n_c,n_k)}$ and $\mu_s^{(n_c,n_k)}$ which are the transition and total exit rate parameter sets optimized over the single cell trajectories in the n_k cluster within the n_c condition. As such, this set of parameter satisfy the following optimization condition:

$$\xi_{ss'}^{(c,k)}(\lambda_{ss'}^{(n_c,n_k)}, \mu_s^{(n_c,n_k)}) = 0 \quad \text{i.e.,}$$

$$\rho_{ss'}^{(k,c)} \left(\frac{1}{\lambda_{ss'}^{(n_c,n_k)}} - \frac{1}{\mu_s^{(n_c,n_k)}} \right) - \sum_{h_{ss'}=1}^{\rho_{ss'}^{n_k,n_c}} t_{h_{ss'}} = 0. \tag{1}$$

The subcluster estimates $\lambda_{ss'}^{(n_c,n_k)}$ can be related to the condition based $\lambda_{c_{ss'}}^{n_c}$ and the cluster based estimates $\lambda_{k_{ss'}}^{n_k}$ as follow:

$$\text{Eq (1)} = \sum_{k\in\mathbb{K}} \text{Eq (1)} \quad ; \quad \left(\frac{1}{\lambda_{c_{ss'}}} - \frac{1}{\mu_{c_s}} \right) \sum_{k\in\mathbb{K}} \rho^{(n_c,k)} = \sum_{k\in\mathbb{K}} \left(\frac{1}{\lambda_{ss'}^{(n_c,n_k)}} - \frac{1}{\mu_s^{(n_c,n_k)}} \right) \rho^{(n_c,n_k)} \tag{2}$$

$$\text{Eq (1)} = \sum_{c\in\mathbb{C}} \text{Eq (1)} \quad ; \quad \left(\frac{1}{\lambda_{k_{ss'}}} - \frac{1}{\mu_{k_s}} \right) \sum_{c\in\mathbb{C}} \rho^{(c,n_k)} = \sum_{c\in\mathbb{C}} \left(\frac{1}{\lambda_{ss'}^{(n_c,n_k)}} - \frac{1}{\mu_s^{(n_c,n_k)}} \right) \rho^{(n_c,n_k)}. \tag{3}$$

We can expand the sum, divide through by the total number of trajectories within a cluster ($\sum_{k\in\mathbb{K}} \rho^{(n_c,k)}$) for Eq (2) and the total number of trajectories within a condition $\sum_{c\in\mathbb{C}} \rho^{(c,n_k)}$) for Eq (3) to further simplify the above system of equations to obtain the following relationships:

$$\frac{1}{\lambda_{c_{ss'}}} - \frac{1}{\mu_{c_s}} = \sum_{k\in\mathbb{K}} \left(\frac{w_c}{\lambda_{ss'}^{(n_c,n_k)}} - \frac{w_c}{\mu_s^{(n_c,n_k)}} \right) \quad \text{and}$$

$$\frac{1}{\lambda_{k_{ss'}}} - \frac{1}{\mu_{k_s}} = \sum_{c\in\mathbb{C}} \left(\frac{w_k}{\lambda_{ss'}^{(n_c,n_k)}} - \frac{w_k}{\mu_s^{(n_c,n_k)}} \right) \quad \text{, where}$$

$$w_c = \frac{\rho^{(n_c,n_k)}}{\sum_{k\in\mathbb{K}} \rho^{(n_c,n_k)}} \quad \text{and} \quad w_c = \frac{\rho^{(n_c,n_k)}}{\sum_{c\in\mathbb{C}} \rho^{(n_c,n_k)}}$$

are the relative occurrence weights of ss' type jump across condition within a cluster and the relative occurrence weights over cluster within a condition respectively. Though these results do not directly relate the condition based estimates to the cluster based estimates, they reveal that the condition based and cluster based estimates importantly differ by the relative occurrence of the transition types within the set of single cell trajectories over which the parameters are optimized.

3 Supplementary Tables

3.1 Phenotypic cluster weights of endothelial cells under increasing VEGF and PF4

Cytokine Condition		Cluster Weights				
VEGF (ng/mL)	PF4 (ng/mL)	w_A	w_P	w_M	w_{Sw}	w_S
0	0	0.186	0.096	0.244	0.292	0.183
10	0	0.086	0.142	0.234	0.173	0.366
20	0	0.030	0.173	0.364	0.236	0.197
40	0	0.037	0.192	0.104	0.160	0.508
0	0	0.125	0.045	0.180	0.220	0.430
20	0	0.045	0.161	0.324	0.154	0.316
20	50	0.069	0.044	0.489	0.182	0.216
20	500	0.097	0.034	0.226	0.260	0.384

3.2 Estimated phenotypic transition rates of single cell trajectories grouped based on cytokine conditions ($\lambda^{(\mathbf{cond})}$)

Cytokine Condition		Transition Rates from S			Cytokine Condition		Transition Rates from P		
VEGF (ng/mL)	PF4 (ng/mL)	λ_{SP} ($\times 10^{-4}$)	λ_{SM} ($\times 10^{-4}$)	λ_{SA} ($\times 10^{-4}$)	VEGF (ng/mL)	PF4 (ng/mL)	λ_{PS} ($\times 10^{-4}$)	λ_{PM} ($\times 10^{-4}$)	λ_{PA} ($\times 10^{-4}$)
0	0	1.782	9.051	6.964	0	0	1.370	0.856	0.000
10	0	3.498	5.441	1.515	10	0	2.841	2.557	0.000
20	0	4.292	5.484	9.358	20	0	2.286	1.231	0.000
40	0	7.827	11.000	2.863	40	0	2.743	2.229	0.000
0	0	0.610	4.799	3.616	0	0	0.832	1.110	0.000
20	0	5.447	7.867	2.128	20	0	2.822	2.328	0.000
20	50	0.809	4.377	3.201	20	50	1.040	0.000	0.000
20	500	0.576	6.744	5.645	20	500	0.000	2.336	0.000
Cytokine Condition		Transition Rates from M			Cytokine Condition		Transition Rates from A		
VEGF (ng/mL)	PF4 (ng/mL)	λ_{MS} ($\times 10^{-4}$)	λ_{MP} ($\times 10^{-4}$)	λ_{MA} ($\times 10^{-4}$)	VEGF (ng/mL)	PF4 (ng/mL)	λ_{AS} ($\times 10^{-4}$)	λ_{AP} ($\times 10^{-4}$)	λ_{AM} ($\times 10^{-4}$)
0	0	1.235	3.107	0.620	0	0	0	0	0
10	0	1.306	0.982	0.000	10	0	0	0	0
20	0	1.056	0.530	0.265	20	0	0	0	0
40	0	0.806	0.270	0.270	40	0	0	0	0
0	0	1.167	0.000	0.780	0	0	0	0	0
20	0	0.792	0.397	0.000	20	0	0	0	0
20	50	1.722	0.433	0.865	20	50	0	0	0
20	500	0.326	0.000	0.000	20	500	0	0	0

3.3 Estimated phenotypic transition rates of single cell trajectories grouped based on phenotypic cluster ($\lambda^{(\mathbf{clust})}$)

Phenotypic cluster	Transition Rates from S			Phenotypic cluster	Transition Rates from P		
	λ_{SP} ($\times 10^{-4}$)	λ_{SM} ($\times 10^{-4}$)	λ_{SA} ($\times 10^{-4}$)		λ_{PS} ($\times 10^{-4}$)	λ_{PM} ($\times 10^{-4}$)	λ_{PA} ($\times 10^{-4}$)
cluster A	0.000	32.071	31.729	cluster A	0.000	0.000	0.000
cluster P	30.238	30.647	0.000	cluster P	15.829	15.131	0.000
cluster M	0.000	0.157	0.000	cluster M	0.000	0.000	0.000
cluster Sw	0.000	0.168	0.000	cluster Sw	0.000	0.000	0.000
cluster S	0.000	0.280	0.000	cluster S	0.000	0.000	0.000

Phenotypic cluster	Transition Rates from S			Phenotypic cluster	Transition Rates from S		
	λ_{MS} ($\times 10^{-4}$)	λ_{MP} ($\times 10^{-4}$)	λ_{MA} ($\times 10^{-4}$)		λ_{AS} ($\times 10^{-4}$)	λ_{AP} ($\times 10^{-4}$)	λ_{AM} ($\times 10^{-4}$)
cluster A	3.596	0.000	3.282	cluster A	0.000	0.000	0.000
cluster P	4.039	3.646	0.000	cluster P	0.000	0.000	0.000
cluster M	0.169	0.000	0.000	cluster M	0.000	0.000	0.000
cluster Sw	0.112	0.000	0.000	cluster Sw	0.000	0.000	0.000
cluster S	0.104	0.000	0.000	cluster S	0.000	0.000	0.000

4 Supplementary Data Sets

4.1 SI Data Set 1: Multichannel live cel imaging data showing cell behavior of GFP expression, RFP expression and unlabeled cells

The SI Data Set 1 contains example multichannel live cell images of a field of mixed population of GFP-labeled, RFP-labeled, and unlabeled hMVECs on collagen I gel. These data sets show that GFP-labeled, RFP-labeled, and unlabeled hMVECs are similar in their phenotypic transition patterns under the same cytokine conditions. This observation in turn suggests that the GFP and RFP reporter protein expression in hMVECs do not significantly affect hMVECs behavior and the interpretation of the imaging results.

4.2 SI Data Set 2: Data analysis scripts

The SI Data Set 2 contains all the data analysis scripts used in contour tracking, automated state classification, and parameter estimation based on CTMC. MATLAB builtin functions are commercially available through MathWorks and are not included here.

4.3 SI Data Set 3: Contour track data

The SI Data Set 3 contains all the tracked contour and centroid trajectories in the MATLAB data file format (.mat) from which the automated state classification is performed.

4.4 SI Data Set 4: Fluorescent live cell images for contour tracking

The dataset containing the original images of all the tracked hMVECs has a large zip file size (4.15 GB), so it is not submitted. However, the dataset is available upon request. The data in this dataset are taken from independent experiments (*exp1*, *exp2*, *exp3*) as specified by the subfolder titiles. The experimental setup from which these data sets are taken is described in the Fig1 and in the Method section.

5 References

References

[1] Vese L, Chan TF (2002) A multiphase level set framework for image segmentation using the Mumford and Shah model. Int'l J. Compo Vis., 50:271-293.

[2] Schapire R (1999) Improved boosting algorithms using confidence-rated predictions. Machine learning. 37(3):297-336.

[3] Rtsch G, Mika S, Schlkopf B (2002) Constructing boosting algorithms from SVMs: an application to one-class classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 24(9):1184 - 1199.

[4] Chung S, Sudo R, Zervantonakis IK, Rimchala T, Kamm RD (2009) Surface-Treatment-Induced Three-Dimensional Capillary Morphogenesis in a Microfluidic Platform. Adv Mater 21:48634867.

[5] Shin Y et al. (2011) In vitro 3D collective sprouting angiogenesis under orchestrated ANG-1 and VEGF gradients. Lab Chip 11:21752181.

[6] Das A, Lauffenburger D, Asada H, Kamm RD (2010) A hybrid continuum-discrete modelling approach to predict and control angiogenesis: analysis of combinatorial growth factor and matrix effects on vessel-sprouting morphology. Philos Transact A Math Phys Eng Sci 368:29372960.

[7] Stratman AN et al. (2009) Endothelial cell lumen and vascular guidance tunnel formation requires MT1-MMP-dependent proteolysis in 3-dimensional collagen matrices. Blood 114:237247.

[8] Wood L, Kamm R, Asada H (2011) Stochastic modeling and identification of emergent behaviors of an Endothelial Cell population in angiogenic pattern formation. The International Journal of Robotics Research 30:659677.

[9] Sudo R et al. (2009) Transport-mediated angiogenesis in 3D epithelial coculture. The FASEB Journal 23:21552164.

6 Supplementary Figures

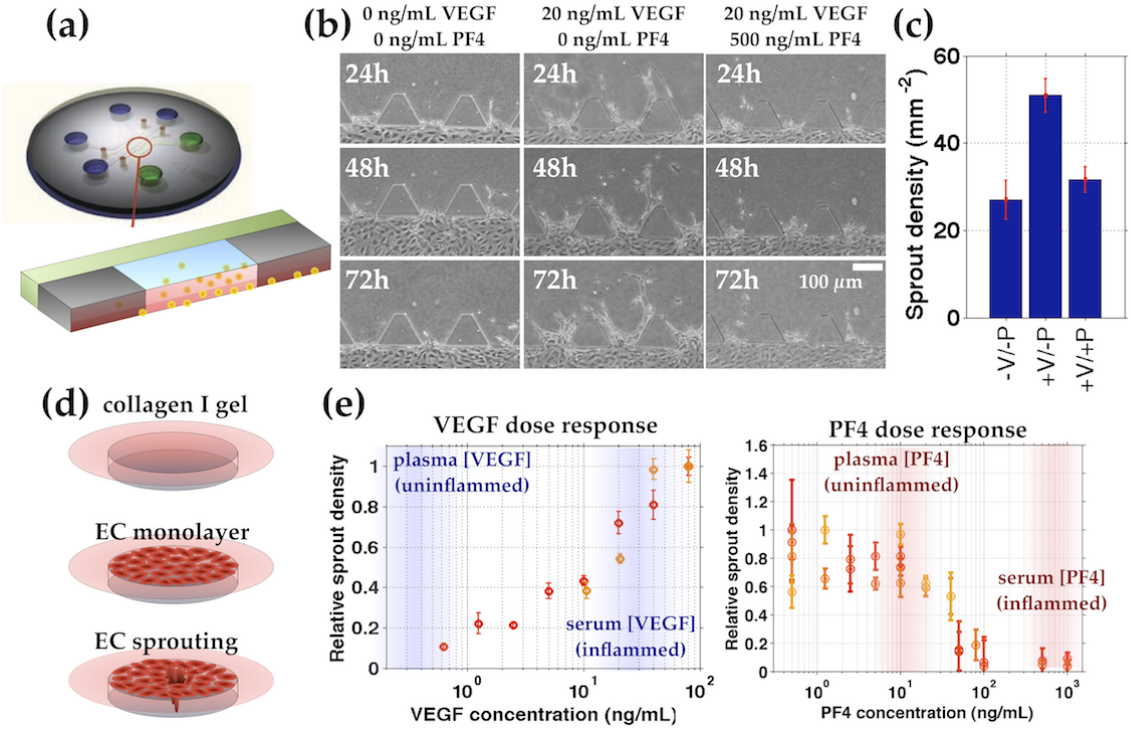


Fig S1: Inflammatory cytokines VEGF and PF4 modulate sprout densities of hMVECs in microfluidic device assay and collagen gel invasion assay. (a) The microfluidic device used in angiogenesis assay consists of two microfluidic channels separated by a middle region in which collagen I gel is casted. hMVECs seeded into one of the two channels form monolayer on the collagen I gel and subsequently send protrusions into the gel region. Device design, fabrication, and cell seeding protocol is previously reported ([4, 9]). (b) Representative images of hMVEC angiogenic sprouts in the two channel devices under different static cytokine conditions (no gradient) at 24 - 72 hrs after treatments. Physiological concentration of VEGF (20 ng/mL) induces extensive sprout formation. Addition of physiological concentration of PF4 (500 ng/mL) suppresses VEGF induced sprout formation. (c) Quantification of the result in (b) at 72 hr after cytokine stimulation. (d) Collagen gel invasion assay set up. Type I collagen gel is injected into a well of multi-well glass bottom plate to yield approximately 1 mm collagen gel slap. After gel polymerization, hMVECs are seeded at an instant monolayer density and are allowed to adhere on collagen gel. Adhered hMVECs send protrusions into collagen gel. (e) Quantified sprout densities of hMVEC in collagen gel invasion assay reveals that VEGF dose-dependently induces sprout formation in collagen invasion assay. Increasing PF4 concentrations in addition to a constant physiological concentration of VEGF (20 ng/mL), lead to dose-dependent decrease in hMVEC sprout density, indicating that PF4 dose-dependently suppresses sprout formation.

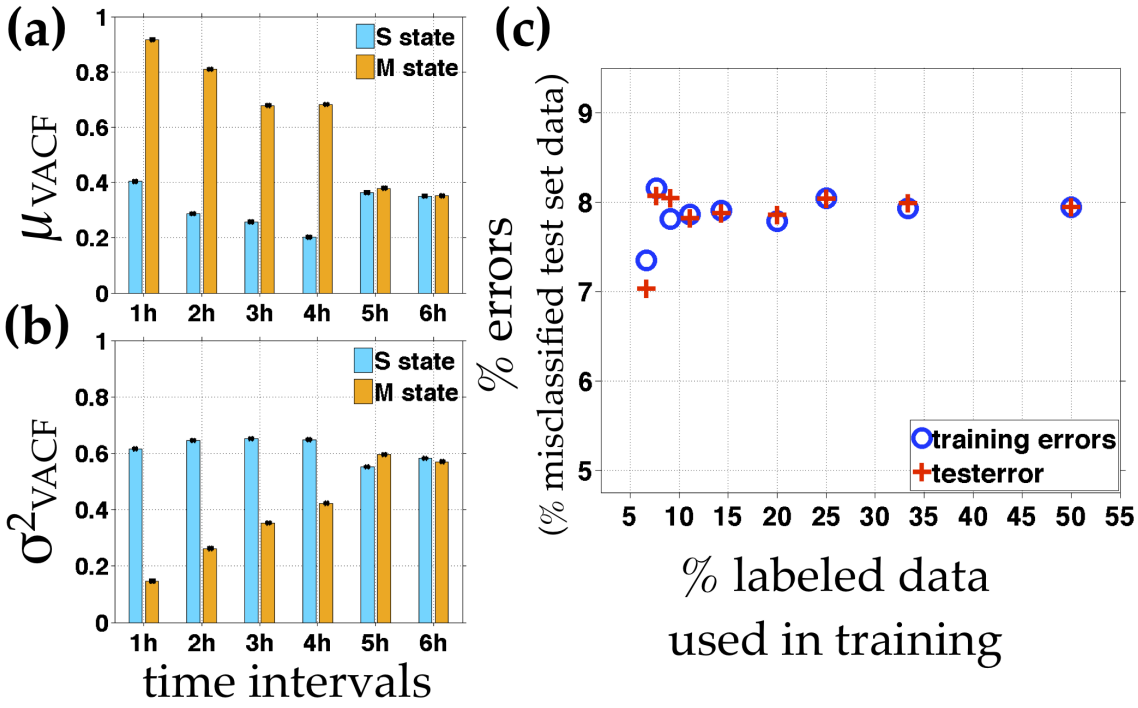


Fig S2: Instances assigned to sessile (S) and migratory (M) state exhibit distinct values of mean and variance of velocity autocorrelation functions (μ_{VACF} and σ^2_{VACF}) for intervals of length 1- 4 h. For the value of mean and variance in VACFs over 1h - 4h intervals, instances classified to M state have high mean VACF of 0.92, low variance of 0.15 for 1h interval (corresponding to 23.44 degree average angle deviation). Over longer time interval, mean VACF of M state instances continually decrease, reaching mean VACF of 0.68, variance VACF of 0.42 at 4 h interval (corresponding to 46.97 degree in angle deviation). On the other hand, the mean and variance VACF of instances in S states are low all across the time interval lengths over which VACFs are computed. For 1 h intervals, S state instances exhibit mean VACFs of 0.40, and variance VACF of 0.62 (corresponding to 66.20 degree angle deviation), while at 4 h intervals, S state instances exhibit mean VACF of 0.20, and variance VACF of 0.65 (corresponding to 78.36 degree angle deviation). (c) The S vs. M state classification errors of ensemble classifier determined by K-fold cross validation where K is the number of evenly divided *'chunks'* of data used for model training and testing. During the training, the algorithm uses a randomly chosen *'chunks'* ($1/K$ fraction) of the data to tune the ensemble model. The optimized model is then tested on the other $K-1/K$ fraction unseen data for their performance. In this study, 2- to 15-fold cross validations are performed on more than 80,000 labeled cell migration instances. For all the cross validation tests performed, approximately only 8% of the cell migration instances are misclassified. Note that the training and test errors are comparable for all cross validation tests with varying amount of labeled data used in training, suggesting that a small fraction of data cannot be correctly classified and that the trained model does not over-fit.

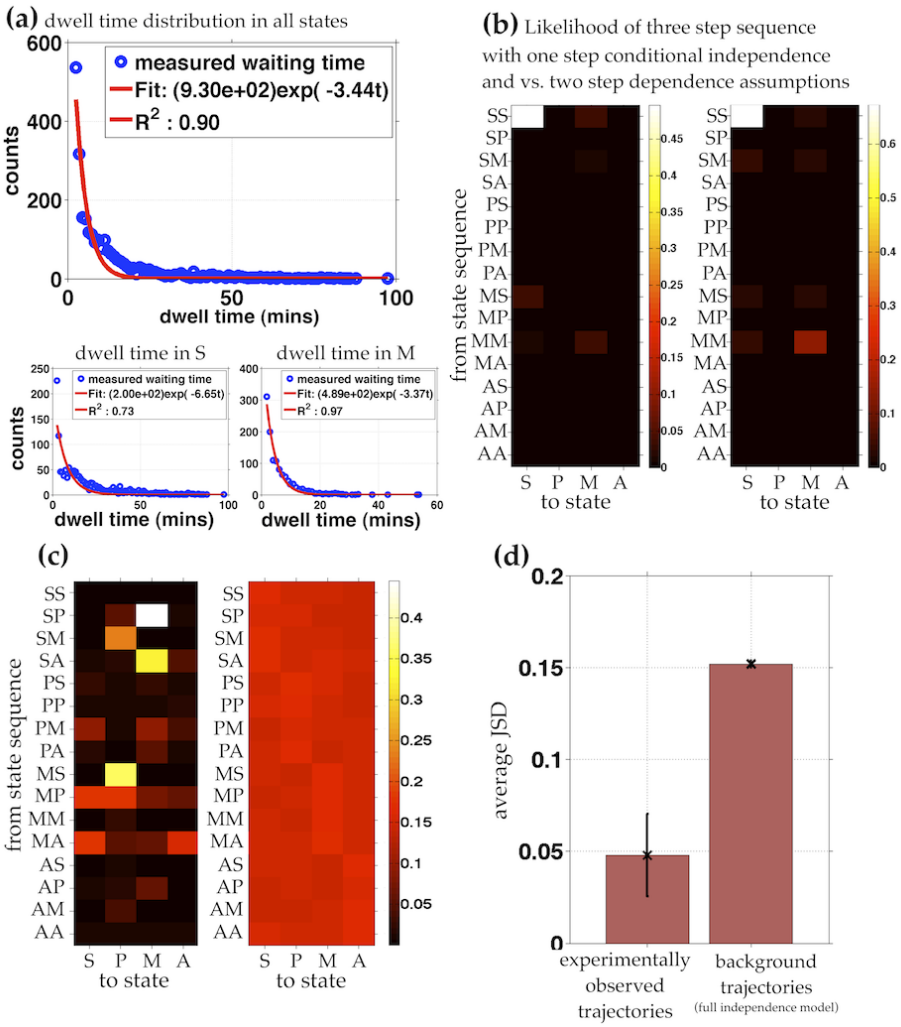


Fig S3: State trajectory can be well approximated by CTMC. (a) The dwell time distribution of in all states (S, P, M, A) can be well described by an exponential distribution with the coefficient of determination R^2 of 0.90. The distributions of dwell time in state S and M fit well to exponential distribution with R^2 of 0.73 and 0.97 respectively. Due to insufficient number of proliferative and apoptotic instances, the dwell time distributions in P and A states do not fit well to most well known statistical distributions. (b) Comparison of likelihood distributions under the one step condition independence assumption to those under the full dependence model of three step subsequence. The state trajectories distribution can be well approximated by both the model with conditional independence assumption (*left*) and the model with no conditional independence assumption (*right*). (c) The Jensen-Shannon divergence (JSD) measures the difference between likelihood distributions predicted by the full dependence and the conditional independence models. We compare JSD of the observed single cell state trajectories to that of a background trajectories generated from a full independence model. Given that sequence of states within these background trajectories are fully independent, likelihood distributions of these data based on one step dependent or full dependent models should be statistically the same. As such, the JSD of likelihood distributions of these background trajectories serves as a relevant control. JSD between the likelihood of full independence and conditional independence model of the data (experimentally observed state trajectories) is comparable to or smaller than that of the computationally generated background data (under the full independence model). (d). The averaged JSD of the experimentally observed trajectories is significantly smaller than that of the background trajectories computationally generated from full independence model.

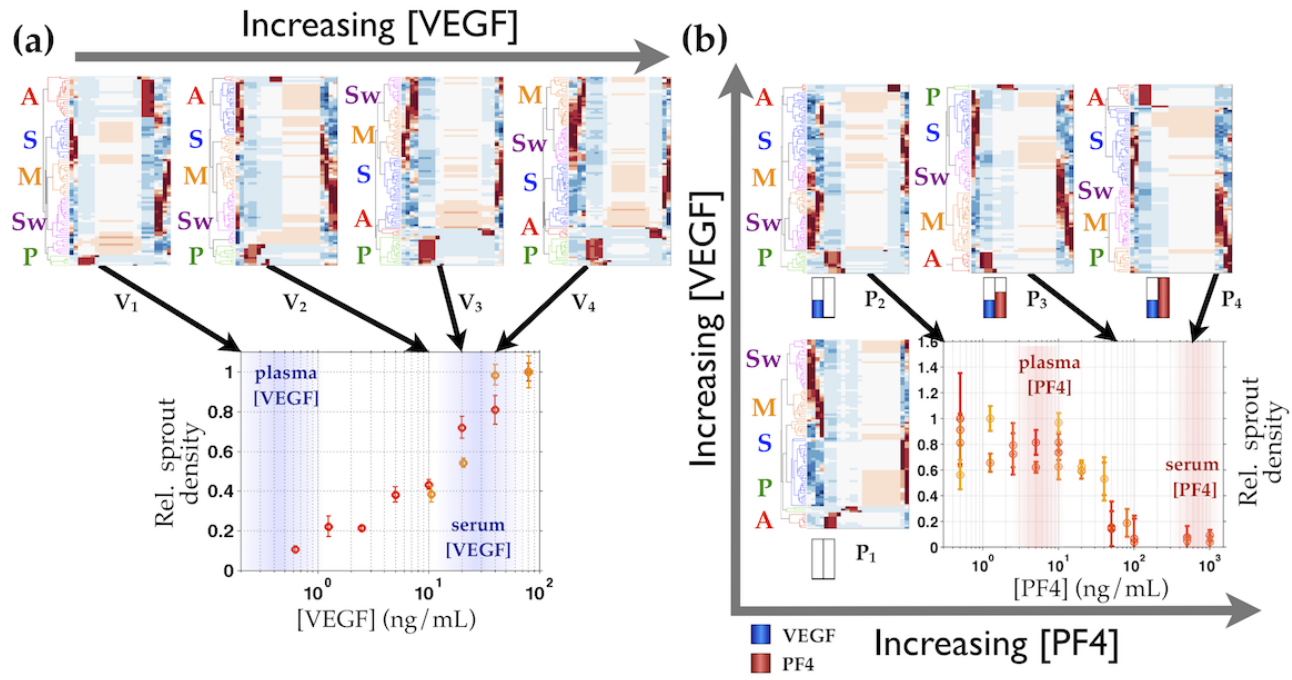


Fig S4: State trajectories of cells in different cytokine conditions cluster into 3-5 identifiable phenotypic subgroups. (a) Hierarchical clustering results of single cell state trajectories under incrementally increasing VEGF concentrations leads to identification of 5 highly phenotypic clusters. (b) Similar clustering results are observed in single cell state trajectories of endothelial cells treated with increasing PF4 concentrations in the background of physiological VEGF concentration (20 ng/mL).

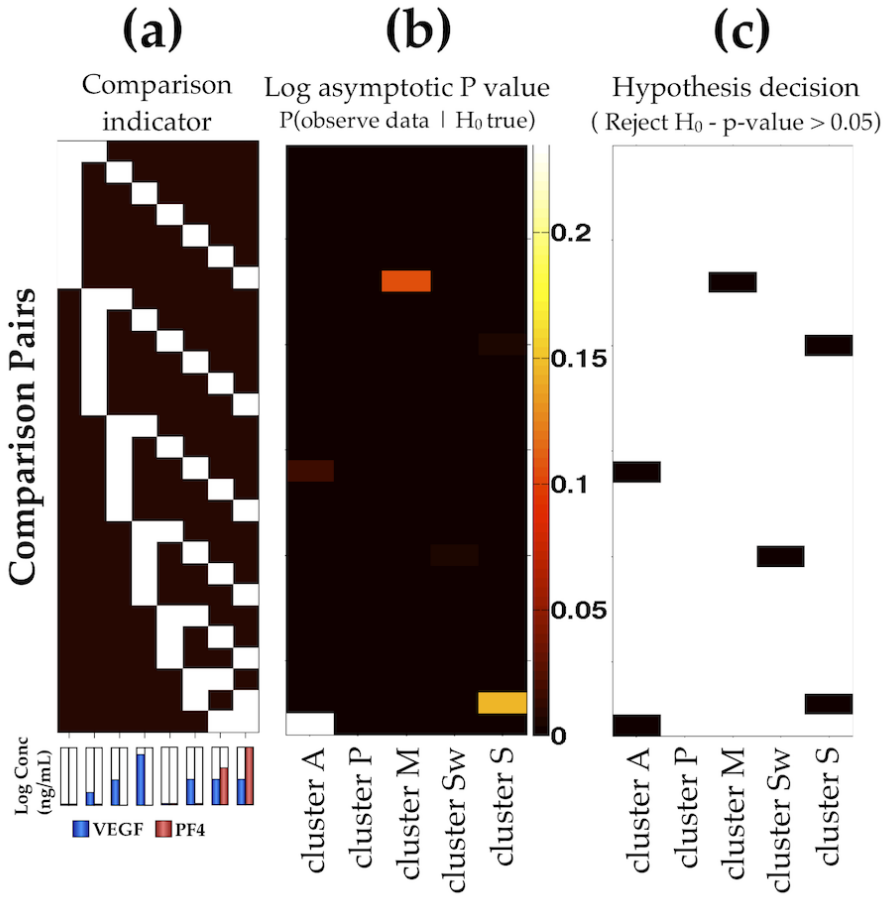


Fig S5: Statistical pairwise comparisons of the cluster weights across different cytokine conditions show that most of the cytokine elicited differences in the cluster weights are statistically significant. (a) Indicator matrix specifying the pairwise comparison. For example, the first row of (a) are true (white) for the first two conditions, indicating that the conditions being compared are conditions 1 and 2 (no cytokine vs. 20 ng/mL VEGF). (b) Log of the asymptotic p-value (probability of having the observed differences in cluster weights given that the null hypothesis is true). (c). The hypothesis decision based on the p-values indicating that most of the cytokine elicited changes in cluster weights are statistically significant.

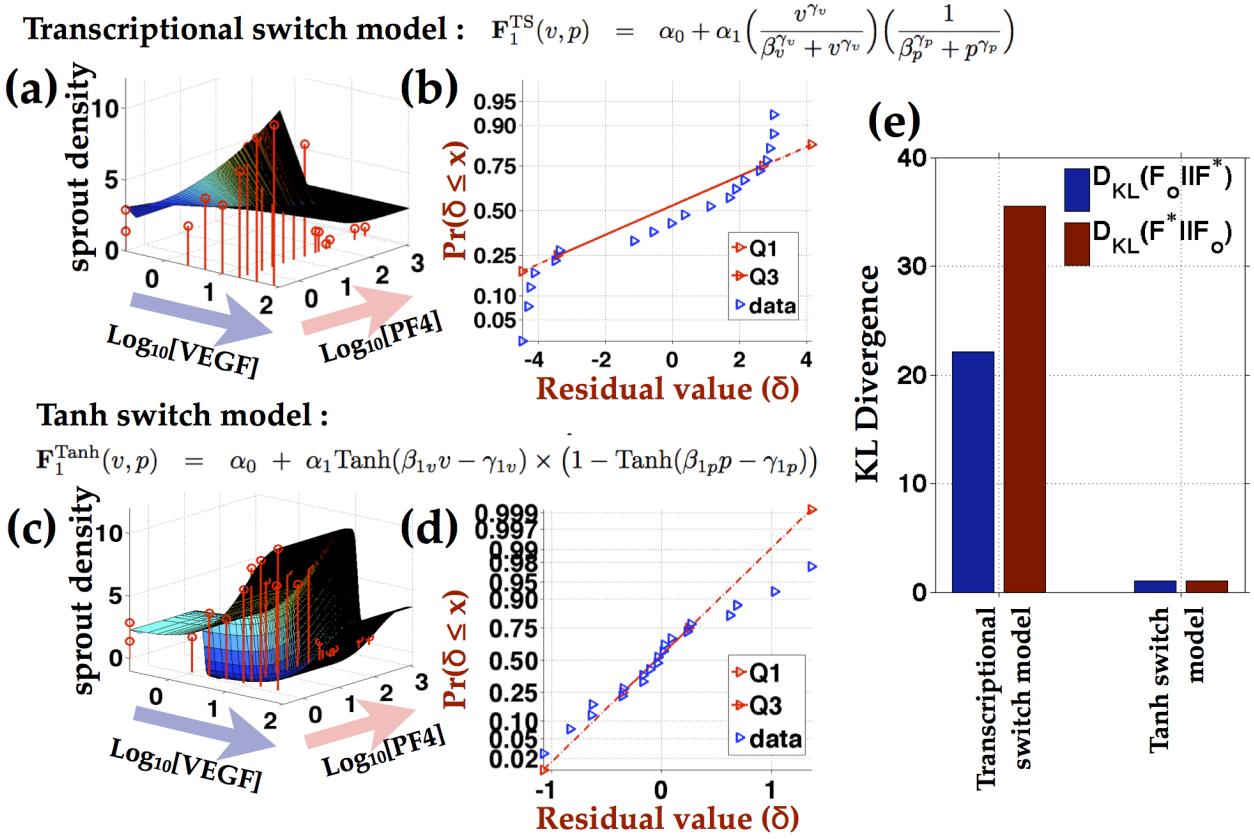


Fig S6: Assessment of switch model performance by normality of residuals and Kullback-Liebler divergence between the switch model estimated and the experimentally observed sprout. Comparison of these measures indicates that, given the same number of model parameters, the Tanh switch function more closely estimates the distribution of sprout density quantified from the confocal data. (a and c) Optimized sprout density distributions estimated by the transcriptional switch model (a) and Tanh switch model (c) are plotted against averaged sprout density quantified from confocal images to depict closer approximation of the Tanh switch model prediction to the data. (b and d) Normal probability plots of residuals in the transcriptional switch model (b) and Tanh switch model (d) is used to assess whether model residuals follow a normal distribution. While there are residual points that fall off the Q_1-Q_3 line (the line connecting the residual values at 1st and 3rd quantiles) in both normality plots, more are observed in the normality plots of the transcriptional switch model residual (b), especially at extreme values. (e) Forward and reverse Kullback-Liebler divergence ($D_{KL}(F^*||F_o)$ and $D_{KL}(F_o||F^*)$ respectively) between the estimated and the observed sprout density distributions over experimental range of VEGF and PF4 values. The Tanh switch model exhibits much smaller forward and reverse KL-divergence, suggesting that it more closely approximate the observed sprout density distribution.

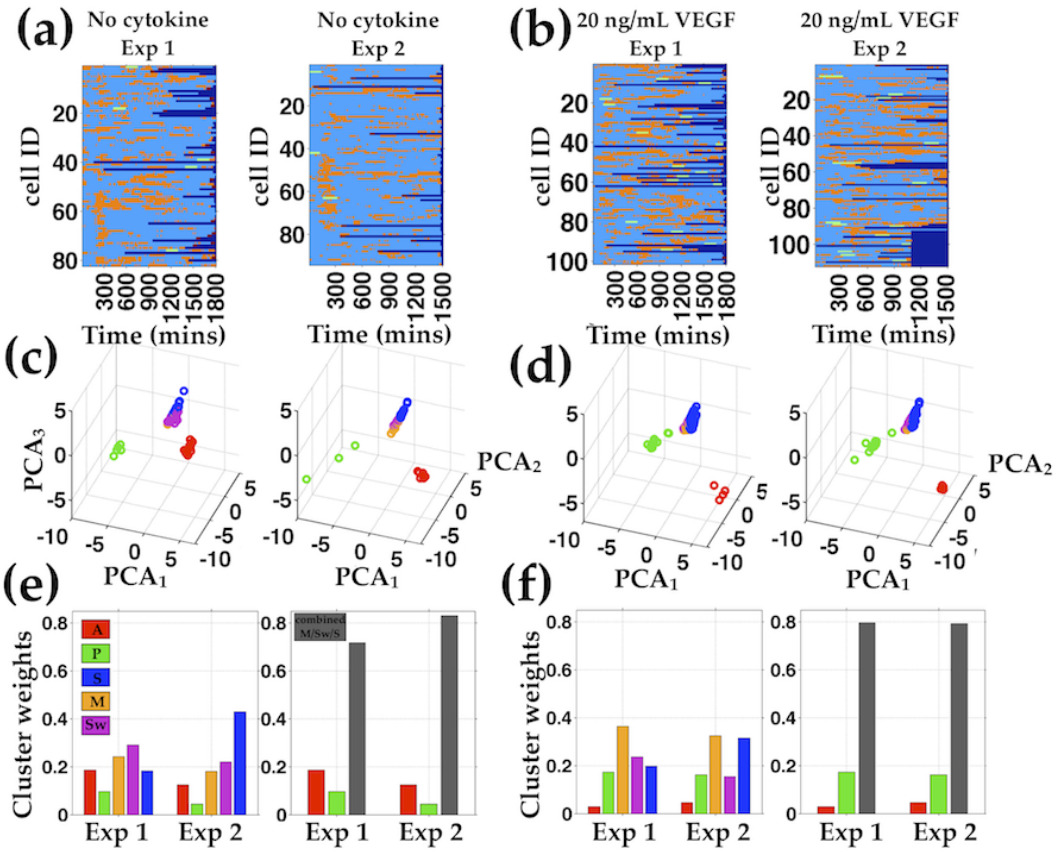


Fig S7: The single cell state trajectories of cells under the same cytokine conditions taken from independent experiments (from different days) exhibit similar phenotypic program based groupings, PCA embeddings, and phenotypic based cluster weights. Phenotypic states are colored labeled as in Fig 3. (a and b) Single cell state trajectories from two independent tracking experiments of hMVEC on type I collagen gel under no cytokine (a) and 20 ng/mL VEGF (b) conditions. (c and d) Independent experiments under the same cytokine conditions exhibit similar three component PCA embeddings. Phenotypic based subgroups are color labeled as in Fig 6-7. (e and f) Similar distributions of phenotypic based cluster weights are observed in both no cytokine (e, *left*) and 20 ng/mL VEGF (f, *left*) conditions. Since the M, Sw, S subgroups cluster closer together than A and P subgroups, the aggregate M,Sw,S cluster weights are also shown (e, *right* and f, *right* subplots).

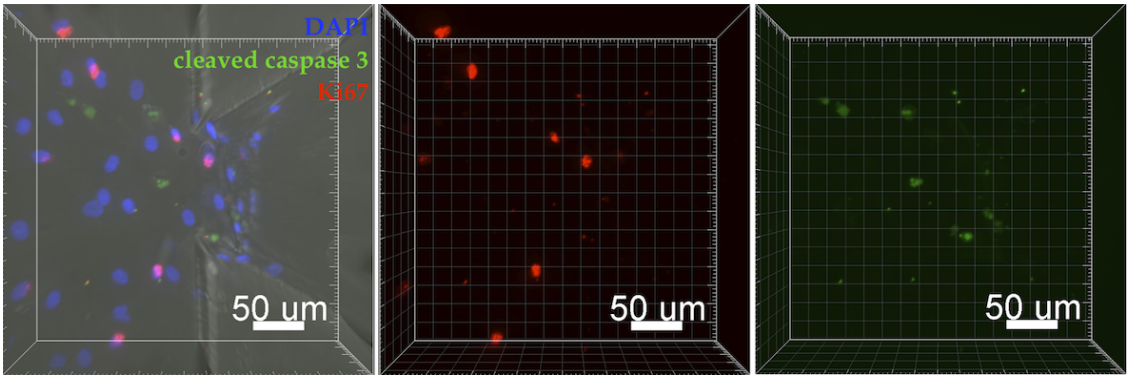


Fig S8: hMVECs adopt both apoptotic and proliferative states in microfluidic device assay. hVMECs were seeded into type I collagen gel containing microfluidic device as described in the Supplementary Methods 1.4 in SI Appendix and maintained in cell culture medium for 72 hrs. hMVECs were fixed in the microfluidic device, permeabilized and stained with anti-Ki67 (red) and anti-cleaved caspase3 (green) to visualize proliferative and apoptotic states of the cells. The fixed samples were also counter stained with DAPI to visualize nuclei. Ki67 positive cells and cleaved caspase 3 positive cells can be detected in the vicinity of the gel-channel interface in the microfluidic device.

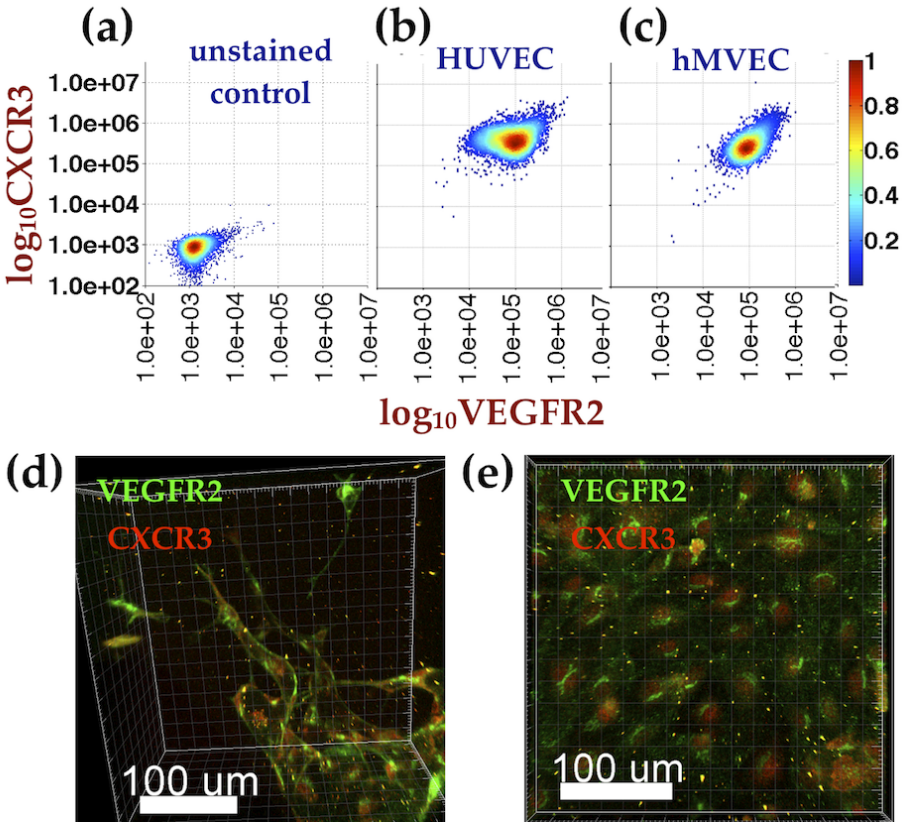


Fig S9: High levels of VEGFR2 and CXCR3 are detected in both human umbilical vein endothelial cells (HUVECs) and hMVECs as determined by flow cytometry (a - c) and immunofluorescent (IF) staining (d - e). (a) Unstained control, (b) doubly stained HUVECs and (c) doubly stained hMVECs. Immunofluorescent staining of cells in microfluidic devices reveals the two receptors are co-expressed in angiogenic sprouts (d) as well as in an endothelial monolayer (e).