cytoNet: Network Analysis of Cell Communities

Arun S. Mahadevan[1], Byron L. Long[1], Chenyue Wendy Hu[1], David T. Ryan[1], George L. Britton[3], Andrew Ligeralde[1], Aryeh Warmflash[1,3], Jacob T. Robinson[1, 2], Amina A. Qutub[1*]

[1]Department of Bioengineering, [2]Department of Electrical and Computer Engineering, [3]Department of Biosciences, Rice University, Houston, Texas, USA

[*]Contact: amina@rice.edu

We introduce cytoNet, a method to characterize multicellular topology from microscopy images. Accessible over the web, cytoNet quantifies the spatial relationships in cell communities using principles of graph theory, and evaluates the effect of cell-cell interactions on individual cell phenotypes. We demonstrate cytoNet's capabilities in two applications relevant to regenerative medicine: quantifying the morphological response of endothelial cells to neurotrophic factors present in the brain after injury, and characterizing cell cycle dynamics of differentiating neural progenitor cells. The framework introduced here can be used to study complex cell communities in a quantitative manner, leading to a deeper understanding of environmental effects on cellular behavior.

A cell's place in its environment influences a large part of its behavior. Advances in the field of phenotypic screening have yielded automated image analysis software that provide detailed phenotypic information at the single-cell level (such as morphology, stain texture and stain intensity) from microscopy images in a high-throughput manner[1,2]. However, current image analysis pipelines often do not account for spatial and density-dependent effects on cell phenotype. Various types of cell-cell interactions including juxtacrine and paracrine signaling are an integral part of biological processes that affect the behavior of individual cells. The recent emergence of technologies for multiparametric mapping of protein and RNA expression in individual cells while preserving the spatial structure of the tissue[3] has further highlighted the need to study single-cell behavior in the context of cell communities.

For these reasons, a robust method to quantify the spatial organization of cell communities and its influence on the behavior of individual cells adds an important, missing component to currently existing image analysis tools. Such a method can be used to enhance image-based biological discovery through phenotypic screens[2] by supplying multicellular metrics, provide a non-invasive means to standardize cell manufacturing for therapeutic purposes[4], and develop a quantitative framework for the analysis of spatially-detailed human cell atlas data[5,6].

Prior reports have accounted for population context in image-based screens by using features such as local cell density or a cell's position on an islet edge, that describe local cell crowding[7,8]. Mathematical graphs, structures that are used to model pairwise relationships between objects, are uniquely suited to cell community analysis. Among image-based methods that employ graph theory to analyze spatial relationships among cells, the cell-graph technique[9] has been employed to great effect in analyzing structure-function relationships in tissue sections. However, coupling single cell data to network structure has been elusive: there remains a need for a broadly applicable, user-friendly tool that enables spatial analysis of various different cell types, integrated with metrics describing the phenotype of individual cells.

Here we introduce an image analysis method called cytoNet for quantification of multicellular spatial organization using a graph theoretic approach. cytoNet is available as a web-based interface, providing significant ease of use compared with other programs that require downloading software. Taking fluorescence microscope images as input, the cytoNet image analysis pipeline identifies cells, creates spatial network representations tailored to the type of image and cell type, and calculates a set of metrics derived from graph theory that describe the network structure of the local multicellular neighborhood of a cell of interest. We define this multicellular neighborhood as a cell's community. Cell community metrics

46 are then integrated with descriptors of cell phenotype, such as morphology and protein expression, to

47 provide a comprehensive description of single- and multiple-cell phenotype states.

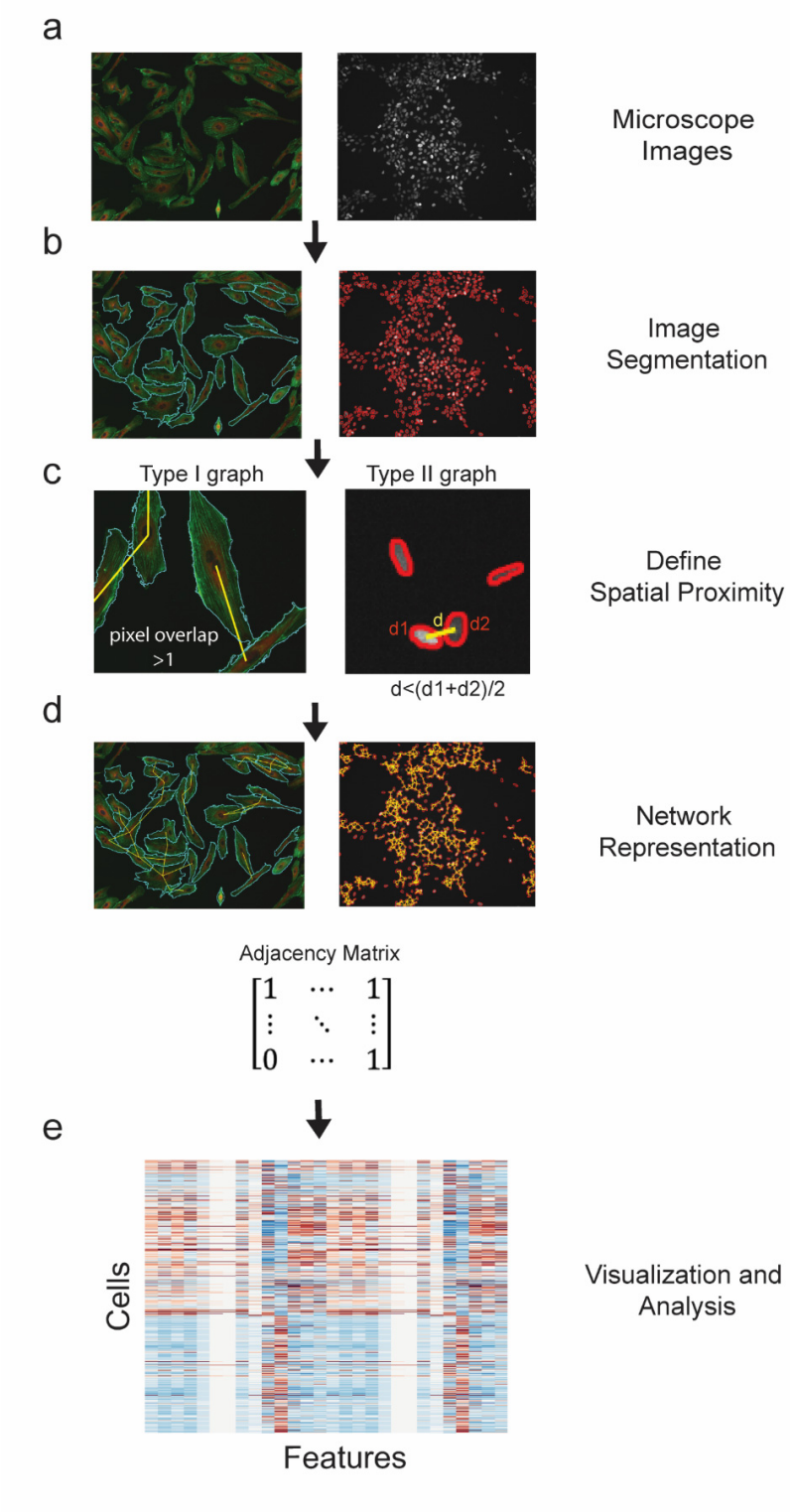


Figure 1. The cytoNet image-processing pipeline. (a) The pipeline begins with microscope images. (b) Segmentation algorithms automatically detect cell boundaries (**Supplementary Fig. 1, 2**). (c) Spatial proximity is determined either by measuring shared pixels between cell pairs – type I graphs (left panel) or by comparing distance between cell centroids to a threshold distance – type II graphs (right panel). (d) We represent the resulting network as an adjacency matrix. (e) Metrics derived from the adjacency matrix are used to describe network information. These metrics are a list of features computed on a per-cell basis.

The cytoNet pipeline begins with microscope images (**Fig. 1a**). Appropriate segmentation algorithms are implemented to detect cells (**Fig. 1b, Supplementary Fig. 1-2**). Upon detection of cells, the next step is to evaluate spatial proximity of cells. We do this in one of two ways – by evaluating the overlap of adjacent cell boundaries (type I graphs), or by evaluating the proximity of cells in relation to a threshold distance (type II graphs) (**Fig. 1c**). The former approach is useful when detailed information of cell boundaries and morphology is available, such as in the case of membrane stains or cells stained for certain cytoskeletal proteins. The latter approach is useful when dealing with images of cell nuclei, where detection of exact cell boundaries is not possible. In both approaches, cells deemed adjacent to each other are connected through edges, resulting in a network representation (**Fig. 1d**). This connectivity is denoted mathematically using an adjacency matrix, A (**Fig. 1d**), where $A_{i,j} = 1$ if there exists an edge between cells i and j, and 0 otherwise. Finally, the extracted metrics are used to visualize and analyze local neighborhood effects on individual cell phenotypes (**Fig. 1e**).

First, we demonstrate the utility of cytoNet in analyzing cell cycle dynamics in communities of differentiating neural progenitor cells. Neural progenitor cells are multipotent cells that can differentiate into neurons, astrocytes or oligodendrocytes. Cell cycle regulation in neural progenitor cells is of interest as it has implications for the genetic basis of brain size in different species[10] and aberrant regulation can cause diseases like microcephaly[11]. Studies in the ventricular zone of the embryonic mouse neocortex have shown that clusters of clonally-related neural progenitor cells go through the cell cycle together[12,13]. However, it is unclear whether this community effect is a ubiquitous feature of neural progenitor cells. To this end, we employed the cytoNet workflow to determine whether cell cycle synchronization is a feature of differentiating neural progenitor cells cultured in vitro.

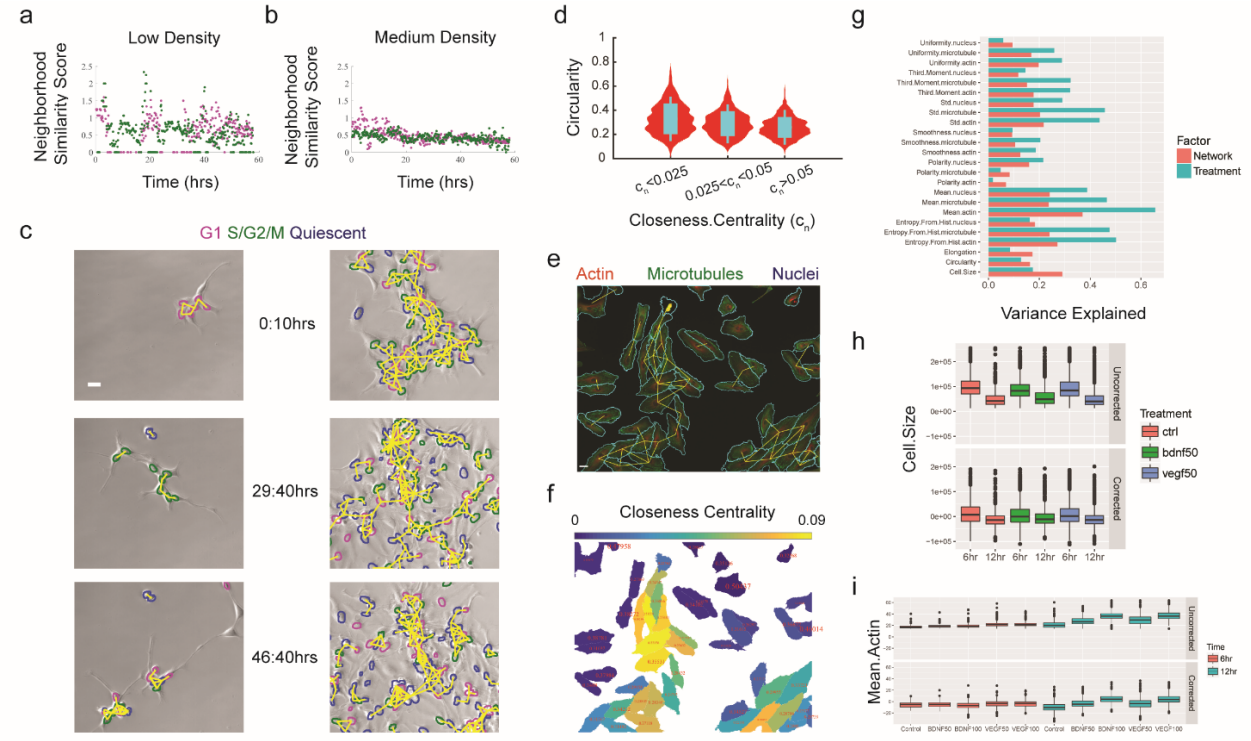


Figure 2. cytoNet reveals dynamic cell community trends and influence of cell density on individual cell morphology. (a-c) Spatiotemporal synchronization of cell cycle in differentiating neural progenitor cells **(a)** Neighborhood similarity score **(Supplementary Table 1)** for low-density culture across time. **(b)** Neighborhood similarity score across time for medium-density culture. **(c)** Frames from time-lapse movies corresponding to **(a)** and **(b)**. Borders of mCherry+ nuclei (G1) are outlined in magenta, Venus+ nuclei (S/G2/M) are outlined in green, and mCherry-/Venus- nuclei (quiescent) are outlined in blue; scale bar = 50µm. **(d-i)** Influence of local neighborhood density on primary human endothelial cell (HUVEC) morphology. **(d)** Distribution of cell circularity values grouped under different levels of closeness centrality; Cohen's d effect size: groups (1, 2) = 0.34, groups (1, 3) = 0.62; sample size, n=786 cells (group 1; $c_n < 0.025$), 741 cells (group 2; $0.025 < c_n < 0.05$) and 782 cells (group 3; $c_n > 0.05$) **(e)** Sample immunofluorescence image with graph representation overlaid; scale bar = 50 µm. **(f)** Heatmap depicting closeness centrality of each cell, with circularity values overlaid in text. **(g)** Bar plot of variance explained by growth factor treatment and local network metrics. **(h)** Box plot of cell size as a function of growth factor treatment. **(i)** Box plot of mean actin intensity as a function of growth factor treatment. Legends and axes in (h-i) contain information on treatment (BDNF, VEGF), concentration (50ng/ml, 100ng/ml) and time of treatment (6 hours and 12 hours). Cohen's d effect size for (h-i) is shown in **Supplementary Table 3**.

For this investigation, ReNcell VM human neural progenitor cells were stably transfected with the FUCCI cell cycle reporters[14] to generate Geminin-Venus/Cdt1-mCherry/H2B-Cerulean (FUCCI-ReN) cells. We captured time-lapse movies of FUCCI-ReN cells after withdrawing growth factors to induce differentiation, and built network representations from nucleus images. Adjacency was determined by comparing centroid-centroid distance to a threshold (type II graphs, **Fig. 1c**).

In order to evaluate spatiotemporal synchronization in cell cycle, for each individual cell in a frame, we evaluated the average fraction of neighboring cells in a similar phase of the cell cycle (G1 phase – mCherry+ and S/G2/M phases – Venus+), normalized by total fraction of that cell type in the population. We called the average value of this fraction across all cells in an image the neighborhood similarity score, N_S (**Supplementary Table 1**). Results for medium and low-density cultures are shown in **Fig. 2a** and **Fig. 2b** respectively. Frames from corresponding time-lapse movies are shown in **Fig. 2c** (see also **Supplementary Videos 1-4**). We observed that groups of cells in the low-density culture moved through the cell cycle in unison, which was reflected in periodically high values of the neighborhood similarity score (**Fig. 2a**, **Supplementary Video 1-2**). In contrast, the composition of cell clusters in the medium density culture was relatively heterogeneous, resulting in relatively low values of the neighborhood similarity score over time (**Fig. 2b**, **Supplementary Video 3-4**). Neighboring cells in very low-density cultures are likely to be derived from the same clonal lineage, which explains the high level of synchronization in these cultures[12]. This example highlights how cytoNet can be used to derive insight into the role of cell-cell interactions on dynamic cell behavior.

Next, we used cytoNet to evaluate the relative influence of local neighborhood density and growth factor perturbations on endothelial cell morphology. From a regenerative medicine perspective, studying the morphological response of endothelial cells to neurotrophic stimuli can help assess the cells' potential angiogenic response following brain injuries that induce growth factor secretion, like ischemic stroke or transient hypoxia[15,16]. Common high-throughput angiogenic assays focus on migration and proliferation as the main cell processes defining angiogenesis, or the growth of new capillaries from existing ones[17]. Distinct morphology and cytoskeletal organization of endothelial cells indicate the cell's migratory or proliferative nature, and hence their angiogenic contribution within a sprouting capillary[18]. Reproducibly

quantifying the morphological response of endothelial cells to neurotrophic factors would enable more targeted approaches to enhancing brain angiogenesis.

We took an image-based approach to this problem, building a library of immunofluorescence images of human umbilical vein endothelial cells (HUVECs) stained for cytoskeletal structural proteins (actin, α-tubulin) and nuclei, in response to various combinations of vascular endothelial growth factor (VEGF) and brain-derived neurotrophic factor (BDNF) treatment. Cell morphology was annotated using 21 metrics described in our previous study[19] (**Supplementary Table 2**), which included cell shape metrics like circularity and elongation, and texture metrics for cytoskeletal stains such as Actin polarity, smoothness etc. Cluster analysis on this dataset revealed dominant morphological phenotypes as a function of treatment conditions (**Supplementary Fig. 3**).

We then used the cytoNet workflow to quantify density-dependent effects on endothelial cell morphology in control cultures (without any growth factor perturbation). Network representations were designated based on shared cell pixels (type I graphs, **Fig. 1c**) and local network properties were described using seven metrics, including degree (number of neighbors) and centrality measures (indicating relative location of cells within colonies) (**Supplementary Table 1**). Our analysis showed correlations between cell morphological features and local network properties (**Supplementary Fig. 4**). Some of these relationships were expected, for instance the positive correlation between shared cell border and cell size. Other relationships, such as the negative correlation between cell circularity and closeness centrality, capture intuitive notions of the influence of cell packing on morphology (**Fig. 2d-f**). The closeness centrality of a cell (**Supplementary Table 1**) describes its relative position in a colony – cells in the middle of a colony will have higher centrality values than cells at the edge of a colony or isolated cells. The negative relationship between circularity and closeness centrality implies that isolated cells and cells located at the edge of

colonies are more likely to have a circular morphology, while more densely packed cells tend to be less circular (**Fig. 2e-f**). Thus, our analysis revealed that local network properties have a quantifiable effect on cell morphology.

Next, we developed a workflow to analyze the effect of growth factor treatments on cell morphology, while correcting for the effect of local network properties. We applied a quantile multidimensional binning approach[20,21] to calculate the variance in morphology metrics that could be individually explained by all local network metrics and growth factor treatments (**Fig. 2g**). We then calculated the values for each morphology metric after correcting for the effect of local network metrics (see **Methods**). The raw and network-corrected values for two metrics, cell size and mean actin intensity, are shown as box plots in **Fig. 2h-i.** The influence of network properties can be clearly seen on cell size, where at 6 hours, large cell sizes are seen in the uncorrected but not corrected plots (**Fig. 2h**). The effect of growth factor treatment can be clearly seen in network-corrected mean actin intensity (**Fig. 2i, Supplementary Table 3**), where VEGF and BDNF treatment have dose-dependent effects on mean actin intensity. Thus, cytoNet detects the independent effects of local neighborhood properties and growth factor perturbations on endothelial cell morphology.

The examples described above illustrate how cytoNet can be used to enhance image informatics for phenotypic screens as well as basic discovery in biology. From the image informatics perspective, cytoNet adds crucial information on local cell density to the suite of metrics that are currently used to characterize individual cells. We illustrated how local network metrics can be used to infer independent effects of cell density and chemical perturbations. This workflow can be used to more comprehensively characterize the response of cells to chemical perturbations, which can aid in drug discovery.

The cytoNet workflow can also be used to quantitatively study biological pathways involved in cell-cell communication. The combination of visualizing dynamic cell behavior through time-lapse movies and quantifying local cell-cell interactions is particularly powerful. This paradigm can be of great benefit in stem cell biology to evaluate environmental effects on cell fate decisions. More broadly, the principle behind cytoNet – treating cell communities as complex ecosystems – will help transition from characterizing cells as independent 'silos' to a more holistic approach, where due importance is given to the environment surrounding cells.

ACKNOWLEDGEMENTS

We thank Dr. David Noren, Dr. André Schultz and Tien Tang for helpful discussions and comments on the manuscript, and Amada Abrego, Becky Zaunbrecher and Grace Ching for technical assistance. This work was supported by NSF Career Grant 1150645 to AAQ, NSF Neural and Cognitive Systems grant 1533708 to AAQ and JTR, NSF award 1553228 to AW and CPRIT award RR140073 to AW. ASM was supported through NSF IGERT training grant 1250104.

AUTHOR CONTRIBUTIONS

ASM, DTR, JTR and AAQ designed the experiments. ASM, GLB and DTR performed the experiments. AAQ, BLL, CWH and ASM analyzed the data. BLL and AL designed and implemented the cytoNet website. All authors contributed to writing the manuscript. AAQ, AW and JTR supervised the work.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

158 **REFERENCES**

159 1. Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A. & Olivo-Marin, J.-C. Imagining the
160 future of bioimage analysis. *Nat. Biotechnol.* **34,** 1250–1255 (2016).

161 2. Bray, M.-A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling
162 using multiplexed fluorescent dyes. *Nat. Protoc.* **11,** 1757–1774 (2016).

163 3. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the
164 mammalian liver. *Nature* **542,** 352–356 (2017).

165 4. Anderson, A. J., Piltti, K. M., Hooshmand, M. J., Nishi, R. A. & Cummings, B. J. Preclinical Efficacy
166 Failure of Human CNS-Derived Stem Cells for Use in the Pathway Study of Cervical Spinal Cord
167 Injury. *Stem Cell Reports* **8,** 249–263 (2017).

168 5. Lavin, Y. *et al.* Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell
169 Analyses. *Cell* **169,** 750–765.e17 (2017).

170 6. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science (80-.).* **356,** (2017).

171 7. Snijder, B. *et al.* Population context determines cell-to-cell variability in endocytosis and virus
172 infection. *Nature* **461,** 520–523 (2009).

173 8. Schapiro, D. *et al.* histoCAT: analysis of cell phenotypes and interactions in multiplex image
174 cytometry data. *Nat. Methods* (2017). doi:10.1038/nmeth.4391

9. Yener, B. & Bülent. Cell-graphs. *Commun. ACM* **60,** 74–84 (2016).

10. Otani, T. *et al.* 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in Progenitor Behavior Contributing to Brain Size. *Cell Stem Cell* **18,** 467–480 (2016).

11. Li, C. *et al. Zika Virus Disrupts Neural Progenitor Development and Leads to Microcephaly in Mice*. *Cell Stem Cell* **19,** (2016).

12. Cai, L., Hayes, N. L. & Nowakowski, R. S. Synchrony of clonal cell proliferation and contiguity of clonally related cells: production of mosaicism in the ventricular zone of developing mouse neocortex. *J. Neurosci.* **17,** 2088–100 (1997).

13. Reznikov, K. & van der Kooy, D. Variability and partial synchrony of the cell cycle in the germinal zone of the early embryonic cerebral cortex. *J. Comp. Neurol.* **360,** 536–54 (1995).

14. Sakaue-Sawano, A. *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* **132,** 487–498 (2008).

15. López-Cancio, E. *et al.* Reported Prestroke Physical Activity Is Associated with Vascular Endothelial Growth Factor Expression and Good Outcomes after Stroke. *J. Stroke Cerebrovasc. Dis.* **26,** 425–430 (2017).

16. Wei, Z. Z. *et al.* Neuroprotective and regenerative roles of intranasal Wnt-3a Administration after

focal ischemic stroke in mice. *J. Cereb. Blood Flow Metab.* 0271678X1770266 (2017). doi:10.1177/0271678X17702669

17. Goodwin, A. M. In vitro assays of angiogenesis for assessment of angiogenic and anti-angiogenic agents. *Microvasc. Res.* **74,** 172–183 (2007).

18. Costa, G. *et al.* Asymmetric division coordinates collective cell migration in angiogenesis. *Nat. Cell Biol.* **18,** 1292–1301 (2016).

19. Slater, J. H. *et al.* Recapitulation and Modulation of the Cellular Architecture of a User-Chosen Cell of Interest Using Cell-Derived, Biomimetic Patterning. *ACS Nano* **9,** 6128–6138 (2015).

20. Snijder, B. *et al.* Single-cell analysis of population context advances RNAi screening at multiple levels. *Mol. Syst. Biol.* **8,** (2012).

21. Gut, G., Tadmor, M. D., Pe'er, D., Pelkmans, L. & Liberali, P. Trajectories of cell-cycle progression from fixed cell populations. *Nat. Methods* **12,** 951–4 (2015).

METHODS

Software. CytoNet is available as a web interface at https://qutublab.org/cytoNet/

See **Supplementary Methods 1** for instructions on using cytoNet.

Cell Culture. Human umbilical vein endothelial cells (HUVECs, Lonza) were cultured in EBM-2 medium (Lonza) supplemented with penicillin-streptomycin (Fisher Scientific) and EGM-2 SingleQuot bullet kit (Lonza). For imaging experiments, cells were cultured for different periods (6, 12 or 24 hours) in different combinations of vascular endothelial growth factor (VEGF, human recombinant; Millipore) and brain-derived neurotrophic factor (BDNF, human recombinant, Sigma-Aldrich). Concentrations used were in the range 50-100ng/mL. Controls were the same culture period without growth factor treatments.

Immortalized human neural progenitor cells derived from the ventral midbrain (ReNCell VM) were obtained from Millipore. Cells were expanded on laminin-coated tissue culture flasks, in media containing DMEM/F12 supplemented with B27 (both Life Technologies), 2µg/ml Heparin (STEMCELL Technologies), 20ng/ml bFGF (Millipore), 20ng/ml EGF (Sigma) and penicillin/streptomycin. For differentiation experiments, cells were cultured in medium lacking bFGF and EGF.

FUCCI Reporter Lines. Stable reporter cell lines (FUCCI-ReN) were generated by sequentially nucleofecting ReNcell VM neural progenitor cells with an ePiggyBac[22] construct encoding mCherry-Cdt, Venus-Geminin, or Cerulean-H2B. Each construct introduced to the cells was driven by a CAG promoter containing a blasticidin (ePB-B-CAG-mCherry-Cdt1), puromycin (ePB-P-Venus-Geminin), or neomycin (ePB-N-Cerulean-H2B) resistance gene. Following each round of nucleofection, cells were cultured in the presence of appropriate antibiotics (2 µg/ml blasticidin, 0.1 µg/ml puromycin and 100 µg/ml neomycin).

HUVEC Immunocytochemistry. For imaging experiments, HUVECs were cultured on glass dishes coated with fibronectin (Sigma-Aldrich). After appropriate growth factor treatments, cultures were fixed with 4% paraformaldehyde, free aldehyde groups were quenched using 1mg/mL sodium borohydride, and

membranes were permeabilized with 0.2% Triton-X-100 solution in PBS. Actin fibers were visualized using an Alexa Fluor 488-phalloidin antibody (1:40, Molecular Probes) and microtubules were visualized using a mouse monoclonal anti-α-Tubulin antibody (1:250, Sigma-Aldrich) followed by a goat anti-mouse Alexa Fluor 647 secondary antibody. Nuclei were stained using Hoescht (Molecular Probes). 16-bit composite immunofluorescence images were acquired through a 20X objective (N.A. = 0.75) on a Nikon Ti-E epifluorescence microscope. Physical pixel size was 0.32µm.

Time-lapse Microscopy. FUCCI-ReN cells were plated at different densities on chambered cover glasses (Fisher Scientific) coated with laminin. Cells were imaged after switching to differentiation medium containing phenol red-free DMEM/F12. Time-lapse imaging was performed using a Nikon Ti-E microscope equipped with a motorized stage, a cage incubator for environmental control (Okolab), a 20X objective lens (N.A. = 0.75), SOLA SE Light Engine for LED-based fluorescence excitation (Lumencor), appropriate filters for visualizing mCherry, Venus and Cerulean fluorescent proteins and a Zyla 5.5 sCMOS camera (ANDOR). 16-bit composite fluorescence images were acquired at 10 minute intervals for a total duration of 57.5 hours.

Image Processing of HUVEC Immunofluorescence Images. Fluorescence images were processed as described previously[23] (**Supplementary Fig. 1**). Briefly, the following steps were used.

1. Contrast was enhanced using histogram equalization.
2. Images were smoothed using a 2D Gaussian lowpass filter.
3. Initial binarization was performed using Otsu's method.
4. The binary image was dilated to fill in individual cell areas.
5. All objects <1% of the total image area were removed. This was called the final binary image.
6. A binary representation of the nuclear and microtubule image layers was generated using a high input threshold value. This was called the marker image.

7. Another binary image was created with values of 0 where either the final binary image (step 5) or the marker image (step 6) had a value of 1.
8. Watershed markers were generated by imposing the minimum of the complement of images obtained in steps 2 and 7. This image had black markers contained within cells to serve as basins for flooding, while cell areas themselves were represented by lighter pixels that served as the rising contours of the basins.
9. The watershed algorithm was implemented using Matlab's built-in function to generate cell boundaries.
10. Masks generated in step 9 were refined by using composite images of microtubules and actin as the marker image (step 6).

In order to automate the threshold generation, the area of cell masks obtained from segmentation were compared to those obtained through thresholding with a high threshold. The entire process was then iterated until an acceptable area ratio was achieved.

Image Processing of FUCCI-ReN Time-Lapse Images. Grayscale images for each channel (H2B-Cerulean, Geminin-Venus and Cdt1-mCherry) were binarized using locally adaptive thresholding. Seeds for the watershed transform were generated using the regional minimas from the distance transform of the grayscale images. Next, the watershed algorithm was applied to detect boundaries between overlapping cell nuclei. Finally, information from different channels were used to correct undersegmented nuclei.

Generation of Network Representation. Type I graphs were generated as follows. Mask boundaries were expanded by 2 pixels and overlap of expanded masks was used to assign edges and build an adjacency matrix (**Fig. 1c**). Cells touching the image border were included in calculations of local network properties (**Supplementary Table 1**) for cells not touching the boundary, but were excluded for the construction of the adjacency matrix.

Type II graphs were generated as follows: For each pair of objects (nuclei), a threshold distance for proximity was defined as the average of the two object diameters, multiplied by a scaling factor (S). If the Euclidean distance between the object centroids was lower than the threshold distance computed, then the pair of objects was connected with an edge (**Fig. 1c**). We chose a default scaling factor S = 2 for all our analysis, through visual inspection of cell adjacency.

Network Metric Computation. All the network metrics described in **Supplementary Table 1** were computed using custom-written code, building upon the routines provided in Bounova et al[24].

Correction of morphology metrics for effects of local network properties and treatment conditions. We performed quantile multidimensional binning[20] of cells for all 7 network metrics (5 bins per metric). The mean of each morphology metric was calculated for each multidimensional bin, and this mean was subtracted from the raw measurements to generate the network-corrected measurements for each cell. Treatment-corrected measurements were generated similarly by calculating the mean of each morphology metric under each treatment condition and then subtracting it from the raw measurements.

Variance explained by local network properties and treatment conditions. The variance explained by each factor was calculated using the following formula[21]

$$1 - V_{corr}/V_{uncorr}$$

V_{corr} is the variance of the corrected measurements, and V_{uncorr} is the variance of the uncorrected measurements.

22. Lacoste, A., Berenshteyn, F. & Brivanlou, A. H. An Efficient and Reversible Transposable System for Gene Delivery and Lineage-Specific Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell* **5,** 332–342 (2009).

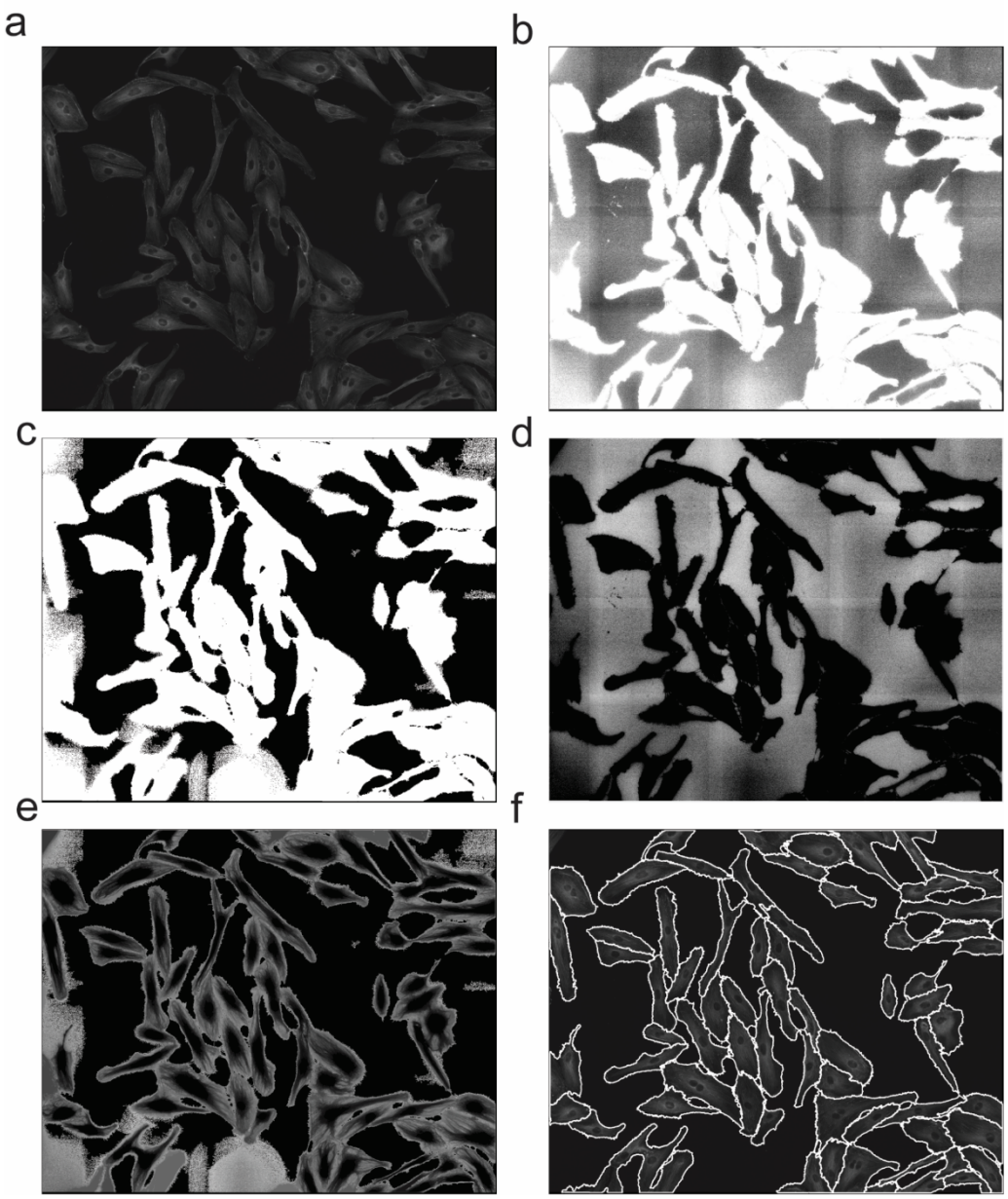
23. Ryan, D. T., Hu, J., Long, B. L. & Qutub, A. A. Predicting Endothelial Cell Phenotypes in Angiogenesis. in *ASME 2013 2nd Global Congress on NanoEngineering for Medicine and Biology* V001T05A013-V001T05A013 (ASME, 2013). doi:10.1115/NEMB2013-93124

24. Bounova, G. & De Weck, O. Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **85,** (2012).

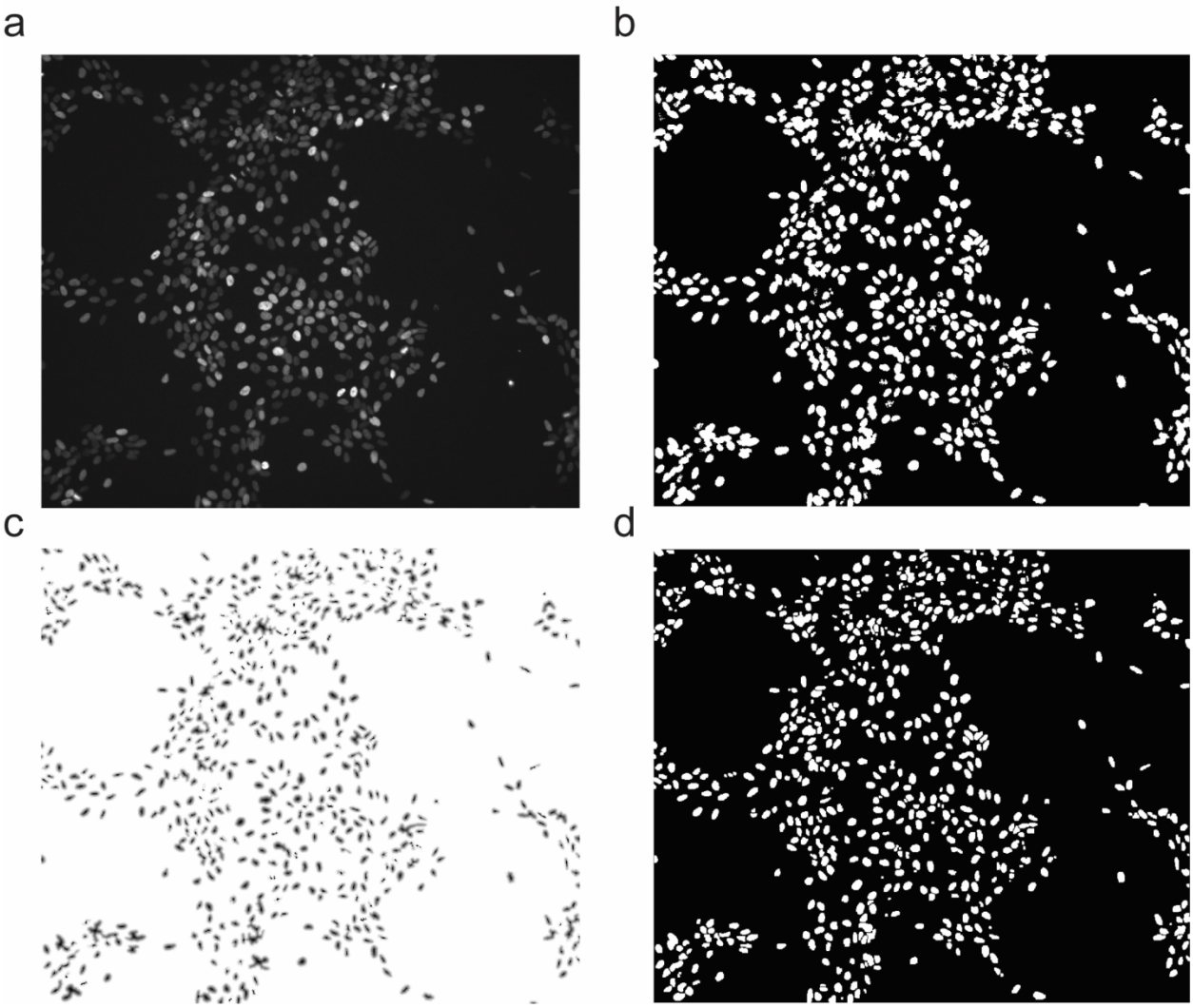
cytoNet: Network Analysis of Cell Communities

Arun S. Mahadevan, Byron L. Long, David T. Ryan, George L. Britton, Aryeh Warmflash, Jacob T. Robinson, Amina A. Qutub

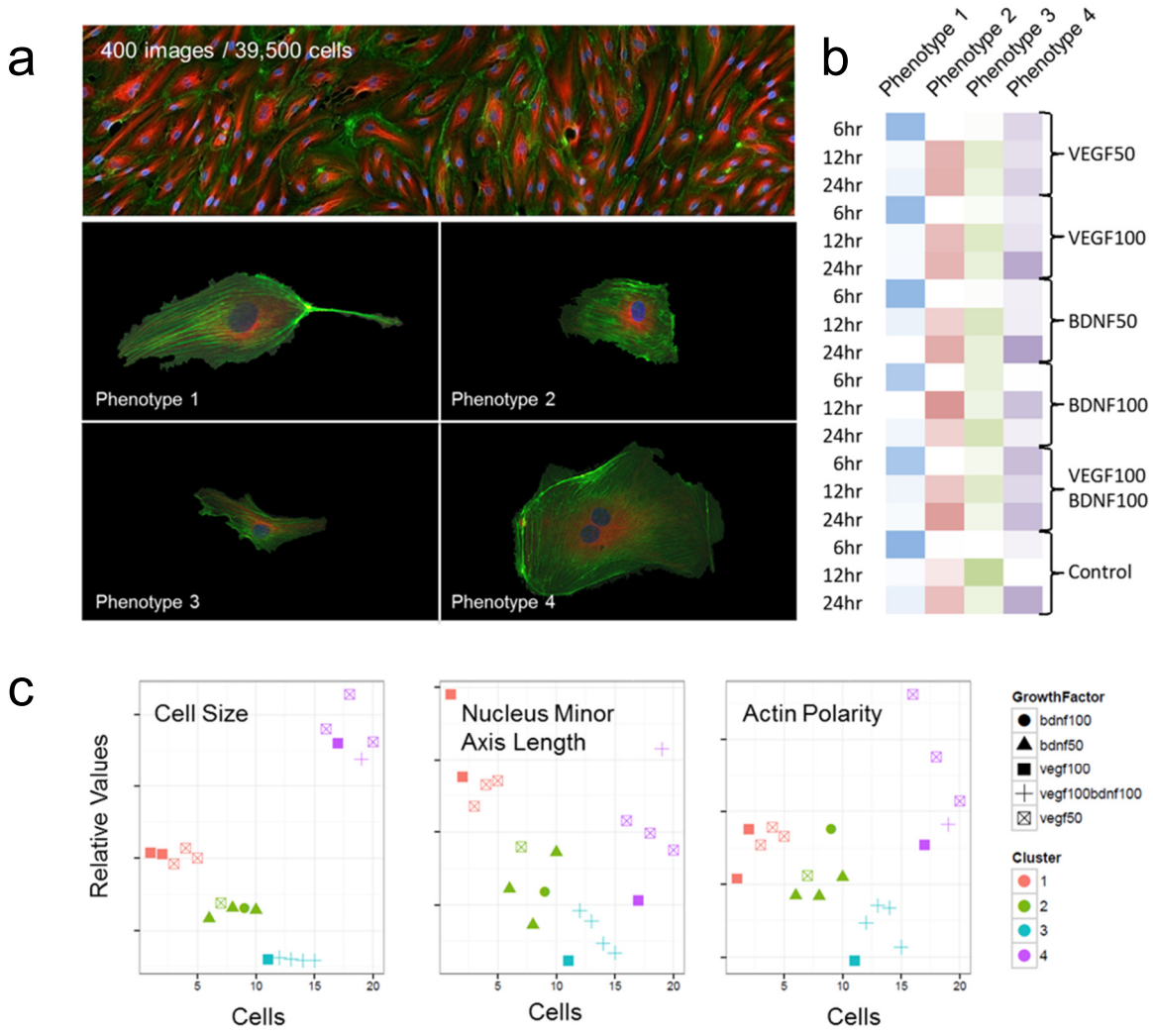
Supplementary Figure 1	Image segmentation of HUVEC immunofluorescence images.
Supplementary Figure 2	Image segmentation of FUCCI-ReN nucleus images
Supplementary Figure 3	Unique HUVEC morphological phenotypes in response to neurotrophic factors
Supplementary Figure 4	Correlation heatmap of local network metrics and morphology metrics for immunofluorescence HUVEC images
Supplementary Table 1	Metrics describing local community, calculated at the level of individual cells
Supplementary Table 2	Metrics used to define cellular architecture.
Supplementary Table 3	Cohen's d effect size for treatment conditions on morphology metrics shown in Figure 2 (h-i) in the main text.
Supplementary Video 1	Time-lapse movie of sparse culture of FUCCI-ReN cells
Supplementary Video 2	Time-lapse movie of sparse culture of FUCCI-ReN cells, with cell boundaries and graph representation overlaid
Supplementary Video 3	Time-lapse movie of dense culture of FUCCI-ReN cells
Supplementary Video 4	Time-lapse movie of dense culture of FUCCI-ReN cells, with cell boundaries and graph representation overlaid
Supplementary Methods 1	Instructions for using the web-based user interface



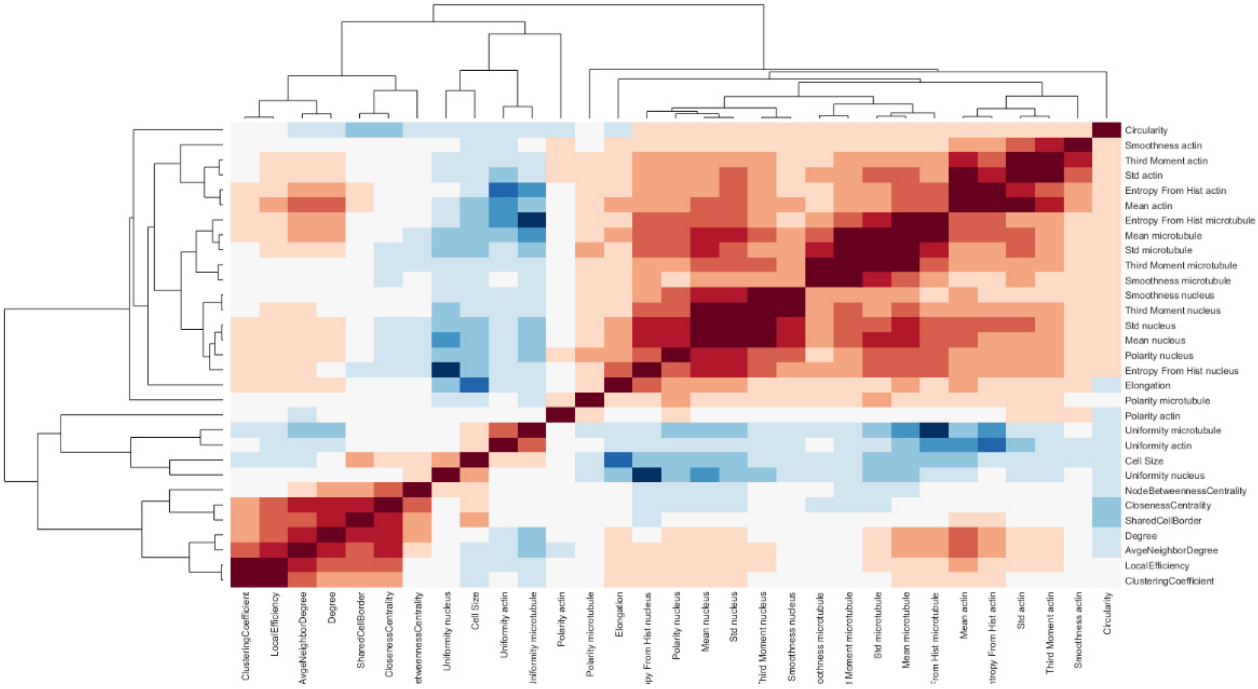
Supplementary Figure 1. Image segmentation of HUVEC immunofluorescence images. (a) Original grayscale image. **(b)** Image after adaptive histogram equalization and Gaussian filtering. **(c)** Binary image obtained using Otsu's threshold, with small objects removed. **(d)** Complement of filtered image in **(b)**. **(e)** Watershed basins obtained through imposing minimum of images in **(d)** and the marker image (obtained by combining the binary image in **(c)** and the image obtained through binarization of microtubules and nuclei). **(f)** Final cell borders.



Supplementary Figure 2. Image processing steps for FUCCI-ReN nucleus images. (a) Fluorescence image from H2B-Cerulean channel marking all nuclei. **(b)** Binary mask obtained through adaptive thresholding. **(c)** Image obtained through imposing minimum of distance transform of binary image in **(b)** and local minima. This image serves as seeds for the watershed algorithm. **(d)** Final mask obtained after watershed transform.



Supplementary Figure 3. Unique HUVEC morphological phenotypes in response to stimulation by neurotrophic factors vascular endothelial growth factor (VEGF) and brain-derived neurotrophic factor (BDNF). (a) The 'average cell phenotype' displayed of each cluster obtained from clustering morphological features of 39,500 cells from 400 monolayer images of endothelial cells stained for actin (green), microtubules (red) and Hoescht (blue). K-means clustering was performed across all cells for all conditions, with the optimal number of four clusters identified by Silhouette scores, using Euclidean distance as the similarity metric. Morphological features within each phenotype captured quantitatively what has been observed qualitatively during coordinated cell processes of angiogenesis. As an example, phenotype 4, a multi-nucleated cell showing nuclear asymmetry, indicates a proliferative state. Elongated cells with asymmetrical actin distribution and relatively large nuclei show both migratory and proliferative properties (e.g., phenotype 1) **(b)** Cluster membership of cells after stimulation by VEGF and BDNF for 6, 12 and 24 hrs is displayed, corresponding to clusters from (a). Cells in each condition change their cluster assignment over time as a function of stimuli. Shading indicates % of cells in that cluster (darker = higher %). **(c)** Example feature metrics of endothelial cell phenotypes resulting from BDNF and VEGF stimulation. Cell size, the minor axis length of the nucleus, and actin polarity were among features that contributed the most variance in the endothelial dataset. Each point in the plot represents metric values for the average cell for all time-points for each of the five growth factor conditions. Cell order (x-axis) is arbitrary, while relative values (y-axis) are normalized to the maximum value per metric across all cells. Color corresponds to the closest cell phenotype from (a) to which an average cell from each stimulation condition belongs.



Supplementary Figure 4. Correlation heatmap of local network metrics and morphology metrics for immunofluorescence HUVEC images. All morphology and local network metrics (**Supplementary Table 1, Supplementary Table 2**) were combined into a single matrix. The cluster dendrogram was obtained through hierarchical clustering of the covariance matrix using Pearson's correlation as the similarity metric.

Supplementary Table 1. Metrics describing local community, calculated at the level of individual cells.

Graph Metrics	Symbol	Definition
Degree*	k	Number of neighbors (one link away)
Average Neighbor Degree	k_n	Average degree of all neighboring cells
Clustering Coefficient	C	Fraction of total possible links among the neighbors of a node that are actually present, averaged across all neighbors
Local Efficiency	E_l	Average shortest path length in local neighborhood
Node Closeness Centrality	c_n	Sum of reciprocal distances to all other nodes
Node Betweenness Centrality	w_n	Number of shortest paths that pass through the node
Shared Border Length**	S_l	Length of border shared with neighboring cells, in pixels

*The fraction of neighbors of a certain phenotype is used to compute the neighborhood similarity score, N_S

**Valid only for type I graphs (adjacency evaluated through shared pixels)

Supplementary Table 2. Metrics used to define cellular architecture.

Metrics	Definition	Mathematical Representation
Cell Size	Cell spread area	A_C
Circularity	Shape factor	$\frac{4\pi . A_C}{P_C}$ Where P_C = perimeter of cell
Elongation	Shape factor	$\frac{P_C}{A_C}$
Polarity*	Distance between center of mass of stain and the centroid of the cell	$\sqrt{\left(\chi_{C,x} - \Omega_{S,x}\right)^2 + \left(\chi_{C,y} - \Omega_{S,y}\right)^2}$ Where Ω_S = center of mass of stain χ_C = centroid of cell
Mean*	First moment of grayscale stain intensity distribution	$\sum_{i=0}^{255} \frac{i}{255} . p$ Where p is the histogram counts of the image for pixel intensities, with 256 possible bins for a grayscale image
Standard Deviation*	Second moment of grayscale stain intensity distribution	$\sqrt{\sum_{i=0}^{255} \left(\frac{i}{255}\right)^2 . p}$
Third Moment*	Third moment of stain intensity distribution	$\frac{1}{255^2} \sum_{i=0}^{255} \frac{i}{255}^3 . p$
Smoothness*	Measure of smoothness of stain	$1 - \frac{1}{1 + \left(\frac{1}{255^2} \Sigma_{i=0}^{255} \frac{i}{255}^2 . p\right)}$
Entropy from Histogram*	Measure of randomness of the stain intensity	$-\sum p . \log_2(p)$
Uniformity*	Sum of squared elements in the histogram counts of the image for pixel intensities.	$\sum p^2$

*Computed for all 3 stains (nucleus, actin and microtubules)

Supplementary Table 3. Cohen's d effect size for treatment conditions on morphology metrics shown in Figure 2(h-i) in the main text.

Morphology Metric	Treatment Condition	Cohen's d Effect Size			
Cell.Size		6hr (uncorrected*)	12hr (uncorrected)	6hr (corrected**)	12hr (corrected)
	BDNF50	0.256	0.217	0.148	0.170
	VEGF50	0.151	0.023	0.093	0.068
Mean.Actin		6hr (uncorrected)	12hr (uncorrected)	6hr (corrected)	12hr (corrected)
	BDNF50	0.381	1.020	0.091	0.873
	BDNF100	0.517	2.522	0.260	1.959
	VEGF50	1.121	1.018	0.348	0.740
	VEGF100	1.267	2.269	0.284	1.808

* no correction for network metrics
** correction applied for network metrics

Supplementary Video 1

Time-lapse movie of sparse culture of FUCCI-ReN cells. Magenta: Cdt1-mCherry, Green: Geminin-Venus. Time stamp is shown on top right corner.

Supplementary Video 2

Time-lapse movie of sparse culture of FUCCI-ReN cells with graph overlay. Movie displays phase contrast frames from movie in Supplementary Video 1, with Cdt1(-)/mCherry(+) nuclei circled in magenta, Geminin(-)/Venus(+) nuclei circled in green and mCherry(-)/Venus(-) nuclei circled in blue. Yellow lines represents proximity edges.

Supplementary Video 3

Time-lapse movie of dense culture of FUCCI-ReN cells. Magenta: Cdt1-mCherry, Green: Geminin-Venus. Time stamp is shown on top right corner.

Supplementary Video 4

Time-lapse movie of dense culture of FUCCI-ReN cells with graph overlay. Movie displays phase contrast frames from movie in Supplementary Video 3, with Cdt1(-)/mCherry(+) nuclei circled in magenta, Geminin(-)/Venus(+) nuclei circled in green and mCherry(-)/Venus(-) nuclei circled in blue. Yellow lines represents proximity edges.

Supplementary Methods 1

Instructions for using the web-based user interface

1. Go to http://qutublab.rice.edu/cytoNet/

Explanation of parameters and input format can also be downloaded there.

2. Select images

a) Select image files by clicking on the 'Choose Files' button to start a file selection dialog box. Multiple files can be selected by: a) clicking on a file while holding down the control key (command key in MacOS; b) clicking and dragging; or c) entering control-a (command-a in MacOS) to select all files in a directory or folder. Color input images are first converted to grayscale images by cytoNet before being processed as previously described. Binary input images are considered to already be binary masks and the segmentation step is skipped.

b) Select image number. Some image file formats such as tiff support the storage of multiple images per file. If your images are not stored in tiff files, you may skip this parameter. Otherwise you can specify which images in each file will be processed by using comma separated (1-based) indices. Hence 1,3 indicates the first and third images in each file. Indices must be specified in increasing order.

c) For demonstration purposes, cytoNet also provides images if you do not have your own.

3. Select edge determination method

a) Edges between nearby objects are determined by the distance between their centroids.

b) Edges between touching objects are determined by the sharing of border pixels.

4. Select adjacency threshold

When edges are determined by the distance between centroids, an adjacency threshold parameter is required. The adjacency threshold determines the maximum distance between two centroids at which an edge is created in the following way. Let a_1 and a_2 be the area of two objects with centroids c_1 and c_2 respectively. For each object, compute its effective radius: $r_i = \sqrt{\frac{a_i}{\pi}}$. A graph edge is placed between two objects (vertices) whenever the distance between their centroids is within the adjusted sum of their effective radii: $distance(c_1, c_2) \leq S \cdot (r_1 + r_2)$ where S is the user defined adjacency threshold parameter.

5. Enter an email address. cytoNet will use this email address to inform you that processing is complete.

6. Click the Submit button.

7. cytoNet will send you an email message indicating that your request has been accepted. This message includes a Request ID that you can use to check on the progress of your request. cytoNet will also send you an email message informing you that processing has ended for your request.

8. When your request has been successfully processed, you may download your results. Note that your results will be available for only a limited amount of time.

Results are formatted as follows. Global metrics are tabulated in a file called 'GlobalMetrics.csv' for all images in the input folder. Local metrics, computed on a per-cell basis are tabulated in a separate file for each image called 'LocalMetrics_filename.csv', where filename is the original file name. Also, basic morphology metrics (size, elongation, circularity and stain intensity) are tabulated in a separate file for each image called SingleCellMetrics_filename.csv, where filename is the original file name. Processed images are also created for each image in the input folder, called 'filename_processed.tif' where the original image is overlaid with cell indices, object outlines (red) and spatial proximity edges (yellow). Cell

indices displayed in the processed images are used in the first column of local metrics and single cell metric files.