Science Signaling
AAAS

www.sciencesignaling.org/cgi/content/full/9/416/ra20/DC1

Supplementary Materials for

Endothelial cells decode VEGF-mediated Ca^{2+} signaling patterns to produce distinct functional responses

David P. Noren, Wesley H. Chou, Sung Hoon Lee, Amina A. Qutub, Aryeh Warmflash, Daniel S. Wagner, Aleksander S. Popel,* Andre Levchenko*

*Corresponding author. E-mail: andre.levchenko@yale.edu (A.L.); apopel@jhu.edu (A.S.P.)

Published 23 February 2016, *Sci. Signal.* **9**, ra20 (2016)
DOI: 10.1126/scisignal.aad3188

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencesignaling.org/cgi/content/full/9/416/ra20/DC1)

Text S1: Detailed description of the probabilistic model

Simulating cell behavior under constant VEGF

We developed a probabilistic model relating the different features of Ca^{2+} signaling and phenotype responses through a directed acyclic graph. The model was not intended to capture the detailed biochemical interactions within the signaling network, but rather describe the probabilistic relationships between quantitatively categorized metrics of the signaling responses measured at different network levels. More specifically, the model integrates information about waveform selection, waveform features, and cell behavior, allowing us to mathematically describe the cell decision-making process. Each node in the network (Fig 4C, expanded model, black schematic) corresponds to a discrete random variable and was described by a multinomial distribution as shown below.

$$Pr(X|n,\mu) = \left(\frac{n!}{x_1!\ldots x_K!}\right)\prod_{k=1}^{K}\mu_k^{x_k}$$

where X is a vector $x_1 \ldots x_k$ giving the number of times each state was selected, n is the number of trials, K is the number of states, and μ_k is the probability the kth state.

Here the variable (**w**) describes which waveform each cell exhibits and can take on one of three exclusive states, NS, RS, or LP. In the case of NFAT activation, none of the cells with LP waveforms were grouped in the high NFAT activation category (Fig 2F), therefore only the RS state allowed further progression down the NFAT network path. Conversely, RS waveforms did not enhance cell migration sufficiently above baseline values to be included in the high activation group (Fig 3A), and therefore only LP states allowed travel down the migration path. Since NS waveforms did not activate NFAT or lead to enhanced cell motility, this state did not lead to either phenotype. The frequency variable (**f**) describes the average spike occurrence exhibited by the cell in the RS state and is categorized into low, moderate, and high frequencies.

These categories were selected to correspond to the different regions shown in Fig 2F, such that low, moderate, and high correspond to f≤3 peaks/hr, 3 peaks/hr < f < 9 peaks/hr, and f≥9 peaks/hr, respectively (illustrated in Fig S3D). Similarly, $\mathbf{t_d}$, the duration of low amounts of sustained Ca^{2+} observed during LP, can take on three states and is categorized into short (<20 min), moderate (20-40 min), and long (>40min) (Fig S3E). Here LP Ca^{2+} signaling is considered sustained as long as the amounts remain at least 15% above baseline following the transient peak. Both the NFAT activation (**N**) and the extent of migration (**m**) were categorized as having one of two states, either high (enhanced) or low (not enhanced), as determined by k-means analysis. The regions corresponding to these categories are shown in Fig S3D &E. (separated by the dotted line)

For each variable, the model parameter μ was derived from experimental data using maximum likelihood (ML) estimates. Here we used repeated random sub-sampling for model training and parameter cross validation as described below. MLs associated with each state of **w**, **f**, $\mathbf{t_d}$, **N**, and **m**, were trained and validated using separate data sets. In this fashion, parameters corresponding to each individual network node were obtained independently, rather than from global training. For each training simulation, 1/2 of each data set was randomly selected to be used as training data whereas the remaining 1/2 was used for validation. To insure each sample was represented, training simulations were conducted until each sample appeared in at least 25% of the total number of simulations. During one training simulation, MLs pertaining to the different states of each random variable were calculated from the training data and used to predict the distribution of states expected in the validation set. The classification error was taken as the difference in the number of cells predicted and observed for each state. MLs from simulations performing in the top 10% were averaged to obtain final model parameters.

Stochastic simulations were conducted to collect simulated data for each node in the graph. For an individual cell, the state of each variable was drawn from a multinomial distribution using the values of μ obtained as described above. Data from an individual cell was considered to represent an independent trial. Here we increased the number of simulated cells until the simulation output varied by less than 1.5% among 10 simulations, corresponding to 10,000 cells for each simulation. As was the case in the corresponding experiments, the concentration of VEGF is considered constant during the simulations. A detailed illustration of the graphical model, including parameters, is shown in Fig S3F. To determine the impact of VEGF dependence on waveform features, the model was re-trained by pooling experimental data for **f** and $\mathbf{t_d}$ for all values of VEGF. Simulations to predict the behavior of the model without stochasticity downstream of waveform selection, **w**, that is, without variability in the interpretation of LP and RS, were conducted by setting the MLs for NFAT and migration to 1 for every value of **f** and $\mathbf{t_d}$ respectively.

Simulating cell behavior in a VEGF gradient

Next we applied the model to predict the behavior of cells migrating in a gradient of VEGF. By virtue of traveling in a gradient, cells experience different concentrations of VEGF in time. Therefore, we used the model described above to conduct a dynamic simulation where cell decisions were evaluated at discrete time steps. Unlike the simulations described above, since the values of VEGF in a gradient are continuous, linear interpolation was used to determine MLs in between the tested VEGF values. At the end of each time step, cells were allowed to evaluate or re-evaluate their choice of waveform. Although the network structure of the model remained the same, the variable **f** was redefined to allow real time evaluation and is referred to as $\mathbf{f_d}$. In this case, $\mathbf{f_d}$ and $\mathbf{t_d}$ were categorized into discrete parts, based on the model time step, with

each section indicating the continued length of time the repeated spikes or sustained amount of Ca^{2+} will be maintained above a minimum amount. For repeated spikes, the minimum spike duration value was approximated from the average frequency at which NFAT activation became enhanced, corresponding to 6 peaks/hour, as shown in Fig 2F. For instance, for a 20 min time step, the first category of $\mathbf{f_d}$ described the probability repeated spikes would persist above 6 peaks/hour for at least 20 min. The second, third, and forth categories held the probabilities that the repeated spikes would continue above this amount for at least 40 min, 60 min, or greater than 60 min respectively, given that they had been sustained during the preceding time interval. As before, $\mathbf{t_d}$ indicated the duration Ca^{2+} is sustained at least 15% above baseline. Values of $\mathbf{f_d}$ or $\mathbf{t_d}$ which were sustained 60 min were assumed to persist during all later time steps (>60 min) unless a different waveform was selected. At each time step, the new values of $\mathbf{f_d}$ and $\mathbf{t_d}$ were evaluated and assessed with those of the previous time step to determine which NFAT or migration category should be used to predict the phenotype activation amount. An illustration of the simulation algorithm is shown in Fig S5A.

Migration speeds were determined by averaging motility values shown in Fig 3A as described below. Cells were first given a baseline migration speed with no directional bias, allowing them to move either up or down the VEGF gradient each time step. This speed was determined from the average non-enhanced migration rate in Fig 3A and was 11 μm/h. This pattern persisted unless the cell achieved the high migration category during a given time step. In this event, cells were assumed to travel with a faster speed and were entirely directed towards the increasing VEGF concentration. These faster speeds were determined by averaging the enhanced migration speeds shown in Fig 3A and corresponded to either 22 μm/h for cells categorized with moderate $\mathbf{t_d}$ values or 30μm/h for cells categorized with long $\mathbf{t_d}$ values. The

simulations predicted cell migration over the course of 5 hours. Here the migration distance, v_{mig}, was normalized by the migration time for comparison with experimental data (see microfluidic chemotaxis assay methods).

NFAT activation was simulated in a similar manner as described for cell migration. Cells were prescribed a baseline NFAT activation corresponding to an NFAT nuclear:cytoplasmic ratio of 5, which was approximately the maximum non-enhanced value experimentally observed. Once a cell maintained an $\mathbf{f_d}$ greater than 20 min, it could reach the enhanced NFAT category. Upon achieving enhanced NFAT, a cell obtained an activation of 20 for that time step, approximately the maximum experimentally observed for the enhanced NFAT group. As noted above, simulations predicted cell behavior over the course of 5 hours. The NFAT activation amount recorded at each time step was averaged over the duration of the simulation. These results are shown for individual cells as a function of $\mathbf{f_d}$ (Fig 5D).

Figure S1

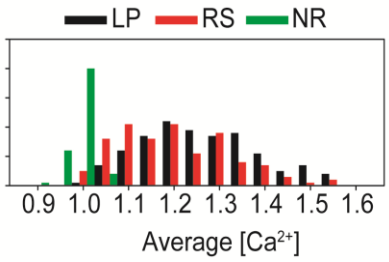

Figure S1: Distribution of averaged Ca^{2+} concentrations. The distribution of time averaged Ca^{2+} concentrations observed for cells following stimulation with 5 ng/ml VEGF for 30 min (n=326 cells from N = 3 independent experiments).

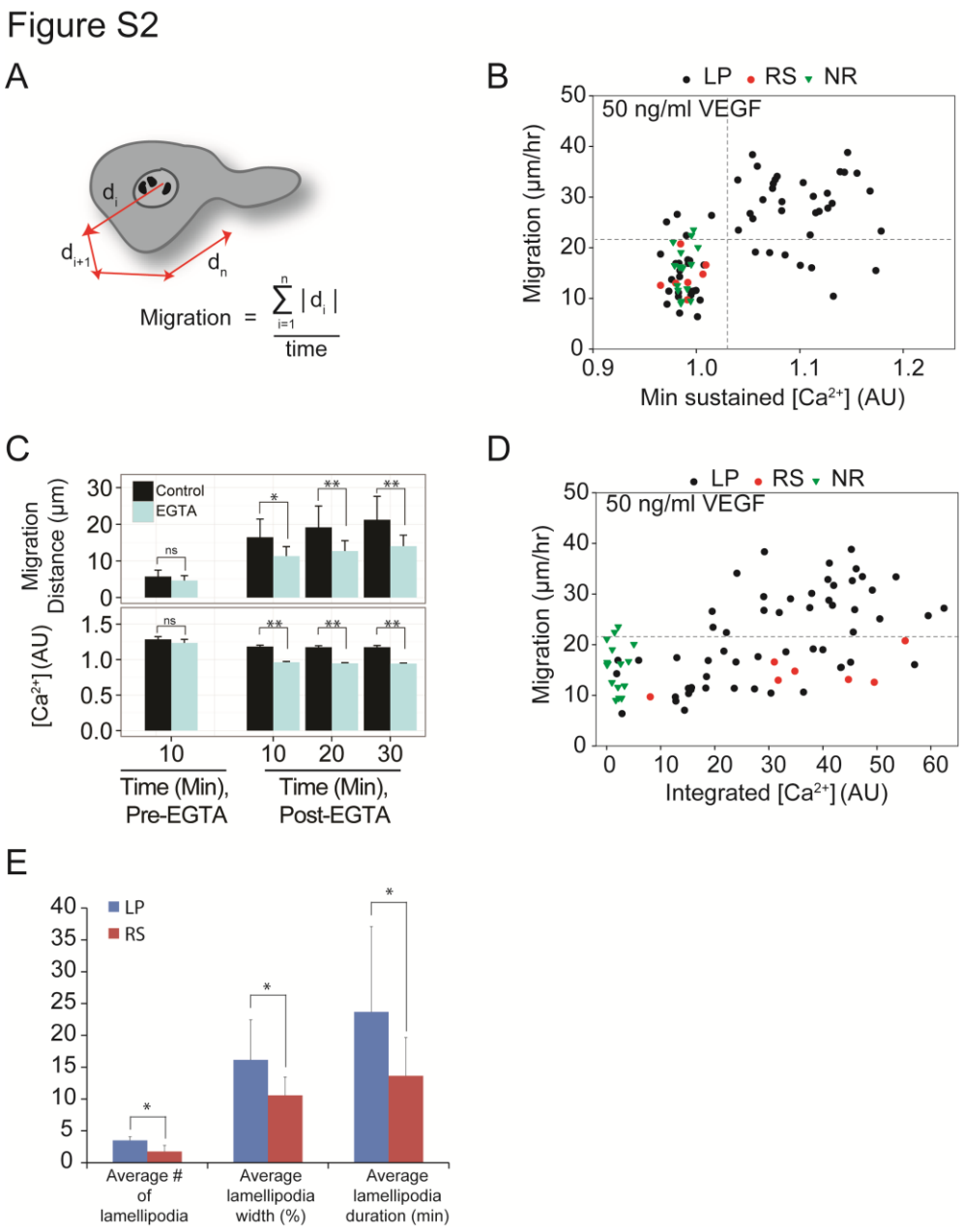

Figure S2: Additional analysis of cell motility. (**A**) Illustration showing the calculation of migration distance. (**B**) Migration as a function of the minimum Ca^{2+} concentration reached during the later portion of Ca^{2+} signaling (20 min – 60 min). Cells that returned to baseline are separated from those with sustained Ca^{2+} concentrations (vertical grey line). Migration values from cells with Ca^{2+} waveforms that returned to baseline, namely NR (green), RS (red), and some LP (black), clustered together, whereas those that were above baseline values achieved higher migration values (grouped by k-means analysis, shown separated by the horizontal grey line). Data was collected for 1 hour following simulation with 50ng/ml VEGF (n=83 cells, N=3 independent experiments). (**C**) Cells exposed to EGTA following 30 min VEGF stimulation (blue) exhibited reduced Ca^{2+} (bottom) and migration distances (top) compared to those without

EGTA (black) (n=45 cells, N = 2 independent experiments, * $p < 0.03$, **$p < 0.016$, Student t-test). (**D**) Migration as a function of integrated Ca^{2+} concentrations determined after 1 hour following stimulation with 50ng/ml VEGF. (**E**) The average number, width measured as a percentile of the cell perimeter, and duration of lamellipodia for LP and RS cells (n = 8 cells, N=2 independent experiments, * $p< 0.0145$, Student t-test). Error bars show 1 full standard deviation above the mean.

Figure S3

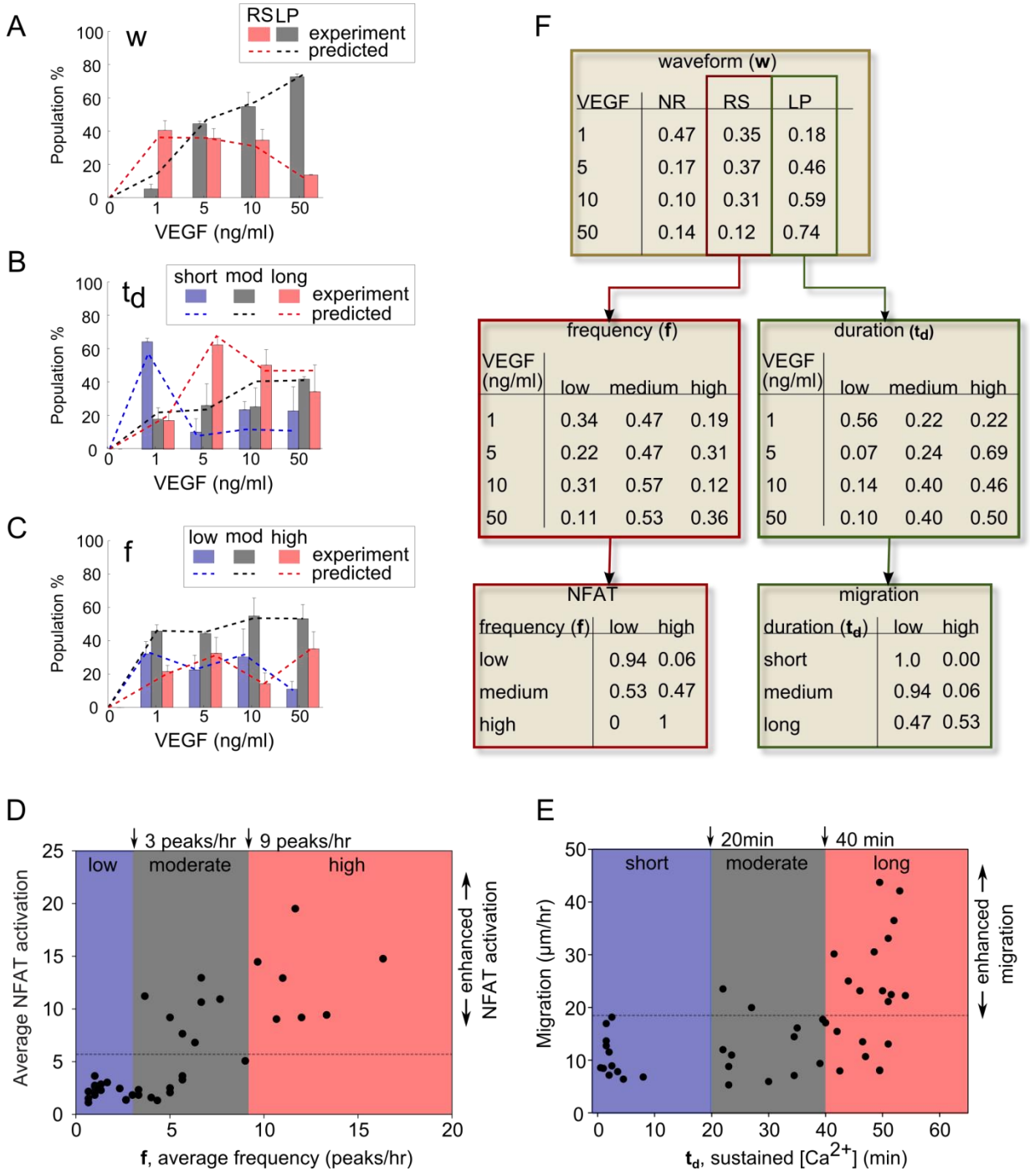

Figure S3: Probabilistic model simulation of responses to spatially uniform VEGF. (**A**) Comparison between predicted distribution of waveforms (dashed lines, black=LP, red=RS), resulting from model training, and experimental measurements (bars, black=LP, red=RS) for each concentration of VEGF (n =780 cells, N = 3 independent experiments). (**B**) Predicted (dashed lines) and observed (bars) percent of cells with t_d values categorized as short (blue), moderate (black), or long (red) (n=197 cells, N=3 independent experiments). (**C**) Predicted

(dashed lines) and observed (bars) percent of cells with f values categorized as low (blue), moderate (black), or high (red) (n=583 cells, N=3 independent experiments). (D) Classification of f into low, moderate, and high categories such that they correspond to only low, high and low, and only high NFAT activation, respectively. (E) Similar classification of td into short, moderate, and long categories based on the differing percentage of cells exhibiting enhanced migration. (**F**) Illustration of detailed model structure including probabilities used to determine the values of each random variable.

Figure S4

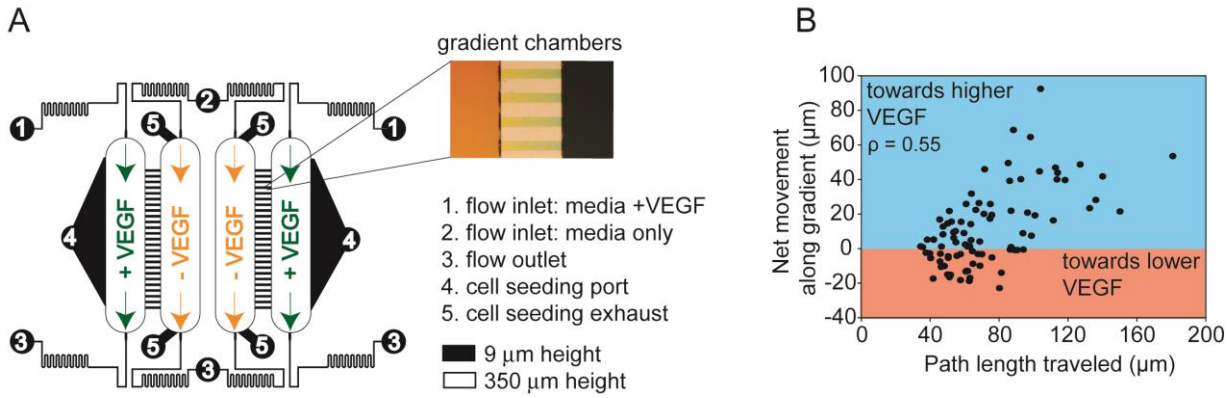

Figure S4: Analysis of cell migration in a microfluidic gradient. (**A**) Design of the microfluidic gradient generating chip. Cells are placed in gradient test chambers using seeding ports (4) and exhaust ports (5). Cell media, with VEGF (port 1) and without (port 2), is continuously perfused through the system. High resistance to fluid flow in the gradient test chambers, due to their shallow height (black), as compared to the flanking flow chambers (white), allows the generation of a stable diffusion based gradient (inset, created using green and yellow food coloring dye) without subjecting the cells to shear stress. (**B**) The net movement of cells in a gradient of VEGF (the difference in final and start position), shown as a function of the total path length traveled, as defined in Fig. 3A (n = 91 cells, N=3 independent experiments).

Figure S5

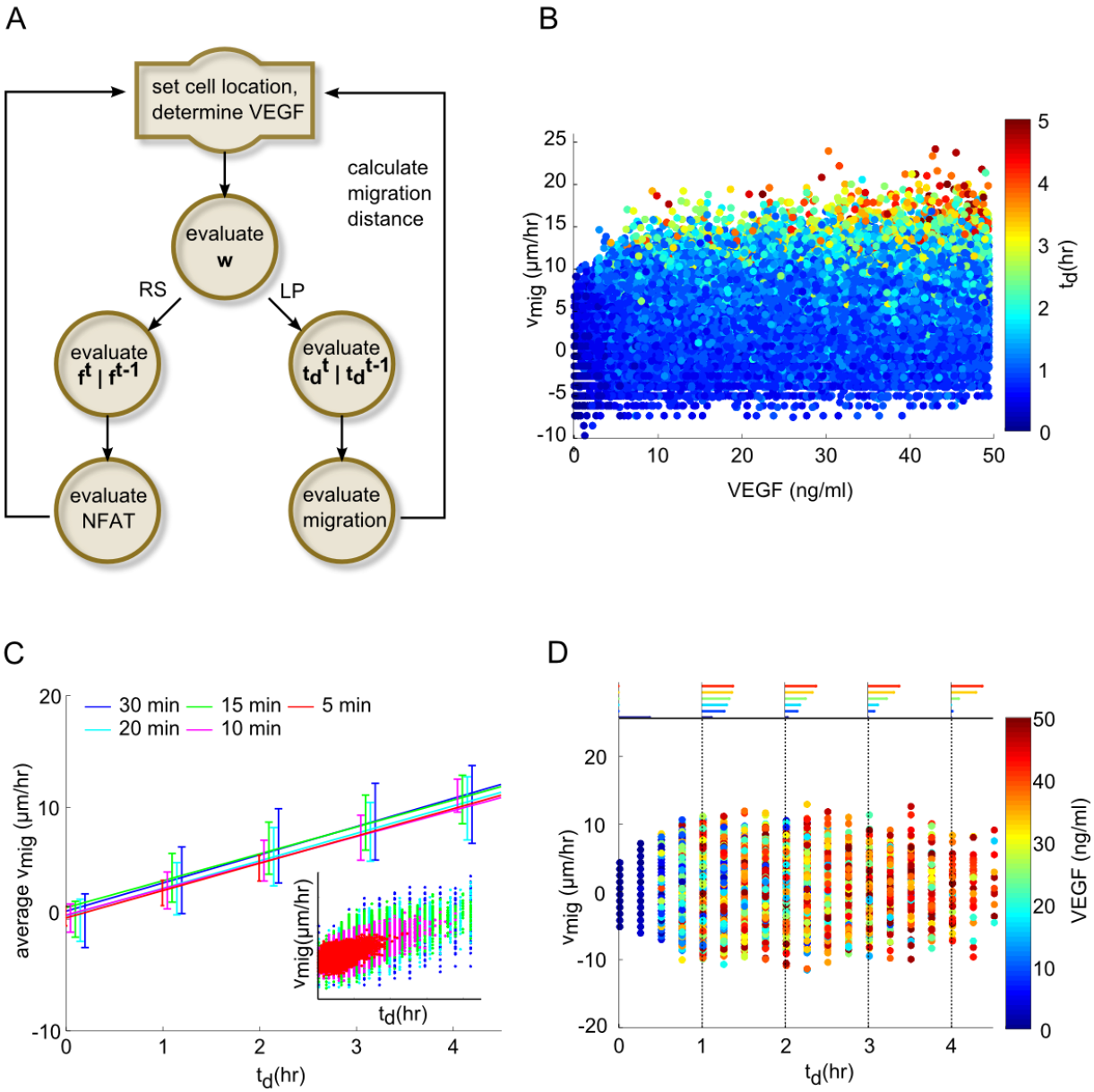

Figure S5: Probabilistic model simulation of cells in a gradient of VEGF. (**A**) Schematic of the gradient response prediction algorithm. (**B**) Predicted v_{mig}, normalized to observation time, as a function of VEGF. Each point represents data from an individual cell and is colored by the initial VEGF concentration experienced by the cell. (**C**) Linear fit to the predicted average v_{mig} as a function of t_d, simulated using different decision-making frequencies (either 5 min, 10 min, 15 min, 20 min, or 30 min), in a gradient of VEGF (decision making events represent potential re-evaluation of the current extracellular VEGF concentration by each cell). Error bars denote standard deviation of v_{mig} averaged at 0, 1 h, 2 h, 3 h, and 4 h, and the inset shows data predicted for individual cells. (**D**) Predicted v_{mig} as a function of t_d simulated without the assumption of a directional bias due to VEGF gradient sensing.

Movie S1: Time-lapse video showing NR, RS, and LP waveforms in PAECs

PAECs expressing YC3.6 were stimulated with 5ng/ml VEGF and imaged every 30 sec for 30 min. The video shows the measured Ca^{2+} regulation (FRET) in cells with a high degree of cell-cell contact.

Movie S2: Time-lapse video showing NR, RS, and LP waveforms in HUVECs

HUVECs were incubated with Indo-1 Ca^{2+} sensitive dye and were stimulated with 5 ng/ml VEGF. Imaged were acquired every 30sec for 30 min. The video shows the measured Ca^{2+} regulation (Indo-1 ratio) in HUVECs with a high degree of cell-cell contact.

Movie S3: Time-lapse video showing simultaneous Ca^{2+} and NFAT signaling. Cells expressing YC3.6 and NFAT-mCherry were stimulated with 5 ng/ml VEGF and imaged for 3 hours at 30 sec intervals. Since CFP intensity decreases upon FRET, increases in Ca^{2+} can be qualitatively visualized by a color change from blue to yellow. Continued NFAT localization (red) to the nucleus and absence from the cytoplasm indicates sustained NFAT signaling. Cell 1 exhibits an RS Ca^{2+} waveform, having a color intermittently changing between blue and yellow, and persistent NFAT localization to the nucleus (quantified in Fig. 2D). Cell 2 exhibits an LP Ca^{2+} waveform, showing a single transient color shift from blue to yellow, and only a transient bias in NFAT nuclear intensity over that observed in the cytoplasm (quantified in Fig. 2E).

Movie S4: Cells exhibiting LP waveforms formed large lamellipodia with pronounced actin translocation. Cells were imaged for 1 hour at 30 sec intervals following stimulation with 5 ng/ml VEGF. Actin translocation was monitored using actin-mCherry (red). Shown is a cell

exhibiting an LP Ca^{2+} waveform, undergoing marked changes in shape and forming large lamellipodia accompanied by actin recruitment.

Movie S5: Cells exhibiting RS waveforms displayed only limited actin recruitment and small filopodia. Cells were imaged as described for movie S3. The cell shown here exhibits an RS Ca^{2+} waveform, does not show pronounced changes in shape, and forms only small filopodia with limited actin recruitment.

Movie S6: Angiogenic sprout formation in zebrafish Time lapse imaging of angiogenic sprout formation originating from the dorsal aorta. Images were taken at a frequency of 1 per minute over the course of 8 hours. GCamp6m was used to monitor Ca^{2+} regulation and is shown in green. mCherry was used as a standard and is shown in red. Sprouts eventually split and fuse to form the dorsal longitudinal anastomotic vessel and remain as intersegmental vessels.