

Supporting Information to "Bursts of activity in collective cell migration"

Oleksandr Chepizhko, Costanza Giampietro, Eleonora Mastrapasqua,
Mehdi Nourazar, Miriam Ascagni, Michela Sugni,
Umberto Fascio, Livio Leggio, Chiara Malinverno,
Giorgio Scita, Stephane Santucci, Mikko J. Alava,
Stefano Zapperi, Caterina A. M. La Porta

Details about the fitting methods

We employ different strategies to obtain the best parameters describing the cluster size distributions in simulations and experiments. We first consider the log-spaced binned probability density functions and perform a least-square fitting with the function

$$f(x) = Ax^{-\tau} \exp(-Cx), \quad (1)$$

To obtain a reliable estimate of the cutoff, we first take the logarithm of the distribution and fit it with the logarithm of Eq. 1. The fitting is performed in python using the pyFitting function (<https://github.com/gdurin/pyFitting>) and the results are reported in table S1 and S2 for experiments and simulations, respectively.

The second strategy relies on the maximum-likelihood estimate. Given the set of measured cluster sizes S_i , we consider the log-likelihood function

$$\mathcal{L} = - \sum_i \log(f(S_i)), \quad (2)$$

where the function $f(x)$ is given by Eq. 1. The function \mathcal{L} is maximized with respect to the parameters A , τ and α . The results are again reported in table S1 and S2. A comparison of the fitted function and the data can be presented in a way that is independent on binning by plotting the cumulative distribution functions in Fig. S10. The theoretical function in this case is given by

$$CDF = \frac{\Gamma(1 - \tau, Cx)}{\Gamma(1 - \tau, Cx_{min})} \quad (3)$$

where $\Gamma(a, b)$ is the incomplete gamma function.

Finally, we collapse together all the experimental probability density function into a unique set. To this end, we first compute the characteristic cluster size $S^* = \frac{\langle S^2 \rangle}{2\langle S \rangle}$ from each experimental data set. We then rescale the cluster sizes, defining a reduced size $s = S/S^*$ and a scaling function $p(s) = P(S) \frac{S^*}{\langle S \rangle}$. Notice that $p(s)$ is not a distribution, since its integral is not equal to one. Once data for different experiments are rescaled according to this prescription, they are joined into a single set which is fitted by the least-square method using the function reported in Eq. 1. The result yields $A = 0.13 \pm 0.01$, $\tau = 1.58 \pm 0.02$ and $C = 0.20 \pm 0.03$.

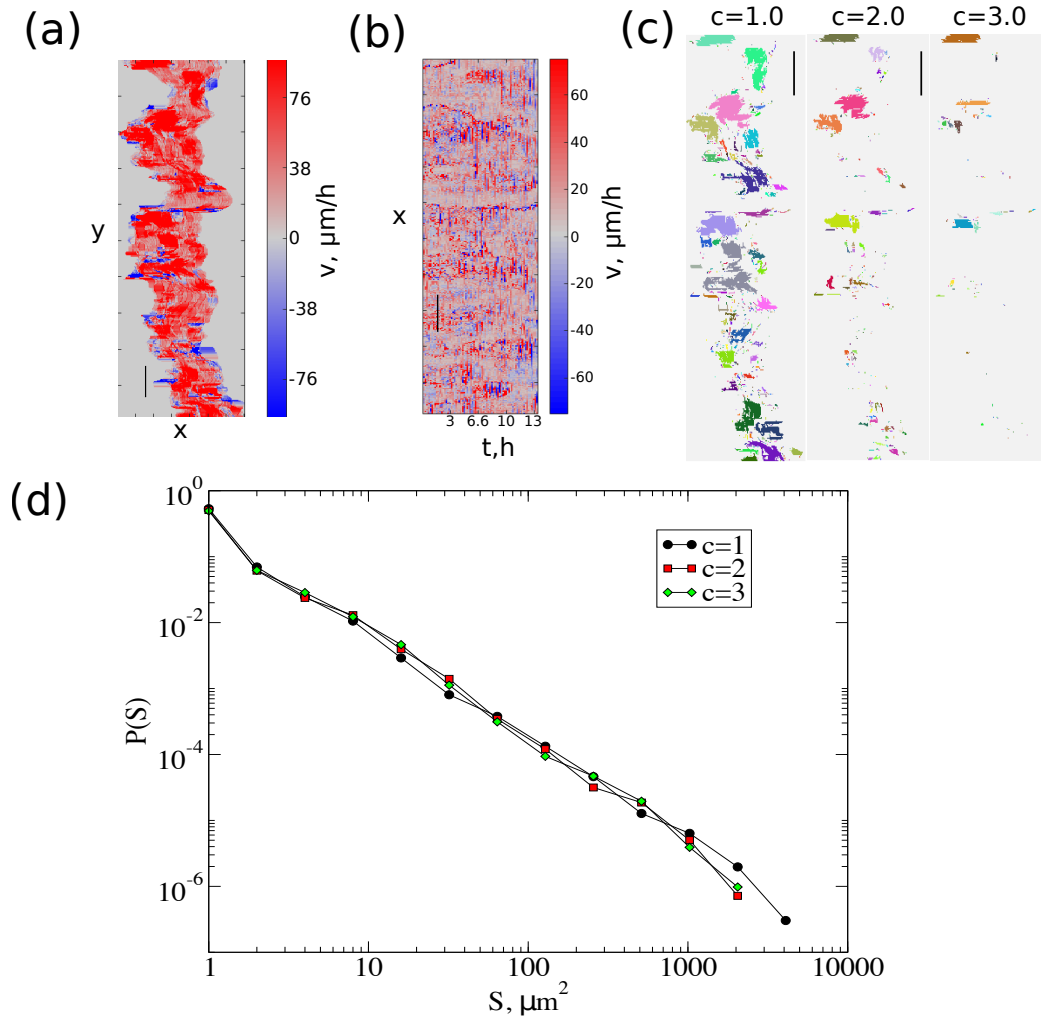


Figure S1: Determination of activity areas from images. Scale bars are $100\mu\text{m}$. a) Spatial front velocity maps. Each pixel is colored according to the velocity of the front once it passes through that location. b) Spatio-temporal velocity map. Each pixel corresponds to the velocity ($v_f(x, t)$) at time t of the front at point x c) Activity clusters are obtained by setting threshold c to the data in panel a. Results obtained with different values of c are reported in three panels. d) Cluster size distributions for different values of c .

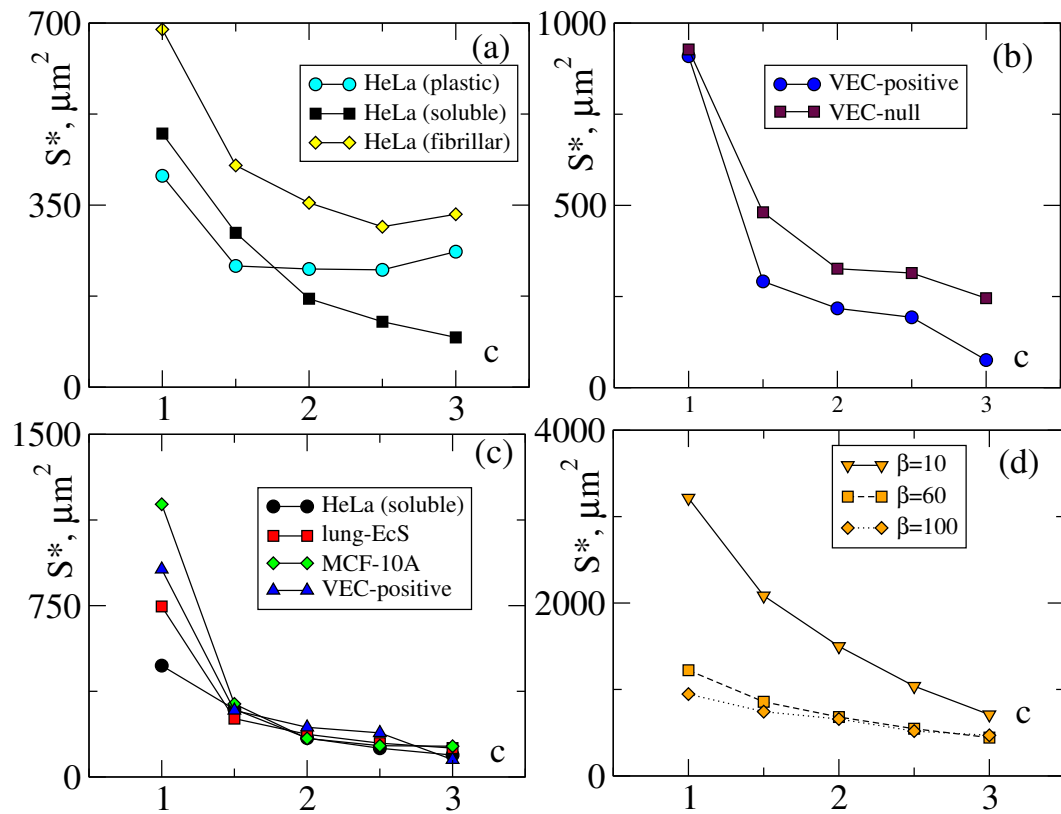


Figure S2: The characteristic cluster size $S^* = 2\langle S^2 \rangle / \langle S \rangle$ as a function of the velocity threshold c . a) HeLa Cells on different substrates. b) Mouse endothelial cells expressing VE-cadherin or not. c) Comparison between different cell lines d) Simulations of the model for different values of β .

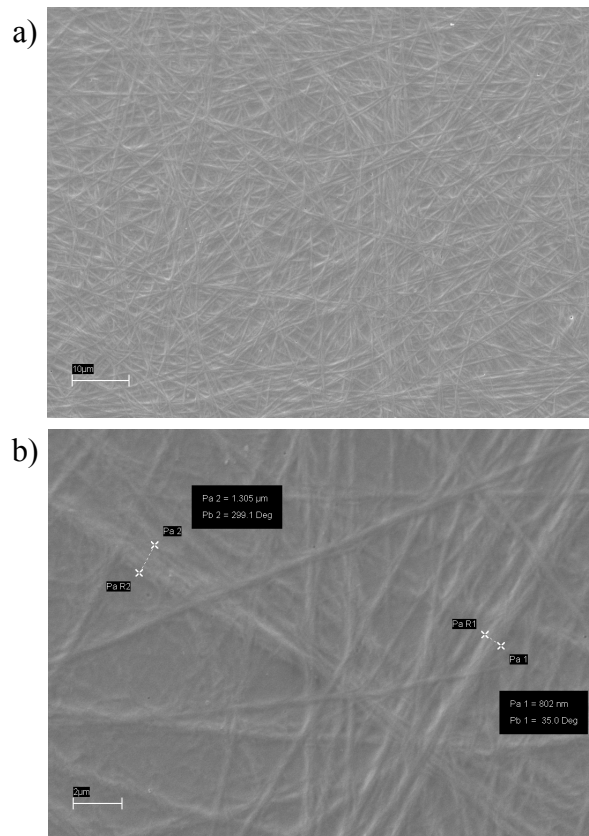


Figure S3: Scanning electron micrograph of the fibrillar collagen substrate at two different magnifications. From the high magnification image b) it is possible to estimate the fiber diameter.

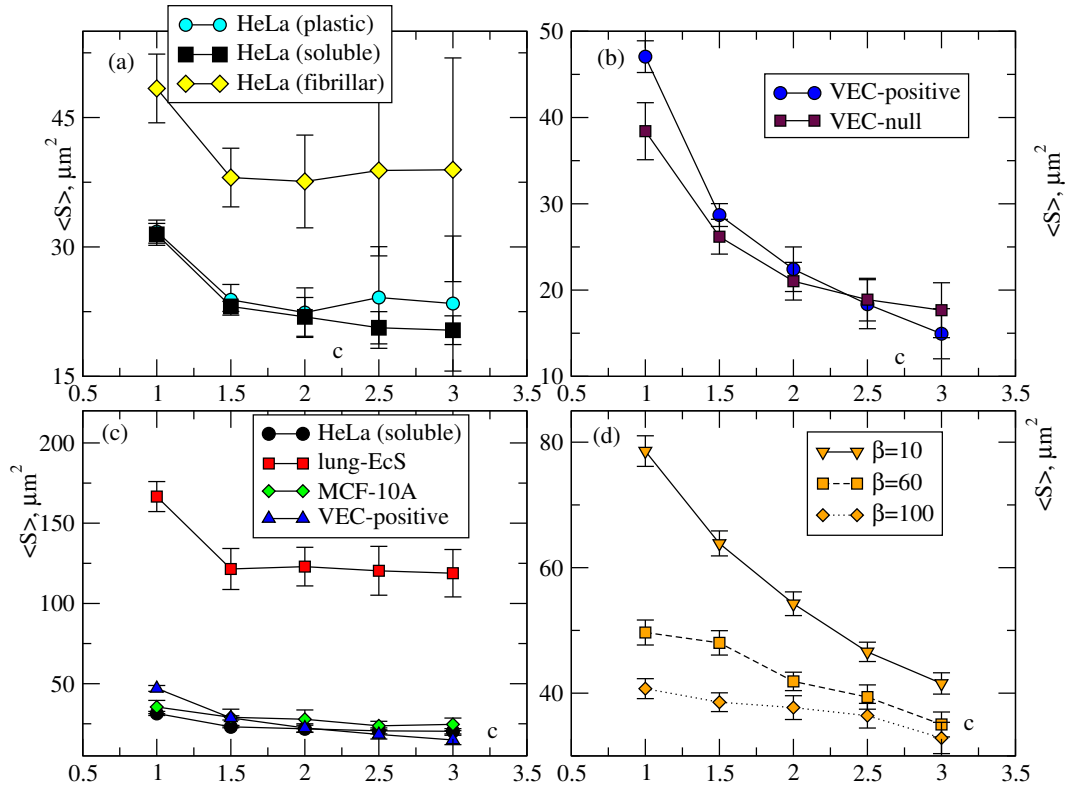


Figure S4: The average cluster size as a function of the velocity threshold c . a) HeLa Cells on different substrates. b) Mouse endothelial cells expressing VE-cadherin or not. c) Comparison between different cell lines d) Simulations of the model for different values of β .

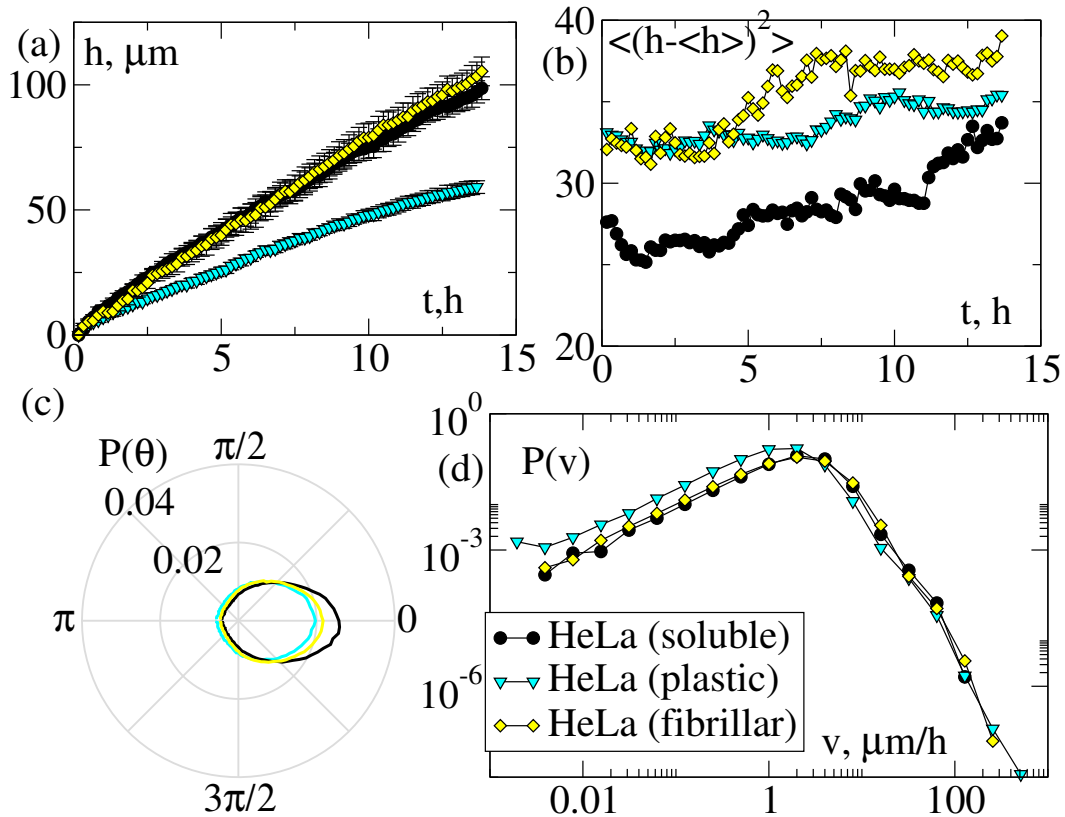


Figure S5: Dependence of the front dynamics on the substrate properties. a) Front position as a function of time indicates that fronts move slowly on plastic substrates. b) The standard deviation of the front (i.e. the front roughness) is larger for cells moving on fibrillar collagen. c) Angular velocity distribution as obtained from PIV. d) Distribution of absolute values of velocities as obtained by PIV.

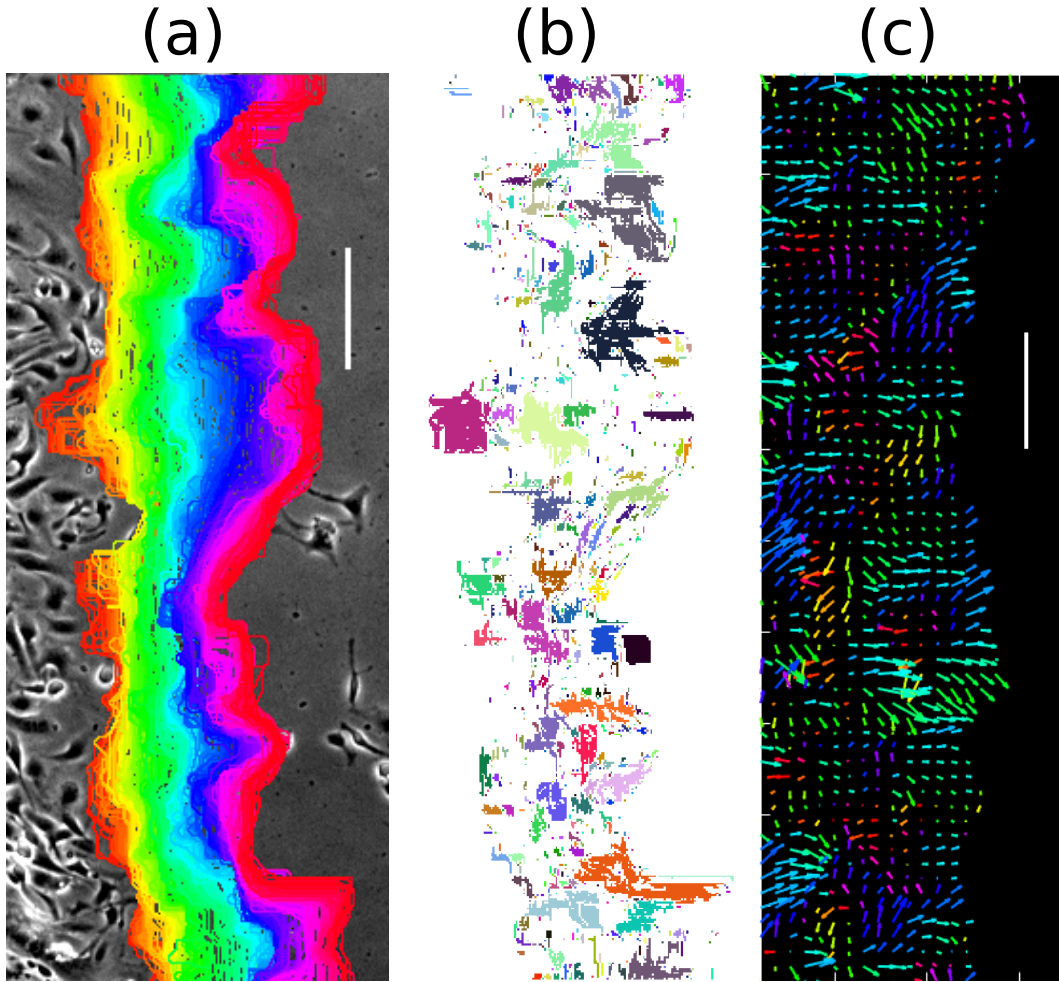


Figure S6: Front dynamics in mouse endothelial cells with VE-cadherin knock down (VEC-null). a) Fronts. b) Activity clusters. c) Results from PIV.

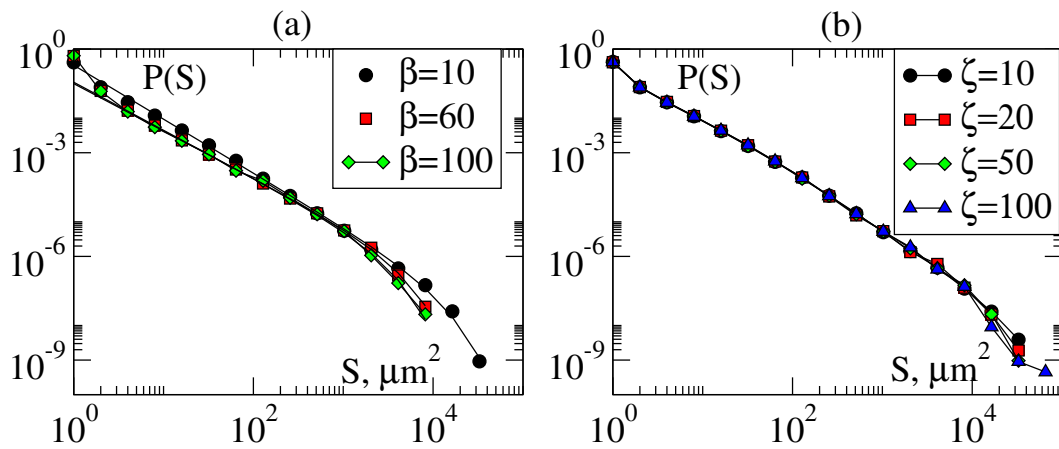


Figure S7: Simulated cluster size distributions a) for different β (the strength of alignment between cells), b) for different ζ (correlation length of the random field).

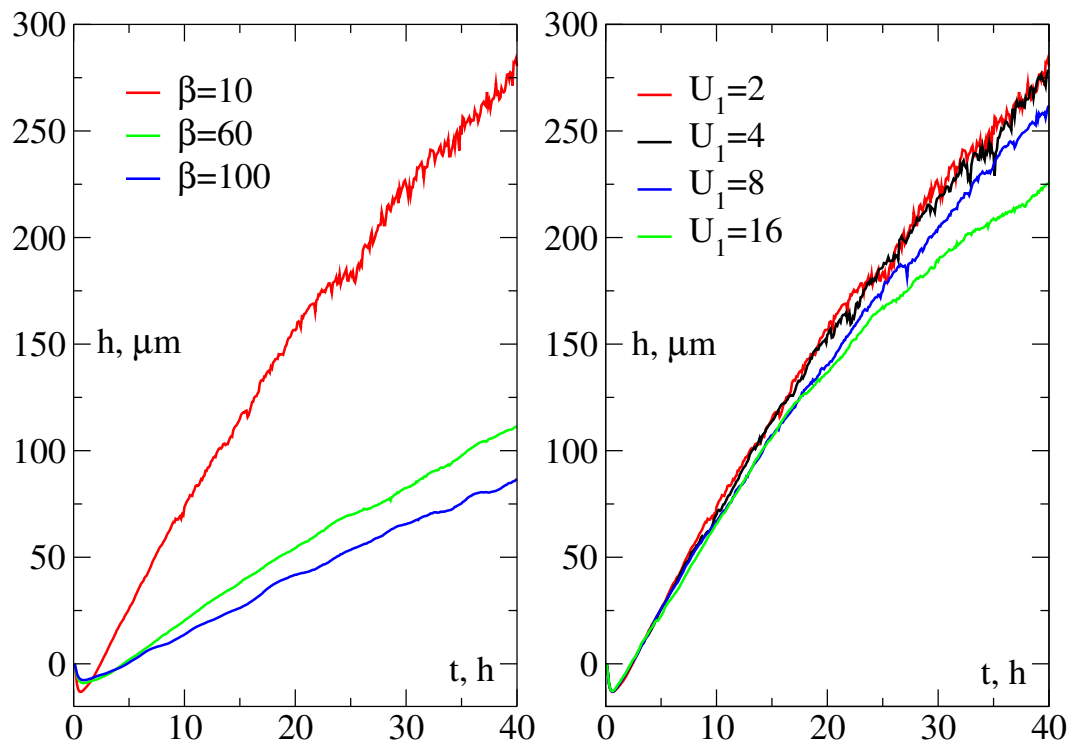


Figure S8: a) Simulations of the model for different values of β show that fronts are slower when β is increased. b) Simulations of the model for different values of the adhesion strength U_1 show that fronts are slower when U_1 is increased.

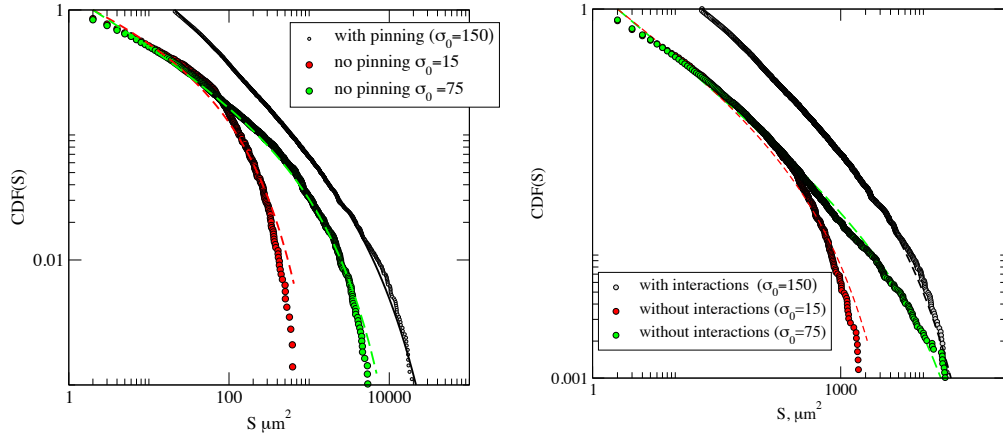


Figure S9: The non-binned cumulative distribution functions (CDF) of the cluster sizes from simulations a) with and without pinning (i. e. no surface particles and no random field) and b) with and without interactions (i.e. $\beta = 0$, $U_0 = U_1 = 0$). The lack of either pinning or interactions leads to clear deviation from a power law distribution, as also shown by fitting the simulation results with the maximum likelihood method. The fit yield scaling exponents $\tau = 1.2$ (panel a, red), $\tau = 1.38$ (panel a, green), $\tau = 1.51$ (panel b, red), $\tau = 1.56$ (panel b, green), but the fitted curves deviate systematically from the data and the scaling regime is extremely small. We also notice that in order to record a front at all in these conditions, one should also reduce the noise parameter σ_0 .

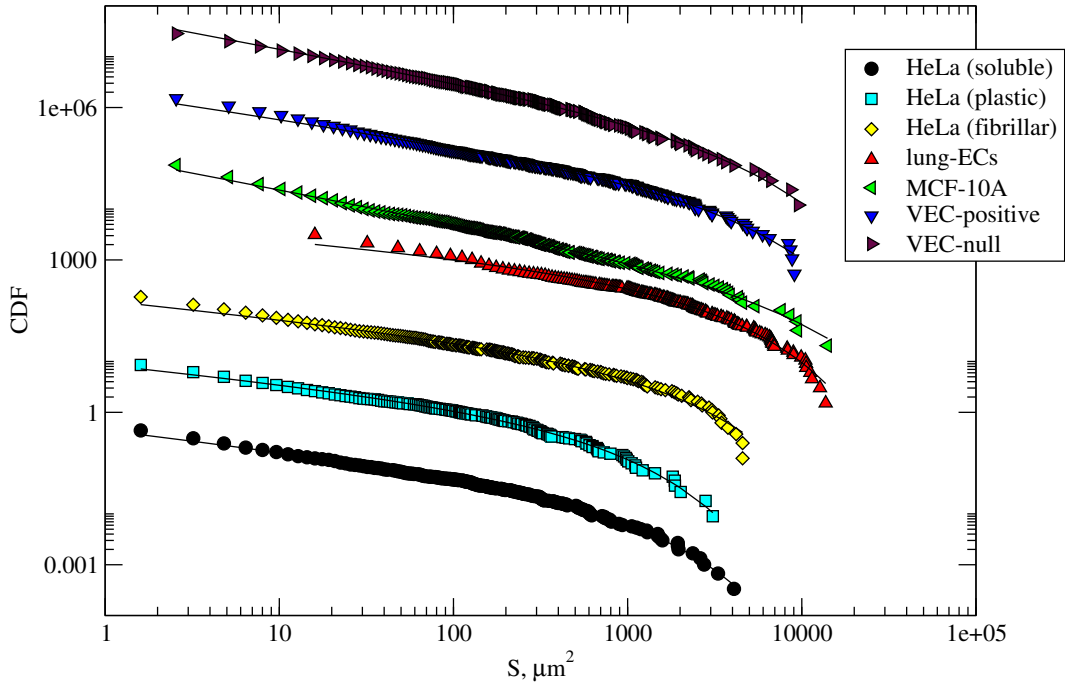


Figure S10: The non-binned cumulative distribution functions (CDF) of the cluster sizes for the different experiments are compared with the results of maximum likelihood estimates.

Name	Fitting			Maximum likelihood		
	A	τ	C	τ	C	x_{min}
HeLa (soluble)	0.18 ± 0.02	1.48 ± 0.03	$0.00054 \pm 5.2 \times 10^{-5}$	1.48 ± 0.10	$(3.2 \pm 2.43) \times 10^{-4}$	12.8
HeLa (plastic)	0.17 ± 0.03	1.47 ± 0.05	0.0006 ± 0.0001	1.49 ± 0.09	$(3.5 \pm 2.8) \times 10^{-4}$	8.0
HeLa (fibrillar)	0.19 ± 0.03	1.49 ± 0.04	$(20 \pm 7) \times 10^{-5}$	1.34 ± 0.14	$(2.82 \pm 2.0) \times 10^{-4}$	28.8
lung-ECs	0.81 ± 0.36	1.58 ± 0.07	$(7.2 \pm 1.4) \times 10^{-5}$	1.45 ± 0.13	$(1.1 \pm 0.6) \times 10^{-4}$	176
MCF-10A	0.17 ± 0.06	1.58 ± 0.08	$(16.7 \pm 9.2) \times 10^{-5}$	1.56 ± 0.12	$(1.2 \pm 1.0) \times 10^{-4}$	43.5
VEC-positive	0.33 ± 0.09	1.62 ± 0.05	$(8.6 \pm 4.9) \times 10^{-5}$	1.55 ± 0.12	$(1.2 \pm 1.0) \times 10^{-4}$	53.8
VEC-null	0.33 ± 0.04	1.62 ± 0.03	$0.0001 \pm 5 \times 10^{-5}$	1.61 ± 0.13	$(1.7 \pm 1.4) \times 10^{-4}$	46.1

Table S1: Best parameters obtained for the cluster distributions for experimental data by least-square fitting of the probability density function and by the maximum likelihood estimate. The smallest value of the cluster size (x_{min}) is also reported in the table.

β	Fitting			Maximum likelihood		
	A	τ	C	τ	C	x_{min}
10	0.34 ± 0.07	1.56 ± 0.04	$(10 \pm 1) \times 10^{-5}$	1.60 ± 0.02	$(6.3 \pm 1.9) \times 10^{-4}$	21
60	0.103 ± 0.007	1.37 ± 0.01	$(3 \pm 0.13) \times 10^{-4}$	1.36 ± 0.05	$(2.5 \pm 0.8) \times 10^{-4}$	16
100	0.11 ± 0.02	1.39 ± 0.04	$(3.8 \pm 0.3) \times 10^{-4}$	1.33 ± 0.03	$(3.9 \pm 0.7) \times 10^{-4}$	6

Table S2: Best parameters obtained for the cluster distributions for simulated data by least-square fitting of the probability density function and by the maximum likelihood estimate. The smallest value of the cluster size (x_{min}) is also reported in the table.

Name	Number of fronts	1 px in μm	1 frame in min.
HeLa (plastic)	6	1.2658	10
HeLa (soluble)	5	1.2658	10
HeLa (fibrillar)	8	1.2658	10
lung-ECs	11	4	5
MCF-10A	8	1.6	5
VEC-positive	4	1.6	2
VEC-null	4	1.6	2

Table S3: Time-lapse parameters. We list the number of fronts analyzed for each case, The size of a pixel in each case and sizes and time step between two frames .

Supplementary video captions

- Video S1 (Left) Time lapse of a representative sheet of Hela cells invading a fibrillar collagen substrate. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S2 (Left) Time lapse of a representative sheet of Hela cells invading a soluble collagen substrate. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S3 (Left) Time lapse of a representative sheet of Hela cells invading a plastic substrate. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S4 (Left) Time lapse of a representative sheet of lung derived endothelial cells. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S5 (Left) Time lapse of a representative sheet of MCF10-A cells invading a plastic substrate. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S6 (Left) Time lapse of a representative sheet of VEC-null mouse endothelial cells invading a plastic substrate. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S7 (Left) Time lapse of a representative sheet of VEC-positive mouse endothelial cells invading a plastic substrate. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S8 (Left) Representative simulation results. The reconstructed front is reported in red. (Right) Evolution of the corresponding local velocity map obtained by PIV.
- Video S9 A representative trajectory (yellow) of an isolated VEC-null mouse endothelial cell.