Supplementary Material for:

Size-based sorting of cancer cells reveals functional heterogeneity among subpopulations

Esra Yilmaz 1 , Zhimeng Fan 2 , Jason P. Beech 1 , Vinay Swaminathan* 2 and Jonas O. Tegenfeldt* 1

1 Department of Physics and NanoLund, Lund University
2 Department of Clinical Sciences, and The Wallenberg Centre for Molecular Medicine, Lund University

* Correspondence:

Jonas Tegenfeldt jonas.tegenfeldt@fysik.lu.se

Vinay Swaminathan vinay.swaminathan@med.lu.se

1. DLD Device Layout

The DLD device features a quadratic array of circular pillars of 50μm diameter (D) with gaps (G) between posts of 48μm.

Inlet channels are designed to have equal fluidic resistances per unit width. Outlets are also designed to have equal resistances per unit width. Balancing resistances at the inlets and outlets respectively ensures that the flow enters and exits the array in the same direction and parallel with the side walls. A schematic of the device is shown in Fig. S1.

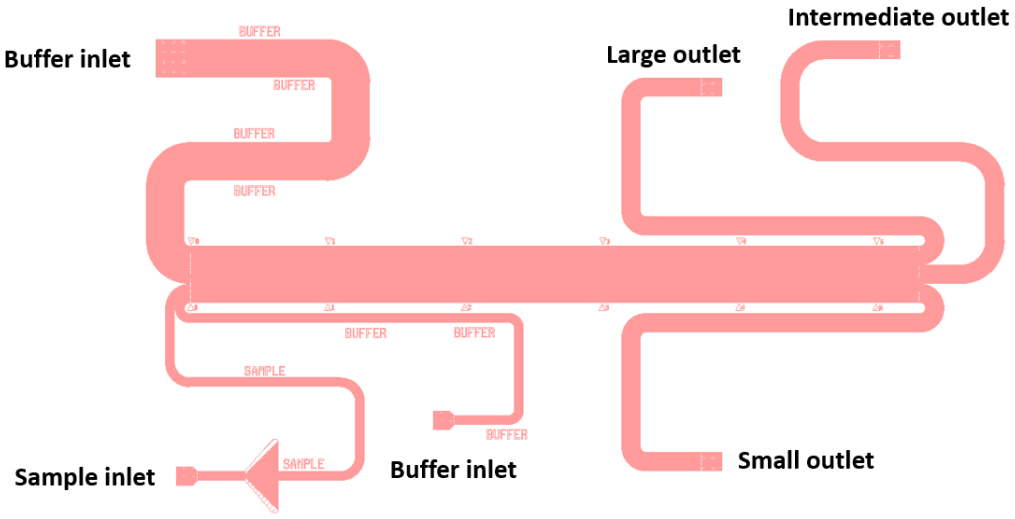

Figure S1. Schematic of the DLD device including arrays of 50μm pillars with N=20, and G=48μm. The device features three inlets, and three outlets.

The inlets consist of one sample inlet and two buffer inlets. The big buffer inlet consists of twenty eight parallel channels, small buffer inlet consists of seven parallel channels, and the sample inlet consists of seven parallel channels. With equal flow speed at the entrance of the DLD array, buffer is added to the sample corresponding to a 16,7% dilution of the initial concentration.

2. PDMS device running protocols

MDA-MB-231 breast cancer cells were suspended in a buffer composed of 10% OptiPrep and 4 mM EDTA in MACS buffer. The final cell concentration was adjusted to 4×10^5 cells/mL. Buffer solution is 10% OptiPrep and 4 mM EDTA in MACS buffer. The sample solution and buffer solutions were pushed from two inlets into the device using pressurised nitrogen gas controlled by an MFCS-4C pressure controller (Fluigent, Paris, France). The outlet reservoirs were kept at ambient pressure. During sorting, the following pressures were applied using a pressure controller. Sample inlet: 80 mbar, small buffer inlet: 80 mbar, and

large buffer inlet: 90 mbar. Sorted cells were collected from three outlets. Following collection, cells from each outlet and the inlet were transferred to Eppendorf tubes and centrifuged at 700 rpm for 3 min. The supernatants were discarded, and cell pellets were resuspended in fresh complete cell media - DMEM. Cells were then seeded onto 35 mm glass-bottom dishes featuring a central glass window and incubated for 1 h at 37 °C to allow cell attachment. To ensure representative statistics from the measurements of the inlet and the three outlets, the sampling and counting of bio particles took place carefully. Note that our device is designed with a sample to buffer ratio of 1:6 at the inlet. This means that the effective concentration of the sample in the device is 16,67% of the initial concentration in the sample inlet.

3.1. Optiprep preparation

OptiPrep™ is a ready-made, sterile and endotoxin tested solution of 60% iodixanol in water with a density of 1.320 ± 0.001 g/ml (20°C), designed for the in vitro isolation of biological particles. Different density solutions were prepared by diluting OptiPrep™ with autoMACS® Running Buffer – MACS® Separation Buffer. AutoMACS® Running Buffer (pH 7.2) contains phosphate buffered saline (PBS), bovine serum albumin (BSA), EDTA, and 0.09% azide. This solution is sterile-filtered. The properties of prepared iodixanol-MACS solutions can be seen from table 1. 40% iodixanol working solution was prepared by mixing 4 volume of OptiPrep™ and 2 volume of MACS. For example, to have 4% iodixanol (w/v) in the solution, 0.4 ml of 40% iodixanol working solution and 3.6 ml of MACS should be mixed. Different Optiprep concentrations were tested with cells then 10% Optiprep was chosen due to its density which is closer to density of cells.

Table S1: Properties of Iodixanol-MACS solutions*

RESULT		MIXING INSTRUCTIONS	
% Iodixanol (w/v)	**ρ (g/ml)**	**40% Iodixanol WS (ml)**	**MACS (ml)**
0.00	1.009	0	4.0
4.00	1.027	0.4	3.6
6.00	1.038	0.6	3.4
8.00	1.049	0.8	3.2
10.00	1.059	1.0	3.9
12.00	1.070	1.2	2.8

* WS = working solution (4 vol of Optiprep™ + 2 vol MACS), ρ = density

4. Characterization of the device

The DLD is characterized by sorting of polystyrene microspheres.

The critical size of the DLD device is defined as the size of a particle for which the probabilities to end up in the small and the large outlets are equal. We refer to these probabilities as the routing probabilities. We can formulate the problem using Bayesian statistics with the routing probability of a particle of size d_i to end up in outlet $k \in$ *{small,intermediate, large}* to be given by the conditional probability $P(k|d_i)$. Basically, we know the size distributions in each outlet $P(d_i|kd_i)$ and the fraction of the inlet that goes to each outlet $P(k)$, and using Bayes' theorem we can calculate the probability that a particle of a given size in the inlet goes to a given outlet $P(k|d_i)$ as follows.

$$P(k|d_i) = \frac{P(k)P(d_i|k)}{\sum_{j\in\{small,intermediate,large\}} P(j)P(d_i|j)}, k \in \{small, intermediate, large\} \quad (1)$$

Assuming $P(small) = P(intermediate) = P(large)$, they all cancel out and we have instead:

$$P(k|d_i) = \frac{P(d_i|k)}{\sum_{j\in\{small,intermediate,large\}} P(d_i|j)}, k \in \{small, intermediate, large\} \quad (2)$$

5. Automated Image Segmentation and Measurement Pipeline

Figure S2 illustrates the processing flow from image acquisition to final plotting of data. The script performs automated segmentation and measurement of microscopy images stored in structured subfolders. Each subfolder contains TIFF images from a large area scan of cells collected from a specific outlet and replicate. The script extracts raw object measurements in pixel units and compiles them into a single CSV file. The basic workflow is the following: 1. Convert images to grayscale and normalize intensity. 2. Enhance local contrast using CLAHE. 3. Compute morphological gradient and apply Otsu threshold. 4. Fill holes to obtain solid object masks. 5. Measure object properties: area, perimeter, bounding box, and centroid etc and calculate equivalent diameter $= 2\cdot\sqrt{(\text{Area}/\pi)}$. 6. Merge results from all subfolders into a unified CSV output. The CSV file can then be filtered to remove non-cell particles and clusters/overlapping particles using circularity ($4 \times \pi \times \text{Area}/\text{Perimeter}$) together with a minimum size threshold, see Figure S3.

In order to maximize the contrast in images of cells, the focus is set such that a bright halo is formed around each cell. This increases the intensity gradient around each cell making segmentation easier. However, it also puts the detected perimeter at the outer edge of the halo, see Figure S4. This effect can be compensated for by capturing a selection of images of the same cells, focused such that the human eye can more accurately estimate the real boundary of the cell and make manual measurements (in ImageJ). A conversion factor can

then be determined and used to improve the diameters calculated from the automated segmentation and analysis. The factor used was 0.75.

The final conversion from pixels to micrometers is done by multiplying pixel values by 0.8 (the physical pixel size on the camera is 16µm and the magnification is 20x, which gives 0.8µm/pixel). The code and exact parameters used are provided in the Supplementary Code archive.

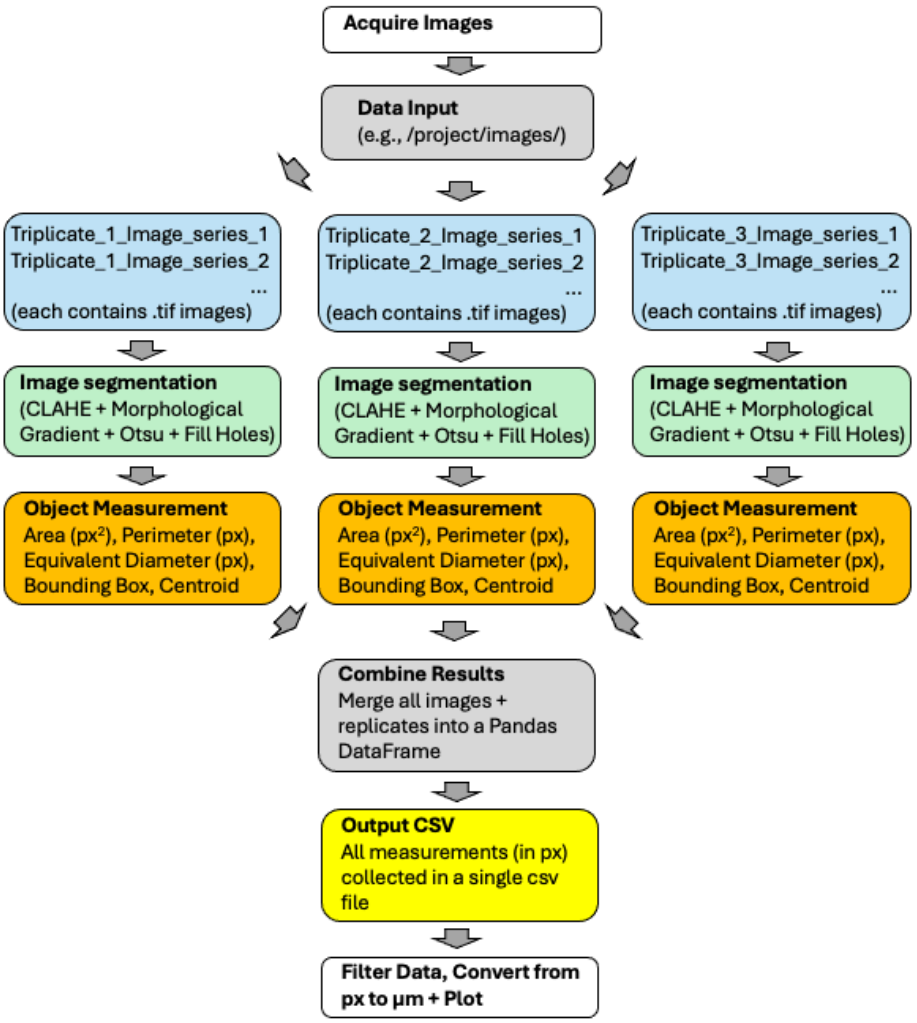

Figure S2. Workflow – Automated segmentation and cell measurement. Each subfolder (e.g., triplicate_A_1, triplicate_B_1) is processed through a standardized pipeline: grayscale conversion, contrast enhancement (CLAHE), morphological gradient, Otsu thresholding, hole filling, and object measurement. All object-level data are consolidated into a single CSV file for downstream analysis. Filtering based on size and circularity is performed to remove non-cell objects and overlapping/clustered cells. The pixel values from this analysis are then converted to micrometers for plotting and routing probability analysis.

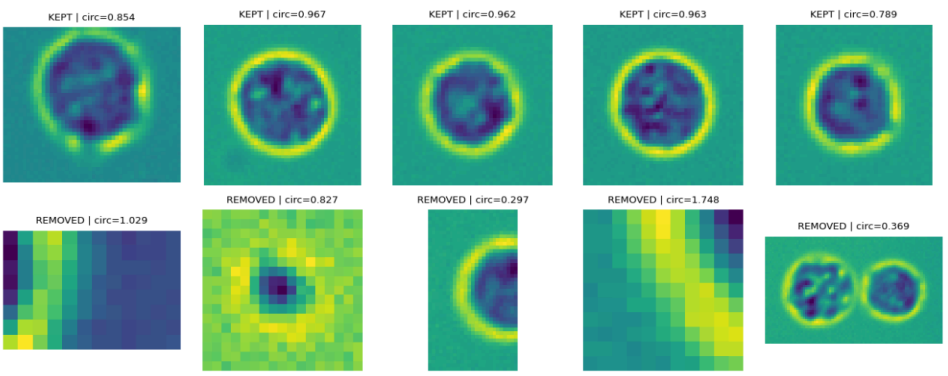

Figure S3. Examples of objects that are kept after filtering (top row) and objects that are rejected (bottom row), including partial cells, objects too small to be cells, and cell clusters.

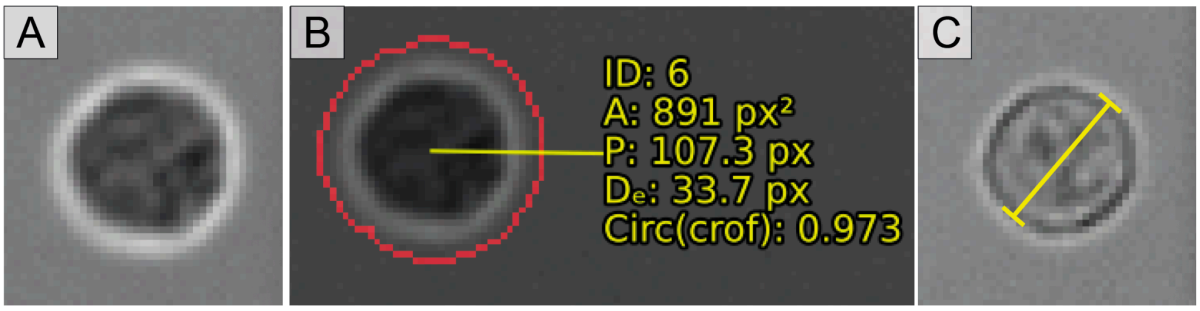

Figure S4. Optimization of image acquisition for segmentation. (A) In order to maximize contrast to improve results from image segmentation, cells are focused such that they have a bright halo. (B) Image segmentation puts the perimeter of the detected object outside of where it actually is. (C) Manual measurements of the same cell with more optimal focus for the human eye can be used to calculate a compensation factor.

6. Shape-independent cell segmentation and area quantification at Day 1 and Day 7 (related to Figure 3 in the main text)

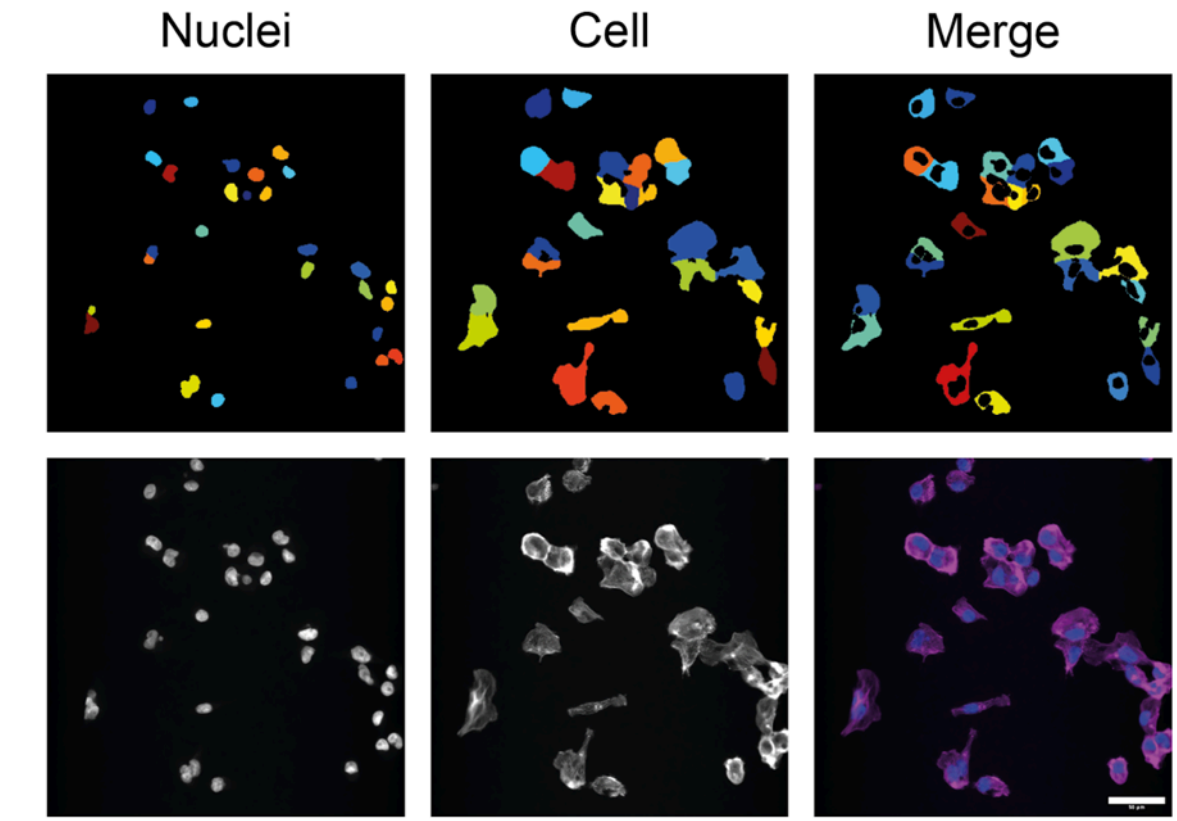

Figure S5. Shape-independent segmentation for accurate cell area quantification. Scale bar represents 50μm.

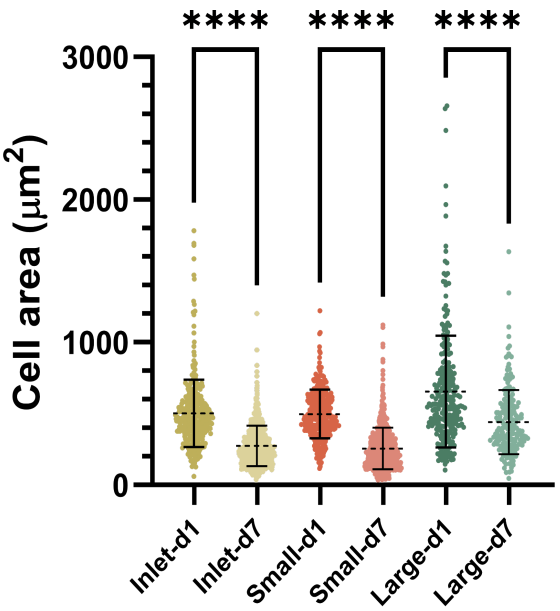

*Figure S6 Comparison of cell area in inlet, small, and large populations at Day 1 and Day 7. Data represent mean ± SD from at least three independent experiments. Statistical comparisons between Day 1 and Day 7 were performed for each population using two-way ANOVA with Tukey's post hoc test; inlet d1 vs inlet d7 (****p < 0.0001), small Day 1 and small Day 7 (****p < 0.0001), and large Day 1 and large Day 7 (****p < 0.0001).*

7. Phenotypic characterization

To validate the linearity and sensitivity of the CellTiter 96® AQueous One Solution Cell Proliferation Assay, a standard calibration curve was generated by correlating known cell numbers with the corresponding absorbance values. Serial dilutions of MDA-MB-231 cells were seeded in triplicate wells of a 96-well plate, covering a range from low to high cell densities. After incubation with the assay reagent, absorbance at 490 nm was measured using a microplate reader.

The resulting calibration curve displayed a strong linear relationship between cell number and absorbance, confirming the reliability of the assay for quantitative assessment of cell proliferation across the tested range. Data are presented as mean ± standard deviation (SD) from triplicate wells. This calibration was subsequently used to ensure accuracy in the proliferation measurements reported in the main text (Figure 3).

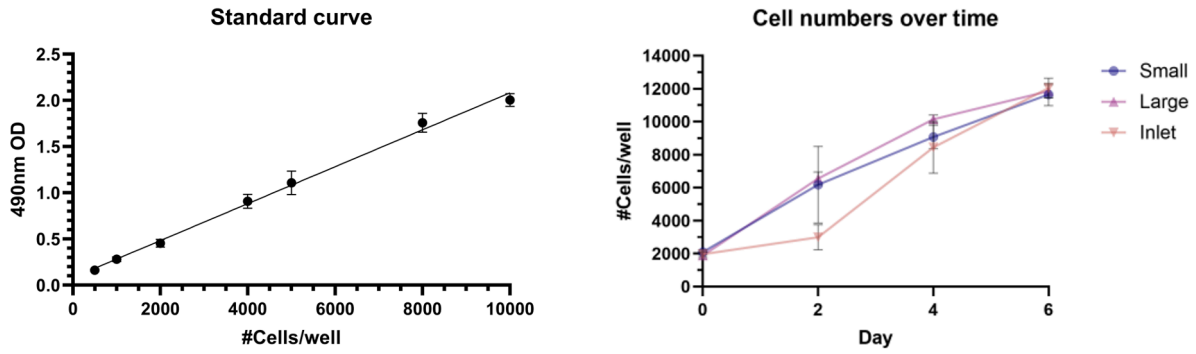

Figure S7 – Standard calibration curve for the CellTiter 96® assay. Standard calibration curve correlating cell number with absorbance values to validate assay linearity. Data are presented as mean ± SD from triplicate wells.

Migration behavior was characterized by simply tracking the cells and noting their position as a function of time. Relevant definitions are given in table S2.

Table S2. Definitions used for analysis of the migration assay

Terminology	**Explanation**
Distance	$\sum_{i=1}^{N-1} \|\|x_{i+1} - x_i\|\|$
Displacement	$\|\| x_N - x_1 \|\|$
speed (velocity)	Instantaneous speed, step length divided by time during step. $\|\|x_{i+1} - x_i\|\|/(t_{i+1} - t_i)$
x	Vector with N elements comprising a trajectory. Each element x_i in the vector corresponds to the position at t_i. **x**=$\{x_1, x_2, \ldots x_i, x_{i+1}, \ldots, x_N\}$

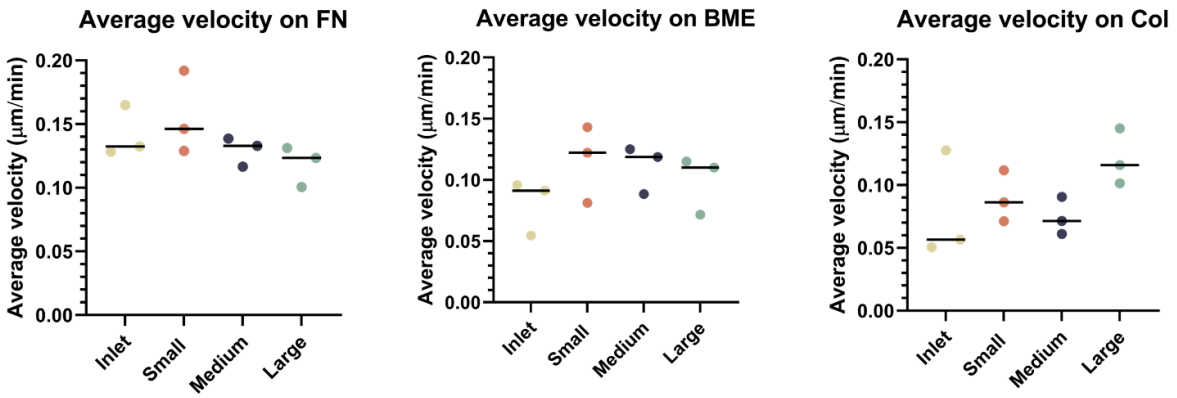

Figure S8 – Migration behavior of size-sorted breast cancer cells on distinct extracellular matrix proteins. Migration assays were performed over 20 hours on surfaces coated with Fibronectin (FN), Basement Membrane Extract (BME), and Collagen I (Col I). For each ECM condition, average velocities were quantified in small, large, and unsorted (inlet) cell populations.

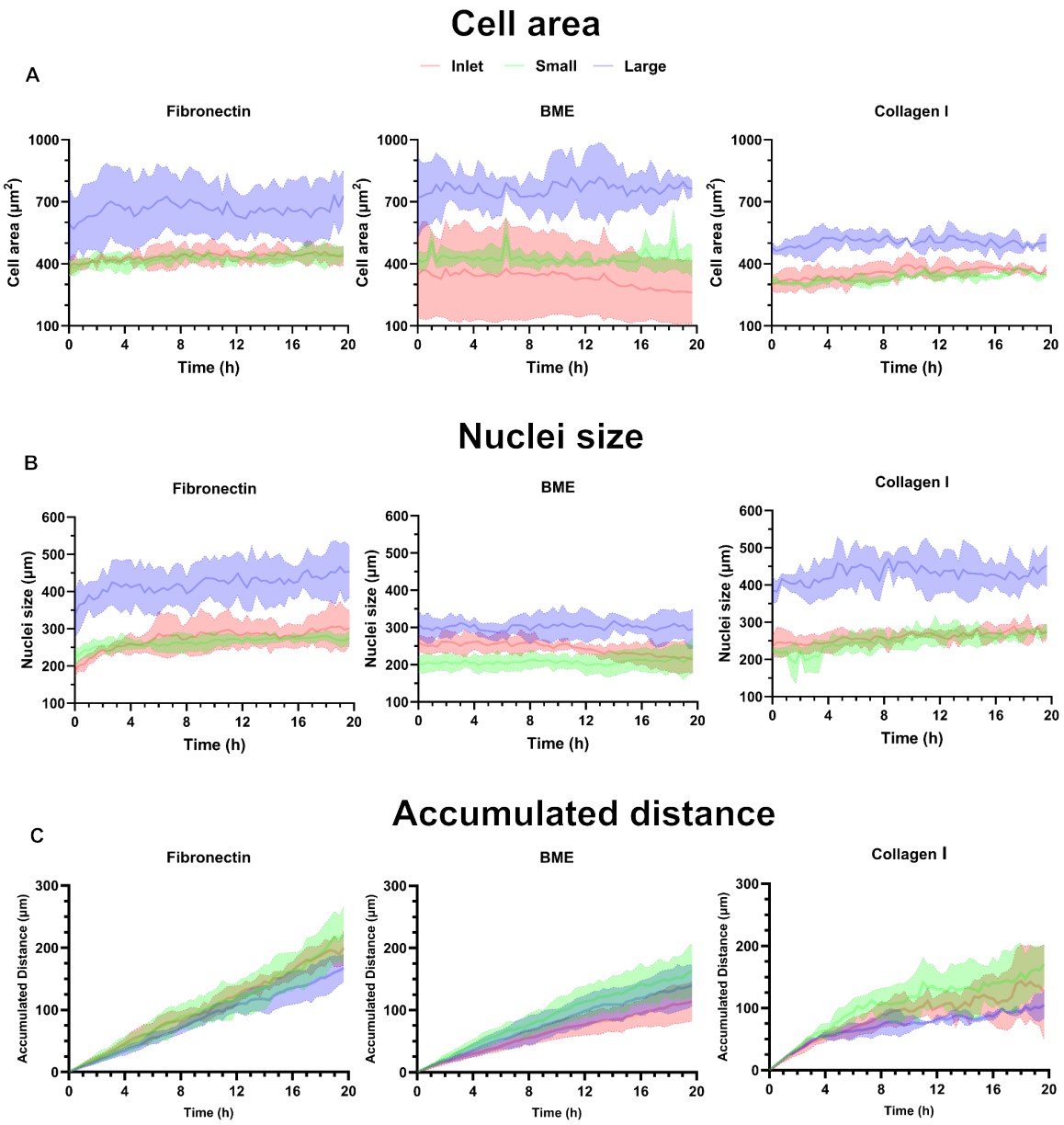

Figure S9 – Migration behavior and morphological characteristics of size-sorted breast cancer cells on extracellular matrix proteins. Migration assays were performed over 20 hours on surfaces coated with Fibronectin (FN), Basement Membrane Extract (BME), and Collagen I (Col I). For each ECM condition, A. Cell area, B. Nuclei size and C. Accumulated distance were quantified in small, large, and unsorted (inlet) cell populations. In all graphs (A–C), the continuous line represents the mean value for the respective parameter (cell area, nuclei size, or accumulated migration distance).

Representative videos

https://doi.org/10.7910/DVN/T1VNMP

Movies are available as follows:

Video 1. MDA-MB-231 cells entering the DLD device – Recording showing unsorted MDA-MB-231 cells entering the DLD array and distributing across the array

Video 2. Sorted MDA-MB-231 cells leaving the DLD array – Recording showing sorted MDA-MB-231 cells exiting the DLD array from three different outlets.

Video 3. Sorted MDA-MB-231 cells in outlet reservoirs – Recording showing the accumulation of sorted cells in the outlet reservoirs of the device.