Electronic Supplementary Material (ESI) for Lab on a Chip.

Supplementary Information

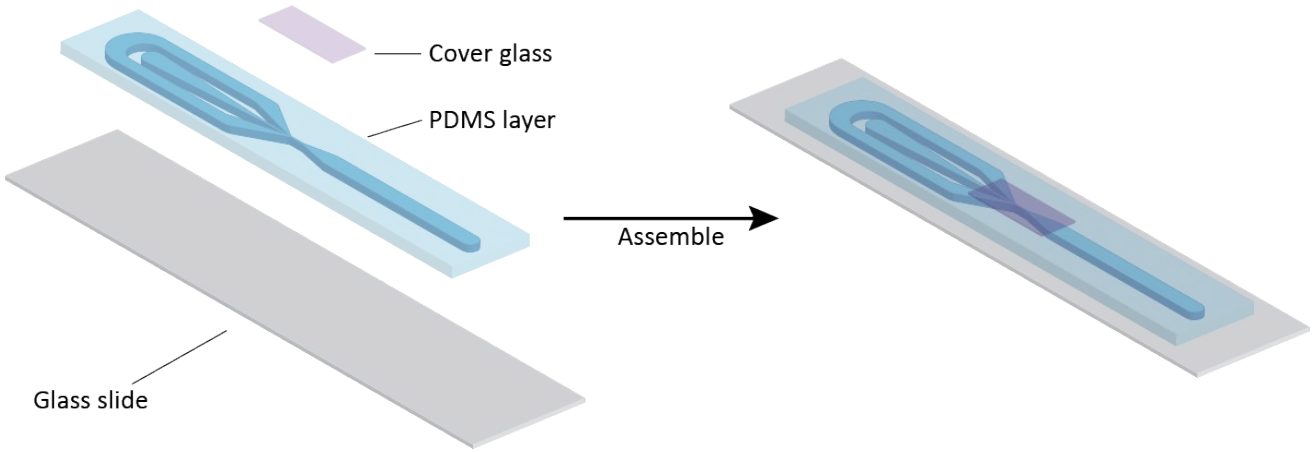

Figure S1. Structure of the microfluidic device. A cover glass is attached on top of the PDMS layer to prevent swelling of the microchannels due to high hydrodynamic pressure.

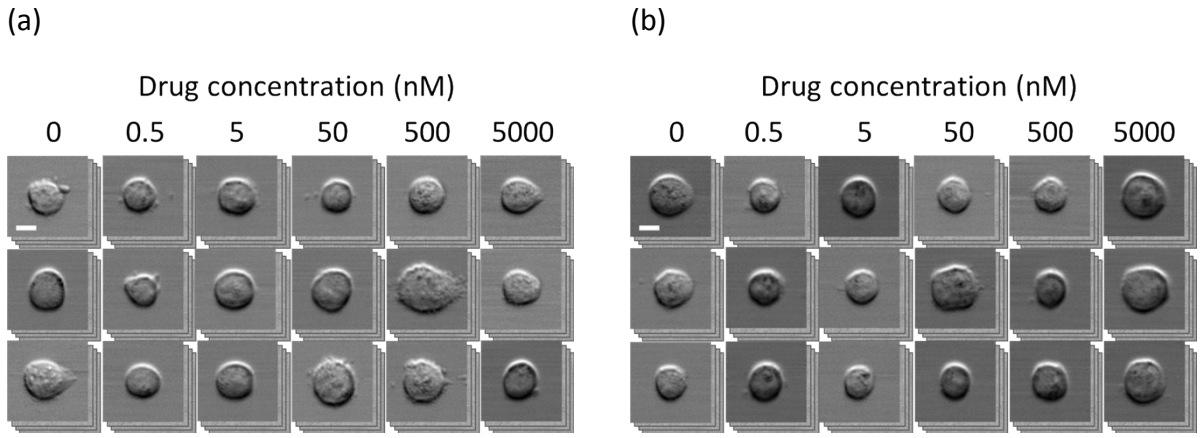

Figure S2. Image libraries of drug-treated K562 and K562/ADM cells. Images of drug-treated (a) K562 and (b) K562/ADM. The images were acquired at a flow speed of 10 m/s. Scale bar: 10 μm.

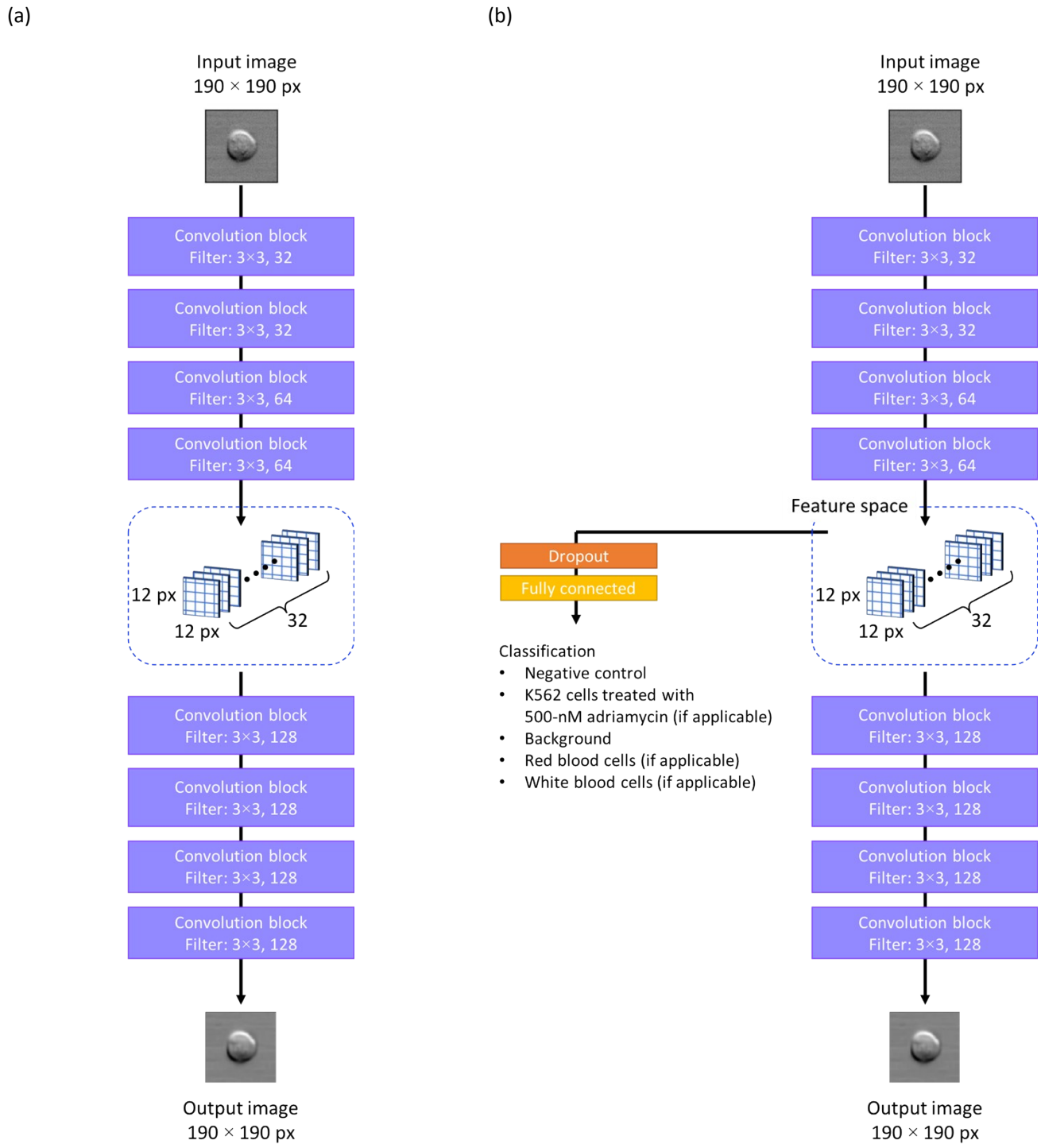

Figure S3. Feature extraction from cell images. (a) Morphological features in a high-dimensional space (2048 dimensions) extracted by a deep convolutional autoencoder. (b) Branch after the encoder is added for a classification task to regularize feature extraction to a specific direction. All convolution blocks and fully connected layers contain convolution layers followed by batch normalization and ReLU activation layers.

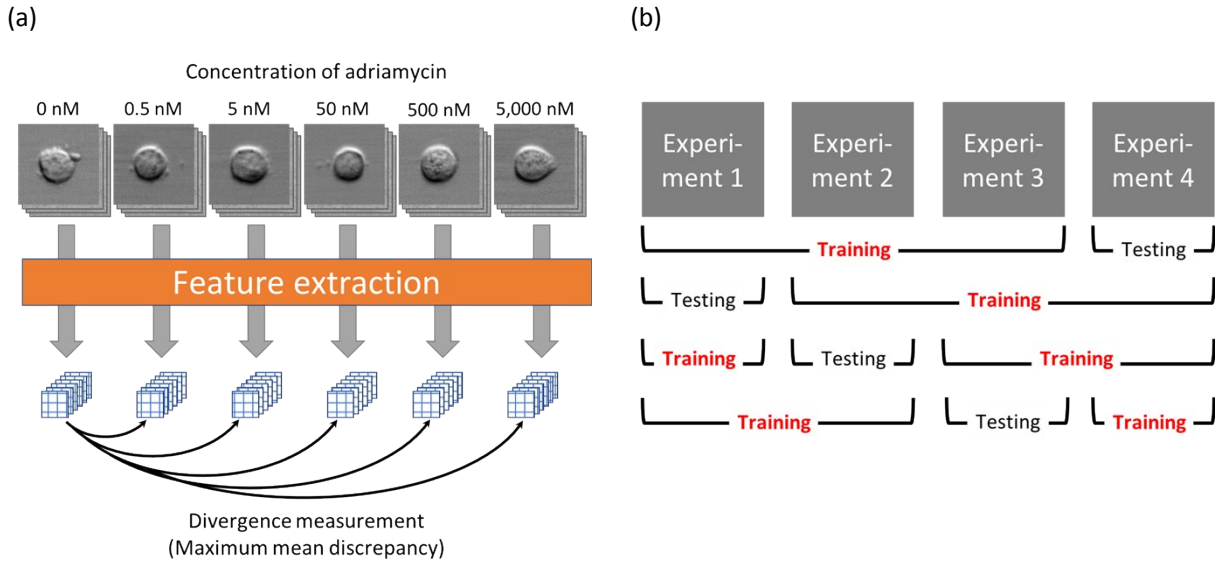

Figure S4. Schematic of the feature analysis and round-robin training and testing. (a) Drug-induced morphological changes evaluated by calculating the MMD between the control population and cells treated with various concentrations of adriamycin. (b) Round-robin training and testing. Among four sets of experimental results, three are used on training and one is used on testing.

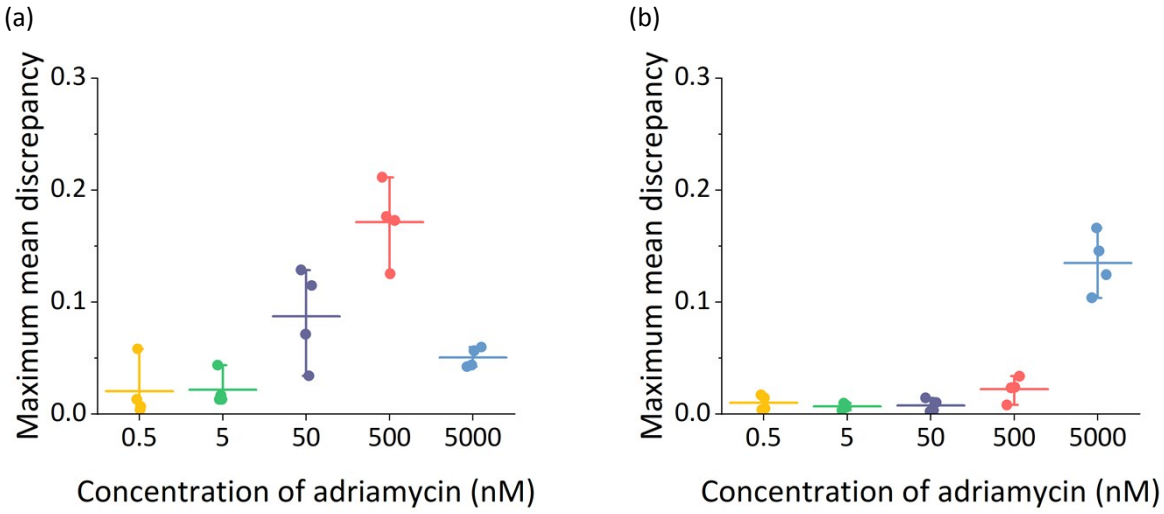

Figure S5. Dose dependence of morphological changes detected by semi-supervised deep convolutional autoencoders. (a) MMD values of K562 cells treated with adriamycin obtained from three-class semi-supervised autoencoder models trained on three trials of K562 cells (Figure S4b). (b) MMD values of K562/ADM cells obtained from a three-class semi-supervised autoencoder model trained on all trials of K562 cells. The error bars represent standard errors of the mean values (n = 4).

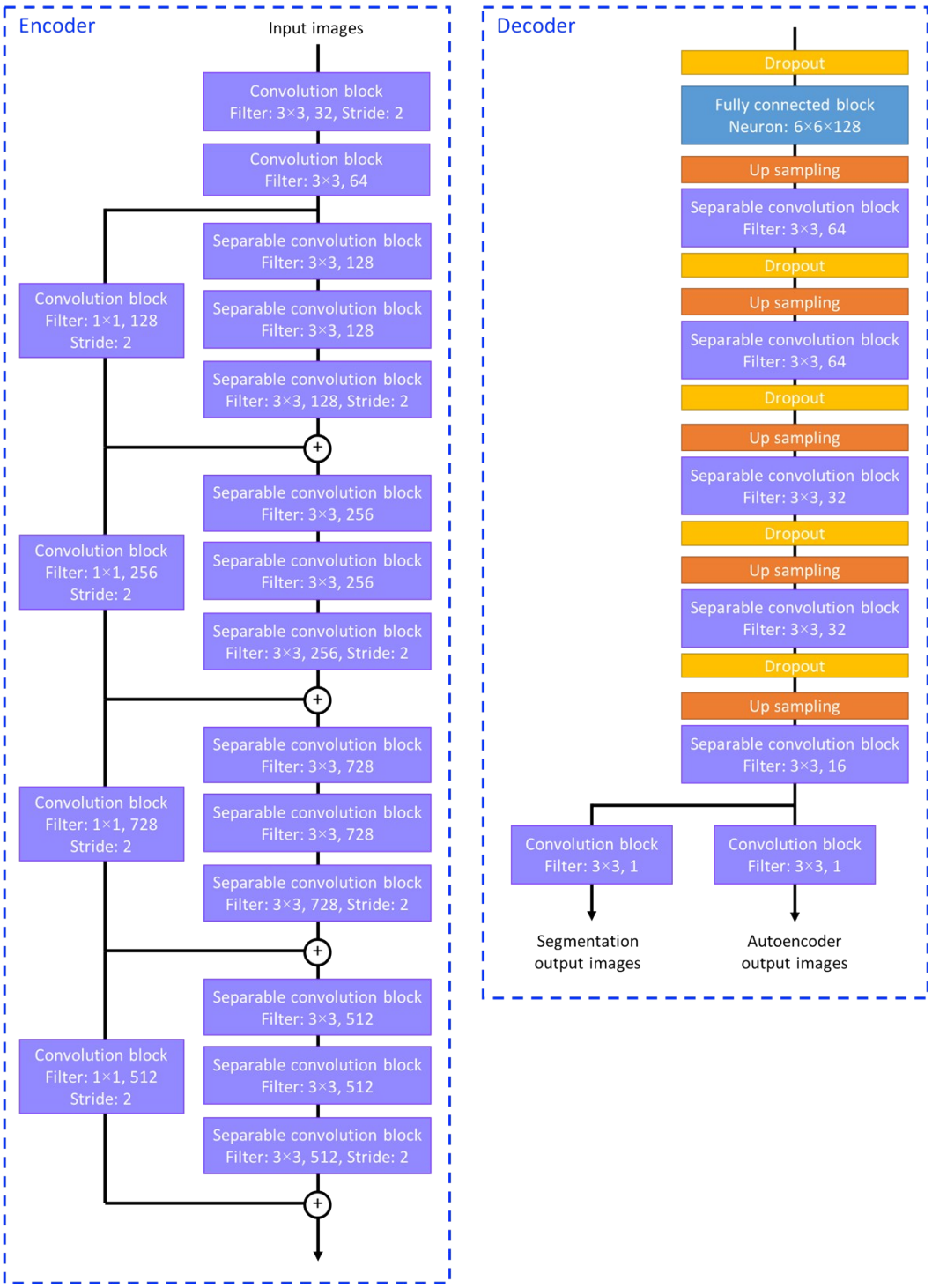

Figure S6. Structure of the deep convolutional autoencoder for segmentation. All convolution blocks and separable convolutional blocks contain convolution layers followed by batch normalization and ReLU activation layers.

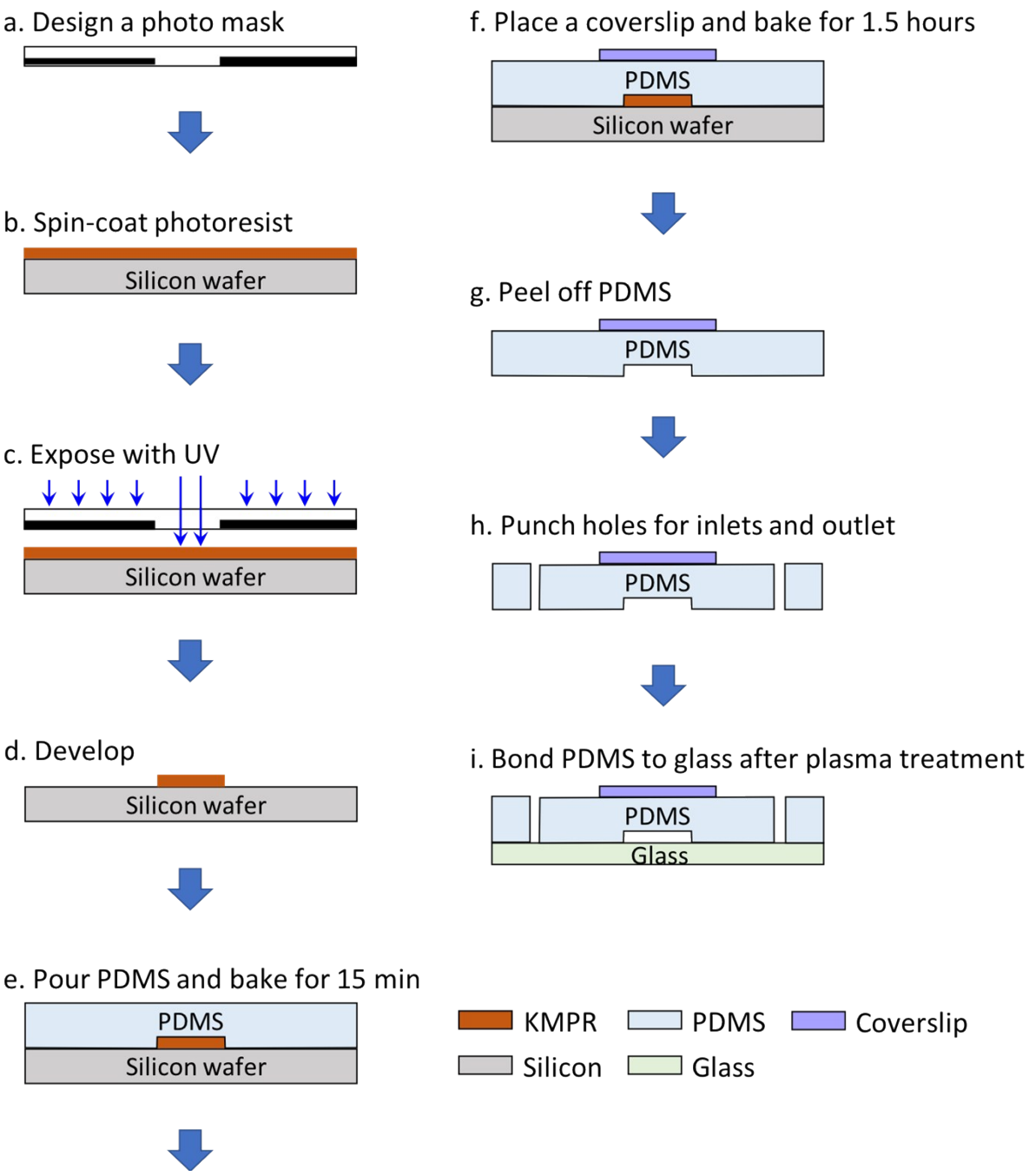

Figure S7. Steps for fabricating the microfluidic device. (a) Print the designed microfluidic device on a film mask. (b) Spin-coat a negative photoresist (KMPR 1035) on a silicon wafer. (c) Expose the photoresist with UV over the photo mask so that the microfluidic device design is transferred onto the photoresist. (d) Develop the photoresist. (e) Pour PDMS (Polydimethylsiloxane) on the silicon wafer, and bake for 15 minutes to partially solidify the PDMS. (f) Place a cover slip on the top of the PDMS, and bake for 1.5 hours. (g) Peel off the PDMS. (h) Punch holes for inlets and an outlet. (i) Bond PDMS to a glass slide after plasma treatment.

Supplementary Note

The MMD is an integral probability metric (IPM) that distinguishes between two probability measures p and q:

$$MMD(p,q) = \sup_{f \in H} E_{x\sim p}[f(x)] - E_{y\sim q}[f(y)]f \quad , \tag{1}$$

where H is a unit ball in reproducing kernel Hilbert space (RHKS) with kernel k. If k is characteristic, then $\mathrm{MMD}(p,q) = 0$ if and only if $p = q$. Given two sets of samples X and Y, the empirical estimate of squared MMD can be computed in quadratic time as

$$MMD^2(X,Y) = \frac{1}{m(m-1)}\sum_{i \neq j}^{m} k(x_i,x_j) + \frac{1}{n(n-1)}\sum_{i \neq j}^{n} k(y_i,y_j) - \frac{1}{mn}\sum_{i}^{m}\sum_{j}^{n} k(x_i,y_j). \tag{2}$$

In all the experiments, we used the Gaussian kernel $k(x,y) = \exp\left(-\frac{\|x-y\|_2^2}{2\sigma^2}\right)$ and selected the Gaussian width $\sigma > 0$ via the median heuristic.

The HSIC measures the dependency between two random variables. Given two measures p and q with joint measure $p \otimes q$, and two RKHSs F, G with kernels $k(\cdot,\cdot)$, $l(\cdot,\cdot)$, respectively, the HSIC is defined as the squared HS-norm of the cross-covariance operator:

$$HSIC(p,q,F,G) = \|C_{xy}\|_{HS}^2 = E_{(x,y),(x',y')\sim p\otimes q}\left[k(x,x')l(y,y')\right] + E_{x,x'\sim p}[k(\quad . \tag{3}$$

Note that HSIC can also be interpreted as the MMD between the joint measure and the product of marginals. Therefore, with the use of characteristic kernels, the HSIC is non-negative and takes zero if and only if the two variables are independent. The normalized empirical estimate of the HSIC is given by

$$HSIC_b(X,Y) = \frac{tr(KHLH)}{\|HKH\|_F \|HLH\|_F}, \tag{4}$$

where $H = I - \frac{1}{n}11^\top$, $K \in R^{n\times n}$ is the Gram matrix of X with $K_{ij} = k(x_i,x_j)$, and $L \in R^{n\times n}$ with $L_{ij} = l(y_i,y_j)$. In all the experiments, *k* and *l* are Gaussian kernels.