PNAS

Supporting Information for

Robotic data acquisition with deep learning enables cell image-based prediction of transcriptomic phenotypes

Jianshi Jin, Taisaku Ogawa, Nozomi Hojo, Kirill Kryukov, Kenji Shimizu, Tomokatsu Ikawa, Tadashi Imanishi, Taku Okazaki, Katsuyuki Shiroguchi*

*Correspondence: Katsuyuki Shiroguchi
Email: katsuyuki.shiroguchi@riken.jp

This PDF file includes:

Other supporting materials for this manuscript include the following:

Table of Contents

Supporting Information Texts

(1) Preparation of cells, organoids, and tissue samples.

Cell lines, mouse T cells (T cells), mouse leukemia cells (leukemia cells), mouse hematopoietic progenitor cells (HPCs), and human T cells (Jurkat cells) were prepared. Mouse T cells were established from BW-1100.129.237 cells (provided by Leszek Ignatowicz, Georgia State University, USA) by retrovirally introducing the genes CD3δ, CD3ζ, CD8α, CD8β, CD28, and 2Cm6α TCR (1). Mouse leukemia cells were established from the C1498 cell line (ATCC) by retrovirally introducing the CD86 gene and a fusion gene consisting of H2-L^d, β2M, a sequence encoding the QL9 peptide, and EGFP. Cell culture, plasmid reconstruction, retroviral gene transfection, and positive cell selection were performed as previously described (2). Mouse hematopoietic progenitor cells (HPCs) were prepared as previously described (3). Jurkat cells were obtained from RIKEN Cell Bank and maintained in Medium A (see Materials and Methods).

Mouse (C57BL6/J) peripheral blood mononuclear cells (PBMCs) were prepared from whole blood as follows. Approximately 500 μl of murine whole blood was collected in a tube containing 30 μl of liquid heparin (1000 units/ml, MOCHIDA PHARMACEUTICAL CO., LTD.) by a syringe washed with heparin via cardiac puncture under anesthesia, then centrifuged using Histopaque-1083 (Sigma-Aldrich) following the manufacturer's protocol. The sedimented cells were incubated with ACK Lysing Buffer (Gibco) to lyse the remaining red blood cells. After washing twice with phosphate-buffered saline (PBS, Nacalai Tesque), the cells were stained with 1 μg/ml 7-aminoactinomycin D (7-AAD) (eBioscience). Viable PBMCs were collected in medium (1× RPMI 1640 medium (Gibco) supplemented with final concentrations of 10% (v/v) fetal bovine serum (Corning), 1× Monothioglycerol Solution (FUJIFILM Wako Pure Chemical Industries), 1× GlutaMAX™ supplement (Gibco), 1× Penicillin/Streptomycin Mixed Solution (Nacalai Tesque Inc.), 5 mM HEPES (Gibco), 1× MEM nonessential AA (Gibco), and 1 mM sodium pyruvate (Gibco)) by cell sorting using a BD FACS Aria III (BD Biosciences) based on the forward and side scatter properties and 7-AAD-negative signals. The collected PBMCs were kept at room temperature until the subsequent experiments.

Mouse small intestinal crypts were isolated and cultured as described with slight modification (4). Briefly, the small intestine was obtained from a C57BL/6 mouse by surgery. After cutting the intestine into 2 mm long pieces and washing using ice-cold PBS, the intestinal pieces were incubated in PBS with 2.5 mM EDTA (Gibco) for 30 min at 4 °C with gentle agitation. Then, intestinal crypts were released from the intestinal pieces by vigorous suspension in ice-cold PBS. The crypts in the suspension were enriched by centrifugation at 300 ×g for 3 min at room temperature after passing through a 70-μm cell strainer (Corning). The collected crypts were incubated in 1× DMEM/F-12 (Nacalai Tesque) with 10% (v/v) fetal bovine serum (Corning) at 4 °C until further experiments.

Mouse small intestinal organoids were grown from the crypts as described with slight modifications (5). The prepared crypts were collected from the medium by centrifugation at 400 ×g for 3 min at room temperature, embedded in growth factor-reduced Matrigel (Corning), and seeded on a 24-well plate. After polymerization, culture medium (1× advanced DMEM/F-12 (Gibco) supplemented with final concentrations of 1× Penicillin/Streptomycin Mixed Solution (Nacalai Tesque Inc.), 250 ng/ml amphotericin B (HyClone), 1× GlutaMAX™ supplement (Gibco), 1 mM N-acetyl-L-cysteine (Nacalai Tesque Inc.), 10 mM HEPES (Gibco), 1× N2 supplement (FUJIFILM Wako Pure Chemical Industries), 1× B-27 supplement (Gibco), 10 μM CHIR99021 (Sigma), and 200 nM LDN193189 (Selleck)) was added and refreshed every 2-3 days. On days 6, 8, and 12,

Matrigel-embedded organoids were incubated with ice-cold cell recovery solution (Corning) for 30 min on ice to release the organoids from the Matrigel. Then, the organoids were transferred to a tube and sedimented by gravity. The supernatant was discarded, and the organoids were resuspended in DMEM/F-12 with 10% fetal bovine serum and kept at room temperature until further experiments.

(2) Isolation of beads, leukemia cells, Jurkat cells, organoids, and crypts using ALPS

Polystyrene beads 1 μm (Sigma-Aldrich), 3 μm (Polyscience), 45 μm (Polyscience), and 200~300 μm (Polyscience) in diameter were separately diluted in Milli-Q water with 0.1% Tween 20 and mixed by pipetting. Then, 3 ml of each diluted bead mixture was added to a 35-mm glass bottom dish (D35-27-1.5-, Matsunami), and the beads were sedimented by gravity for approximately 1 h, 15 min, 15 min, and 15 min, respectively, before isolation. The final densities of the beads on the bottom of the dish were approximately 2, 50, 50, and 10 per mm^2, respectively.

For leukemia cells, Jurkat cells, organoids and crypts, 3 ml of each prepared sample was added into a 35 mm glass bottom dish and sedimented for approximately 15 min before isolation.

First, a randomly selected sample was moved to the center of the field of view by the microscopic motorized stage. Second, the sample was picked at the center of the field of view using a 40-μm needle with 9 nl of solution for a 1-μm bead, a 40-μm needle with 9 nl of solution for a 3-μm bead, a 70-μm needle with 9 nl of solution for a 45-μm bead, a 400-μm needle with 340 nl of solution for a 200~300-μm bead, a 60-μm needle with 11 nl of solution for a leukemia cell or Jurkat cell, a 400-μm needle with 110 nl of solution for an organoid, and a 40-μm needle with 11 nl of solution for a crypt. Third, the needle with the bead, cell, organoid, or crypt was moved to a well of a 96-well glass bottom plate (Greiner Bio-One) that contained 40 μl of Milli-Q water with 0.1% Tween 20 for beads, organoids, and crypts or 40 μl of conditioned medium for leukemia cells and Jurkat cells (the conditioned medium was basically Medium A filtered using a 0.22 μm Steliflip filter unit after culturing 10^5 (initial) leukemia or Jurkat cells for two days). Finally, 15 nl of solution was deposited into the well for 1-μm, 3-μm, and 45-μm beads, 340 nl for 200~300-μm beads, 19 nl for leukemia cells and Jurkat cells, 110 nl for organoids, and 19 nl for crypts. All the isolation processes were recorded by the two cameras and the screen capture software (see section "Hardware and software of ALPS" in Materials and Methods).

(3) Growth of cells isolated by ALPS

After a leukemia or Jurkat cell was isolated into a well of a 96-well glass bottom plate, 60 μl of conditioned medium was added. After the plate was maintained at 37 °C and 5% CO_2 using an incubator (MCO-170AIC-PJ, PHCbi) for three days, each well was imaged using a Nikon microscope by bright-field illumination (30% power), 10 ms exposure time, and a 20× objective (see section "Hardware and software of ALPS" in Materials and Methods). For the negative control, the same procedures were performed except for cell deposition for wells in the same plate.

(4) Single-cell RNA sequencing

An automated library preparation system for RNA-seq was set up using the Bravo NGS workstation (Agilent Technologies): multiple steps, including cell lysis, reverse transcription, template switching, and bead purification, were programmed to perform scRNA-seq (SI Appendix Fig. S21). Two types of barcodes were introduced to

identify individual cells in multiplex sequencing: 96 well-position barcodes and 96 plate barcodes distinguished a total of 9,216 (=96 × 96) cells (Dataset S13). It took less than 2 days for one technician to perform library preparation for four 96-well plates using four thermal cyclers, including quality check and concentration measurements of the generated libraries.

Bravo NGS workstation (Agilent Technologies) was programmed as follows. (1) Frozen picked or sorted cells in 96-well PCR plates were thawed and kept at 4 °C on Bravo's deck cooler. (2) Then, 0.8 µl of a lysis and fragmentation solution (0.4 U RNAsin plus RNase inhibitor (Promega), final concentrations of 2.5 × Single-cell lysis buffer (Takara), 1 ×Maxima H Minus Reverse Transcriptase buffer, 2.5 nM polyT14V_A, 2.5 nM polyT14V_T, 2.5 nM polyT14V_G, 2.5 nM polyT14V_C (Integrated DNA Technologies, Inc.) (Dataset S13), and 1,280,000-fold diluted ERCC RNA Spike-In Mix (Thermo Fisher Scientific)) was added to each well of the plates. (3) After the plate was sealed and centrifuged for 30 s at 4,000 rpm using Plate Centrifuge PlateSpin II (KUBOTA), cell lysis and RNA fragmentation were performed at 94 °C for 3 min using thermal cyclers (Mastercycler X50s, Eppendorf). (4) Then, 1.2 µl of a reverse transcription solution (8 U Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific), 0.4 U RNAsin plus RNase inhibitor, final concentrations of 3.3 mM dNTPs (NEB), 17 mM DTT (Sigma-Aldrich), 6.7 µM TSO (Integrated DNA Technologies, Inc.) (Dataset S13), and 25% (w/v) PEG 8000 (Promega)) was added to each well of the plates. (5) After the plates were sealed, centrifuged for 30 s at 4,000 rpm using Plate Centrifuge PlateSpin II, vortexed for 15 s at 3,000 rpm using MixPlate, and centrifuged again for 30 s at 4,000 rpm at room temperature, reverse transcription was performed at 42 °C for 90 min and inactivated at 85 °C for 10 min. (6) The unreacted primers were digested with 1 U of Exonuclease I (NEB) at 37 °C for 30 min followed by inactivation at 80 °C for 20 min. (7) After adding 15 µl of an amplification solution (10 µl of 2 × SeqAmp CB PCR buffer (Takara), 0.25 U of SeqAmp DNA polymerase (Takara), and final concentrations of 0.4 µM well-barcode primer and 0.4 µM plate-barcode primer (Integrated DNA Technologies, Inc.) (Dataset S13)) to each well of the plates, amplification was performed: 1 min at 94 °C; 21 cycles of 10 s at 94 °C, 15 s at 60 °C and 2 min at 68 °C; and 5 min at 68 °C. (8) Then, 10 µl of the amplified PCR products of each of 12 wells (i.e., each row) of the 96-well PCR plate were pooled into one well of a 96-deepwell plate (Eppendorf). We note that for eight plates of the three-type-mixed cell experiments and all plates of the PBMC experiments, the amplified products of the wells were discarded before pooling when any cell was not picked, more than one cell was deposited, or the linking with time-lapse imaging failed (see image analysis section). (9) The pooled amplified products were purified by three rounds of AMPure XP Bead purification (Beckman Coulter). For the first, 240 µl of bead buffer (final concentrations of 8% (w/v) PEG 8000 and 1.4 M NaCl) and 120 µl of AMPure Beads were added, and 60 µl of elution buffer (Qiagen) with 0.1% (v/v) Tween 20 was used for elution; for the second, 120 µl of bead buffer, 60 µl of AMPure beads, and 60 µl of elution buffer was used; and for the third, 120 µl of bead buffer, 60 µl of AMPure beads, and 15 µl of elution buffer were used.

The quality of the libraries was checked by an Agilent 2100 Bioanalyzer (Agilent Technologies), and the concentration was determined by qPCR (KAPA SYBR Fast qPCR kit, KAPA Biosystems) using primers P1_qPCR_Fw and P2_qPCR_Rv (Dataset S13). For the libraries that had a high proportion of products between 100~200 bp, two more rounds of AMPure bead purification were performed. The twelve purified libraries from the same plate were mixed in equal volumes. The libraries of different plates (8 or 10 plates for each sequencing run) were mixed by equal volume or by equal concentration, and 20~25% Phix (Illumina) was added. The libraries of samples were paired-end sequenced on an Illumina MiSeq (two runs for PBMC) or HiSeq (three runs for three-

type-mixed cells) platform by allocating 102 cycles for Read 1, 8 cycles for Index 1, 8 cycles for Index 2, and 56 cycles for Read 2 (SI Appendix Fig. S21). The Illumina Index 1 and Read 2 sequencing primers were replaced by custom primers named I1_primer and R2_primer (Dataset S13).

(5) Sequence analysis

The sequencing results were processed to count the number of RNA molecules for detected genes in each index (unique pair of a well-position barcode and plate barcode), which basically corresponded to each single cell. For HiSeq runs, the output bcl files were converted to fastq format and demultiplexed into different fastq files based on indexes using bcl2fastq (Illumina). For the fastq files from each sequencing run of HiSeq or MiSeq, the following steps were performed. First, a read pair (Read 1 and Read 2 in pairs) that had correct fixed bases in the molecular barcode (24 bp from the 5' end of Read 1; SI Appendix Fig. S21 and Dataset S13) was selected for the next step (6). Second, the first 38 bases from the 5' end of Read 1 that corresponded to the designed molecular barcode and poly T primer were deleted, and the poly T sequences in the remaining reads were replaced by N using tantan 23 (7) with the parameters *"*-x N *-r 0.0005"*. Third, the processed Read 1 and Read 2 were mapped to both the mouse genome (mm10) and the sequences of the ERCC Spike-In RNA using STAR 2.7.3a (8) with the parameter *"*--outFilterMultimapNmax 50*",* which allowed mapping to multiple genes. The mapping results outputted from STAR were reformatted using samtools 1.10 (9). The mapped Read 1 and Read 2 were annotated with gene name(s) using the GTF file from the UCSC genome browser (refFlat, downloaded using "Table Browser" on Jan. 21st, 2018). We note that for the sequencing results of PBMCs, 270 T cell receptor-related genes registered in the GTF file (version M25) from Gecode were also used for annotation. If Read 1 and Read 2 for the same read pair were mapped to the same gene(s), the read was annotated to the gene(s); if not, the read was annotated to all mapped genes of both Read 1 and Read 2 (the annotated gene(s) were defined as a gene set for the next steps). Fourth, the annotated reads were divided into groups, each of which had different combinations of gene sets and indexes. Within each group, the molecular barcode sequences (first 24 bp of Read 1) were clustered using Nucleotide-Sequence-Clusterizer with the parameter "--distance 2" (6). Fifth, for each cluster, the sequence at each position in the barcode with the highest number of reads was used as a representative sequence for the cluster (when multiple sequences had the highest number of reads, one of them was used). The clusters that had the same representative sequences among all groups from different gene sets or indexes were merged. In a merged cluster, a representative index with the highest number of reads was used as an index for this cluster. When multiple indexes had the highest number of reads in a merged cluster, the cluster was annotated by all these indexes. In a merged cluster, a representative gene set that had the highest number of reads was used as a gene set for this cluster. When multiple gene sets had the highest number of reads in a merged cluster, the cluster was annotated by all these gene sets. We note that each (merged) cluster corresponded to one detected RNA molecule. Sixth, the number of clusters for each gene and index was counted as follows: a coefficient of 1 was given to each cluster; when a cluster was annotated by multiple gene sets and/or indexes, a coefficient of 1/(the number of different combinations of gene sets and indexes) was given for each combination of the gene sets and indexes; then, for each combination of a gene set and an index, a coefficient of 1/(the number of different combinations of gene sets and indexes)/(the number of genes in the gene set) was given for each gene; the total number of clusters (RNA molecules) for each combination of a gene and an index was determined as the sum of the coefficients given to the combination of the gene and index. Finally, the number of reads for each combination of a gene and an index

in each cluster (RNA molecule) was determined as the total number of reads in the cluster multiplied by the given coefficient of the combination of the gene and index. The average detection efficiency of ERCC Spike-In RNA was 6.7% for three-type-mixed cells and 6.6% for PBMCs based on the nominal concentration of ERCC Spike-In RNA. We confirmed that the sequencing depths in all runs and the number of random bases in our designed barcode were sufficient (SI Appendix Figs. S22 and S23).

(6) Gene expression analysis

Single cells were clustered based on RNA expression. The indexes (i.e., wells) with failed time-lapse tracking (see next section) and/or isolation by ALPS (Datasets S1 and S2) were filtered out. The remaining indexes (corresponding to single cells) whose detected number of ERCC RNA molecules was < 1,000 or detected number of molecules of murine genes was < 10,000 for three-type-mixed cells (< 1,000 for PBMCs) were removed from further analyses. For the three-type-mixed cells, the indexes whose detected number of genes was < 2,500 were removed from further analyses. The gene expression matrices of the remaining cells were normalized by the NormalizeData function using the parameters "*normalization.method = LogNormalize"* and "*scale.factor = 10,000"*. For dimensional reduction, 2,000 genes with high cell-to-cell variation were found using the FindVariableFeatures function. The dimensional reduction was performed by the RunPCA function with default parameters on the scaled data which were linear-transformed by the ScaleData function. The cells were clustered into different numbers of clusters using the FindNeighbors and FindClusters functions by adjusting the parameter "*resolution*". For visualization, nonlinear dimensional reduction by the RunTSNE function was conducted. All functions were used with the instructions from Seurat 3.2.2 (10).

(7) Image analysis

Determination of cell types by fluorescent images

To determine the cell types of each single cell based on the fluorescent images obtained by the ALPS using imaging method 3 in the assays of single isolated cells using multimode imaging (Materials and Methods), the mean intensities of GFP (marker for leukemia cells), PE (marker for T cells) and Alx648 (marker for HPCs) of the cells were determined. The mean intensity was calculated from all pixels in a circular area (radius, 40 pixels (12.8 μm)) that covered the given cell. Then, the calculated mean intensity was subtracted by the background, which was the average intensity of the same image excluding the circular area. The subtracted mean intensities of all cells were divided by the highest mean intensity among these cells for each color respectively. Finally, the cell type of each cell was determined to be the one that had the highest intensity among the three colors. When the values of all colors were less than 0.04, meaning that the cell was not efficiently labeled by any of the three markers, the cell type was defined as "undetermined" (Fig. 2A and Dataset S3).

Linking time-lapse image to each isolated single cell

The cells in the images obtained by 29 min time-lapse imaging (1 min interval, 30 time points) and nine scans during cell isolation at focal position 2 (Materials and Methods) were tracked and linked to the picked cells by the following steps. First, after normalizing the background using the "Shading correction" job (NIS-Elements), the images of 6 × 6 squares for each time point were stitched, and all 39 stitched frames were stacked based on the time series of acquisition using the "stitch Multipoint to Large Image" job (NIS-Elements). Second, the

intensity of each pixel of the first 31 frames was subtracted by a constant value of 500 because frames 32-39 had lower background intensity due to the effect of the needle position. Third, the cells in each image were segmented by the "Define Threshold" job (NIS-Elements) using the following parameters: "Intensity = 0~1600, Fill holes = On, Separate = 1 ×, and Size = 5~25 μm for cell lines or 5~15 μm for PBMCs". Fourth, the cells were tracked along the time series by the "Tracking" job (NIS-Elements) using the following parameters: "Radom motion = On, Constant speed = Off, Allow Gaps in tracks = Off, Allow New Tracks After First Frame = Off, Standard Deviation Multiplication Factor (Stdev) = 20". We note that when the "Tracking" job failed the tracking of a cell after frame 31 (until the cell was picked), we tracked it manually between two continuous frames based on the closest distance if the distance from the second closest cell was more than 5 times the distance from the first closest cell. Fifth, the picked cell was linked with the closest cell in the last frame before picking.

When the cell was not tracked in all frames from the first to the last frame before the cell picking, the tracking of this picked cell was considered a failure and was not used for further analyses (Dataset S2).

Cropping cell images for machine learning

For successfully tracked cells, stacked cell patches (100 × 100 pixels, 30 frames ordered by time) containing the cell at the center of the image were obtained by the following steps. First, the images obtained by time-lapse imaging (1 min interval, 30 time points) were stitched using the "stitch Multipoint to Large Image" job for focal position 1 (Materials and Methods). Second, for each tracked cell, cell patches (100 × 100 pixels) were cropped from each stitched image using the tracked coordinates of the cell (this was determined in the third step of the section "linking time-lapse image to each isolated single cell") as the center and were aligned by time series.

Conventional image analysis

To cluster cells based on morphological and dynamical features of cells, first, 19 features which are defined by NIS-Elements (version 5.11.00) (Datasets S7 and S10) were extracted from the time-lapse cell images at focal position 2 using the same procedures as "*Linking time-lapse image to each isolated single cell*".

Second, 82 morphological and dynamical features which are defined by CellProfiler 4 (version 4.2.4) (11) (Datasets S8 and S11) were extracted from the stitched time-lapse cell images at focal position 2 (see "*Linking time-lapse image to each isolated single cell*") by cell detection and tracking using the "Object Tracking" example pipeline (version 5) obtained from a website (https://cellprofiler.org/examples). The "ColorToGray" module was replaced by the "ImageMath" module with the "Invert" operation since the analyzed time-lapse cell images were in grey format and the cells in the images were darker than the background. The "Typical diameter of objects, in pixel units (Min, Max)" in the "IdentifyPrimaryObjects" module was set as 20 and 100 for the 3mix-ALPS-timelapse cells and 10 and 50 for the PBMCs to fit the sizes of the analyzed cells (1 pixel = 0.324 μm). The "Lower and upper bounds on threshold" in the "IdentifyPrimaryObjects" module was set as 0.01 and 1 for the 3mix-ALPS-timelapse cells and 0.5 and 1 for the PBMCs. The features were extracted by the "MeasureObjectSizeShape" module with the "Calculate the advanced features" option and the "MeasureObjectIntensity" module. Then, the cells were tracked between frames using the "TackObjects" module with parameters: "Choose a tracking method = Distance, Maximum pixel distance to consider matches = 500". The tracks of three 3mix-ALPS-timelapse cells (in the 1,008 cells linked with scRNA-seq data) and two PBMCs

(in the 346 cells linked with scRNA-seq data) were not completed for all 30 time-lapse images and were not included in the further analyses.

Third, 24 morphological and dynamical features which are defined by TrackMate 7 (version 7.7.2) (12) (Datasets S9 and S12) were extracted from the stitched time-lapse cell images at focal position 2 (see "*Linking time-lapse image to each isolated single cell*") by cell detection and tracking using the "TrackMate 7" Plugin in Fiji(ImageJ) (13). After the images were inverted by the "Invert" module in Fiji, the cells in each image were detected using the "Thresholding detector" module in TrackMate 7 with an automatically generated intensity threshold and the "Simplify contours" function. Then, the cells were tracked between frames using the "Kalman tracker" module in TrackMate 7 with parameters: "Initial search radius = 500 (pixel), Search radius = 500 (pixel), Max frame gap = 0 (frame)". The features were automatically calculated by TrackMate 7 after detection or tracking. The tracks of 60 3mix-ALPS-timelapse cells (in the 1,008 cells linked with scRNA-seq data) and 10 PBMCs (in the 346 cells linked with scRNA-seq data) were not completed for all 30 time-lapse images and were not included in the further analyses. We note that the "Thresholding detector" were chosen since other available detectors (which do not use deep learning) in TrackMate 7 did not detect the cells correctly judged by eye, and the "Kalman tracker" were chosen since it tracked the highest number of cells that linked with scRNA-seq data among other available trackers in TrackMate 7.

(8) Sorting-based single cell isolation

A mixture of T cells, leukemia cells, and HPCs was prepared as described in the section "Automated single cell isolation using the ALPS" (Materials and Methods). Then, using BD FACS Aria III (BD Biosciences), 36 leukemia cells, 36 T cells, and 24 HPCs were sorted in this order into one 96-well PCR plate (one cell in one well) based on the signals of EGFP (positive for leukemia cells and negative for the others), PE (positive for T cells and negative for the others), and Alx648 (positive for HPCs and negative for the others), respectively. Two plates were prepared.

(9) Deep learning

Cell type classification was performed for three-type-mixed cells and PBMCs respectively. For each classification, a 2D convolutional neural network (CNN) that extracts image-based features of the cells, combined with a recurrent neural network (RNN) that models the temporal information of the cells, was performed by two architectures respectively: a deep residual network (ResNet) (14) architecture combined with a long short-term memory (LSTM) (15) architecture (ResNet-LSTM classification architecture) and a LeNet architecture (16) with three convolutional layers combined with an LSTM architecture (LeNet-LSTM classification architecture). Regression of each of the 300 genes with high cell-to-cell variation (determined by the same method described in SI Appendix Text 6) using a ResNet architecture combined with an LSTM architecture (ResNet-LSTM regression architecture) was performed for three-type-mixed cells. For both classification and regression, the leave-one-out cross-validation strategy (17) was applied; cells in one plate were used for testing, and cells in all other plates were used for training.

The ResNet-LSTM classification architecture (SI Appendix Fig. S9A) was modified from the combined architecture (18) of ResNet-18 and LSTM, which was implemented via PyTorch (19): first, the pixel intensity of each frame (100 × 100 pixels) for the red, green and blue channels was normalized to the mean = [0.485, 0.456,

0.406] and standard deviation = [0.229, 0.224, 0.225], respectively, using the transform.Normalize function so that transfer learning (20) with the pretrained weights of ResNet-18 could be used in the next step. Second, each frame was processed through ResNet-18 to extract image features using the pretrained weights of ResNet-18. To transfer the extracted image features to the LSTM in the next step, the last two layers of ResNet-18 (the fully connected layer and the softmax output layer) were removed, and a max-pooling layer (21) with a kernel size of 3 and a stride of 2 was added so that 512 features were output from each frame. Third, the output features of all 30 frames (512 × 30) were input to the LSTM with 2 layers containing 100 hidden units. Fourth, the output of the last frame from the LSTM was connected with a number of units equal to the number of classes (i.e., number of cell types) by a fully connected layer.

For training using the ResNet-LSTM classification architecture, the training data were divided into a training set (80%) and a validation set (20%). In total, 100 epochs of training with batch size 8 were performed, and the weights of the LSTM and the last fully connected layers were updated for each epoch using the Adam algorithm with a learning rate of 0.001 to decrease the training loss, which was calculated using the cross-entropy loss function. The weights of the network with the lowest loss among all 100 epochs in the performance on the validation set were chosen for the final test on the testing data.

The ResNet-LSTM regression architecture (SI Appendix Fig. S9C) and its training were basically the same as the ResNet-LSTM classification architecture except for the following modifications: the output of the last frame from the LSTM network was connected with a single unit by a fully connected layer, and the cross-entropy loss function was replaced by the mean-squared-error loss function. The log-transformed gene expression levels were used as input.

The LeNet-LSTM classification architecture (SI Appendix Fig. S9B) was modified from the architecture reported (22) for cell type classification based on time-lapse cell patches and was implemented via Keras (23). The architecture was sequentially composed of the following layers: an input layer (30 × 100 × 100 × 1, frames × height × width × channel (grayscale)), a convolutional layer with 32 filters and a kernel size of 9 using Rectified Linear Unit (ReLU) (24) activation, a BatchNormalization layer (25), a max-pooling layer with a kernel size of 2, a convolutional layer with 64 filters and a kernel size of 3 using ReLU activation, a BatchNormalization layer, a max-pooling layer with a kernel size of 2, a convolutional layer with 64 filters and a kernel size of 3 using ReLU activation, a max-pooling layer with a kernel size of 2, a dropout layer with a ratio of 0.25, a Flatten layer, a fully connected layer with 400 units using ReLU activation, a LSTM layer with 500 hidden units, a TimeDistributed layer, and a fully connected layer with a number of units equal to the number of classes (i.e., number of cell types) using softmax activation. For training, the training data were divided into a training set (85%) and a validation set (15%), and all gray values of the images were divided by the highest gray value of the imaging system (65,535). In total, 200 epochs of training with batch size 8 were performed, and the weights of the whole network were updated for each epoch using stochastic gradient descent with the default parameters to decrease the training loss, which was calculated using the categorical cross-entropy loss function. The weights of the network that achieved the highest-accuracy performance on the validation set among all 200 epochs were chosen for the final test on the testing data.

(10) Selected cells shown in Fig. 2C

For each cell type, three cells correctly predicted by deep learning and two cells incorrectly predicted were randomly selected from the all PBMCs. We performed this process five times and showed one of them as a typical example. For each cell in Fig. 2C, Plate_ID (PBMC-timelapse-plateX), Well_ID_in_Plate, and prediction result are as follows. $CD4^+$ T cell 1-5: 3, 2, correct; 3, 31, correct; 7, 36, correct; 8, 6, incorrect; 8, 81, incorrect. $CD8^+$ T cell 1-5: 2, 8, correct; 3, 77, incorrect; 6, 56, correct; 7, 48, correct; 8, 58, incorrect. B cell 1-5: 3, 11, correct; 4, 31, correct; 5, 23, incorrect; 7, 68, incorrect; 8, 79, correct.

Figures

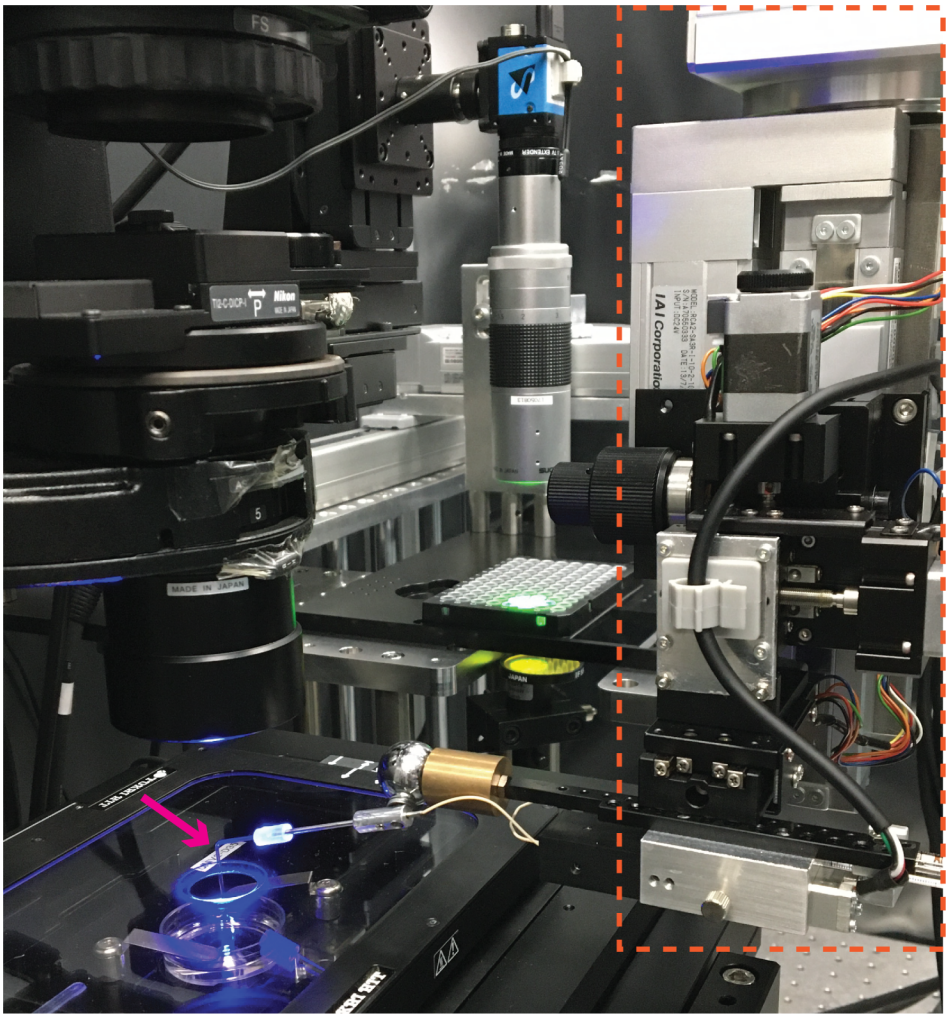


Figure S1: Photo of ALPS.

A cell picker was equipped on a Nikon microscope (Eclipse Ti2-E). The sample dish on the microscope was illuminated by blue LED light. A glass needle (magenta) located in the dish was controlled by an automated robot arm (orange). A 96-well plate on the motorized plate stage was illuminated by green LED light under a camera (sky blue) (Fig. 1A, Materials and Methods).

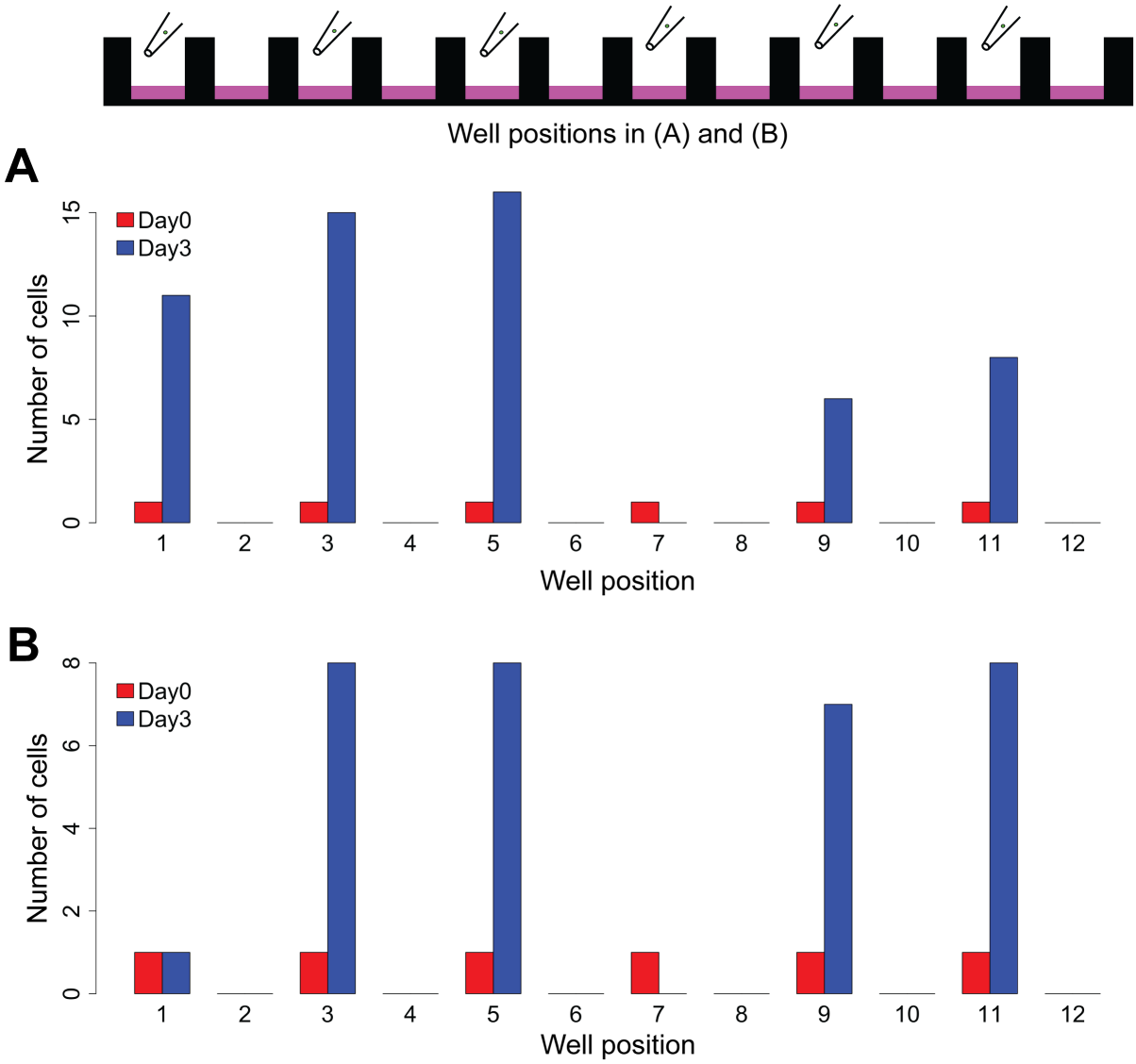


Figure S2: The number of cell(s) at day 0 (isolated by ALPS) and day 3 in each well in the same raw of a 96-well plate (SI Appendix Text 3).

(A) Mouse leukemia cells. **(B)** Human Jurkat cells. One cell was deposited by ALPS in the 1st, 3rd, 5th, 7th, 9th, and 11th wells, and no cells were deposited in the 2nd, 4th, 6th, 8th, 10th, and 12th wells.

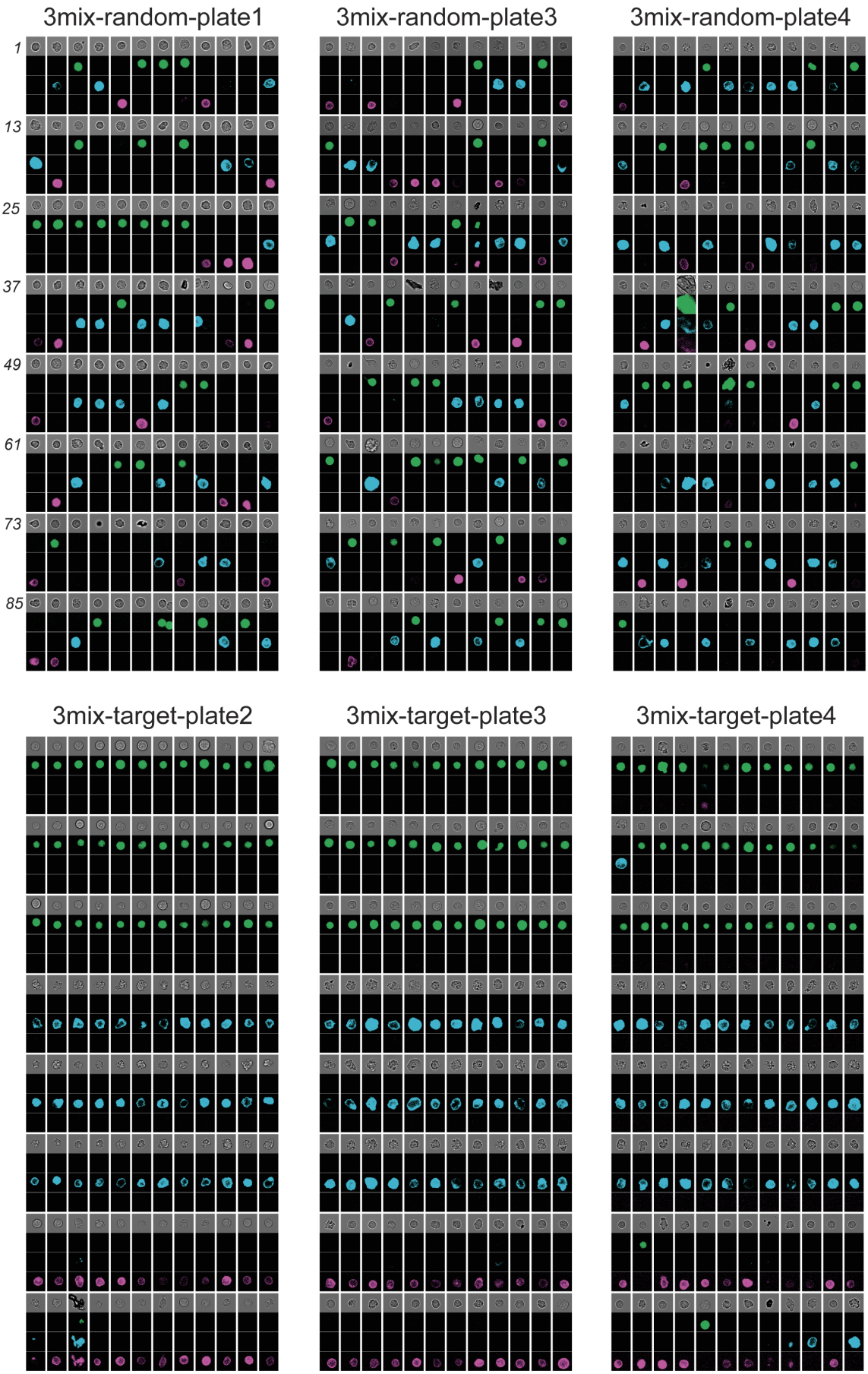


Figure S3: Images of the cells from 3mix-ALPS-random&target-multimode.

Four images (as column), bright-field image (top) and fluorescent images of GFP (green), PE (cyan), and Alx648 (magenta) of each cell from 3mix-ALPS-random-multimode (3mix-random-plate1, 3, and 4 shown here, plate 2 shown in Fig. 1C) or 3mix-ALPS-target-multimode (3mix-target-plate2, 3, and 4 shown here, plate 1 shown in Fig. 1D). Numbers on the left indicate the order of isolation. Image size, 26 μm × 26 μm. The contrast of the shown images was changed linearly.

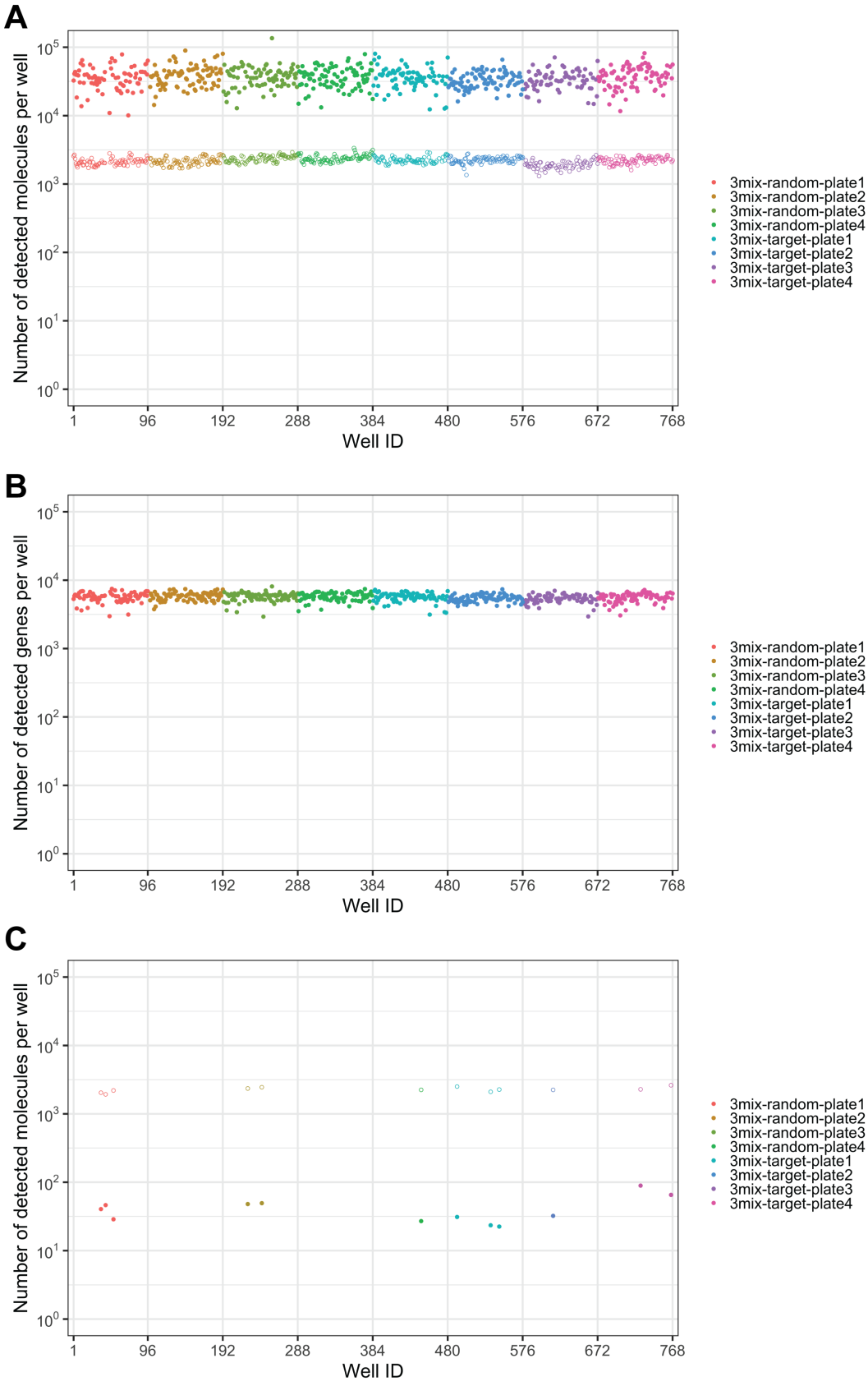


Figure S4: The number of detected RNA molecules and detected genes of each well for the cells from 3mix-ALPS-random-multimode (4 plates) and 3mix-ALPS-target-multimode (4 plates).

The order of the well IDs in each plate corresponded to the order of the 96 isolations. Open circles, ERCC Spike-In RNA; filled circles, RNAs mapped to murine genomes. **(A** and **B)** The numbers of detected RNA molecules (A) and detected genes (B) in the wells in which one cell was isolated and the sequencing quality was high (SI Appendix Text 6). **(C)** The numbers of detected RNA molecules in wells where no cell was picked and deposited. The small number of detected RNA molecules in these wells suggested that the contamination of RNA molecules by ALPS&RNA-seq was low.

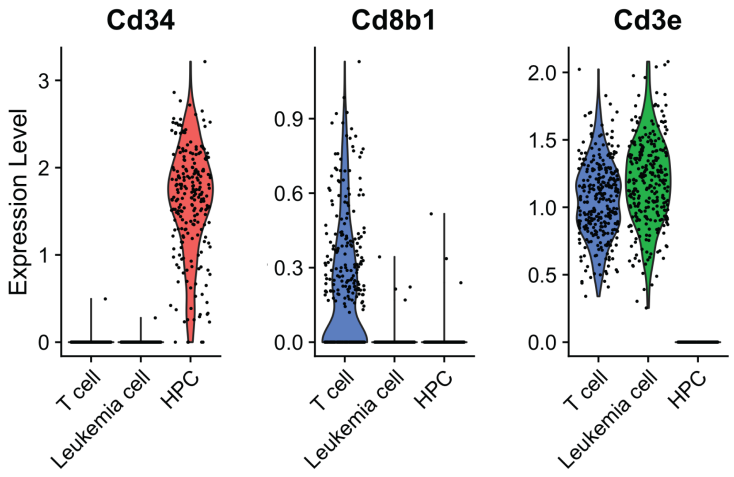


Figure S5: Violin plots showing the expression distributions of three cell marker genes in the three clusters of the cells from 3mix-ALPS-random&target-multimode.

The cell type of each cluster was determined by the following marker genes: T cells were identified by the expression of Cd8b1 because we introduced this gene into this cell line; leukemia cells were identified by the expression of Cd3e and without the pronounced expression of Cd8b1 and Cd34 (26); HPCs were identified by the expression of Cd34 (27).

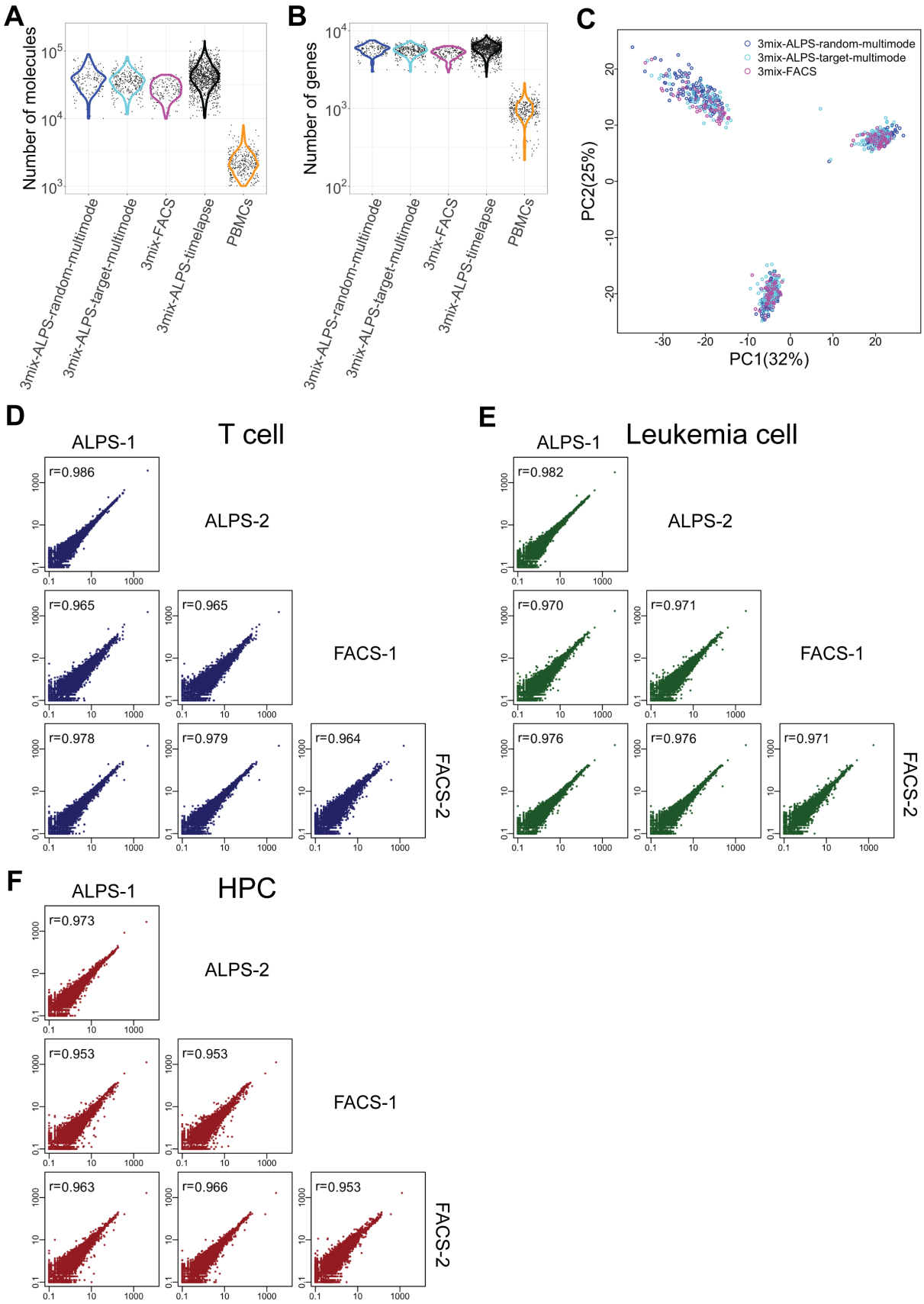


Figure S6: Comparison of the scRNA-seq results among different isolation methods.

(A and **B)** Violin plots showing the numbers of detected RNA molecules (A) and the numbers of detected genes (B) of single cells from 3mix-ALPS-random-multimode, 3mix-ALPS-target-multimode, and 3mix-ALPS-

timelapse, of three-type-mixed cells isolated by a sorter (3mix_FACS), and of PBMCs. **(C)** PCA of single cells from 3mix-ALPS-random-multimode, 3mix-ALPS-target-multimode, and 3mix-FACS. (**D-F)** Comparison of average RNA expression levels of detected genes between the cells from 3mix-ALPS-random-multimode (ALPS-1: 3mix-random-plate1; ALPS-2: 3mix-random-plate2) and 3mix-FACS (FACS-1: 3mix-FACS-plate1; FACS-2: 3mix-FACS-plate2) for T cells (D), leukemia cells (E), and HPCs (F). The RNA expression levels were averaged from the first 16 cells in each plate for each cell type. *r*, Pearson coefficient calculated in log scale.

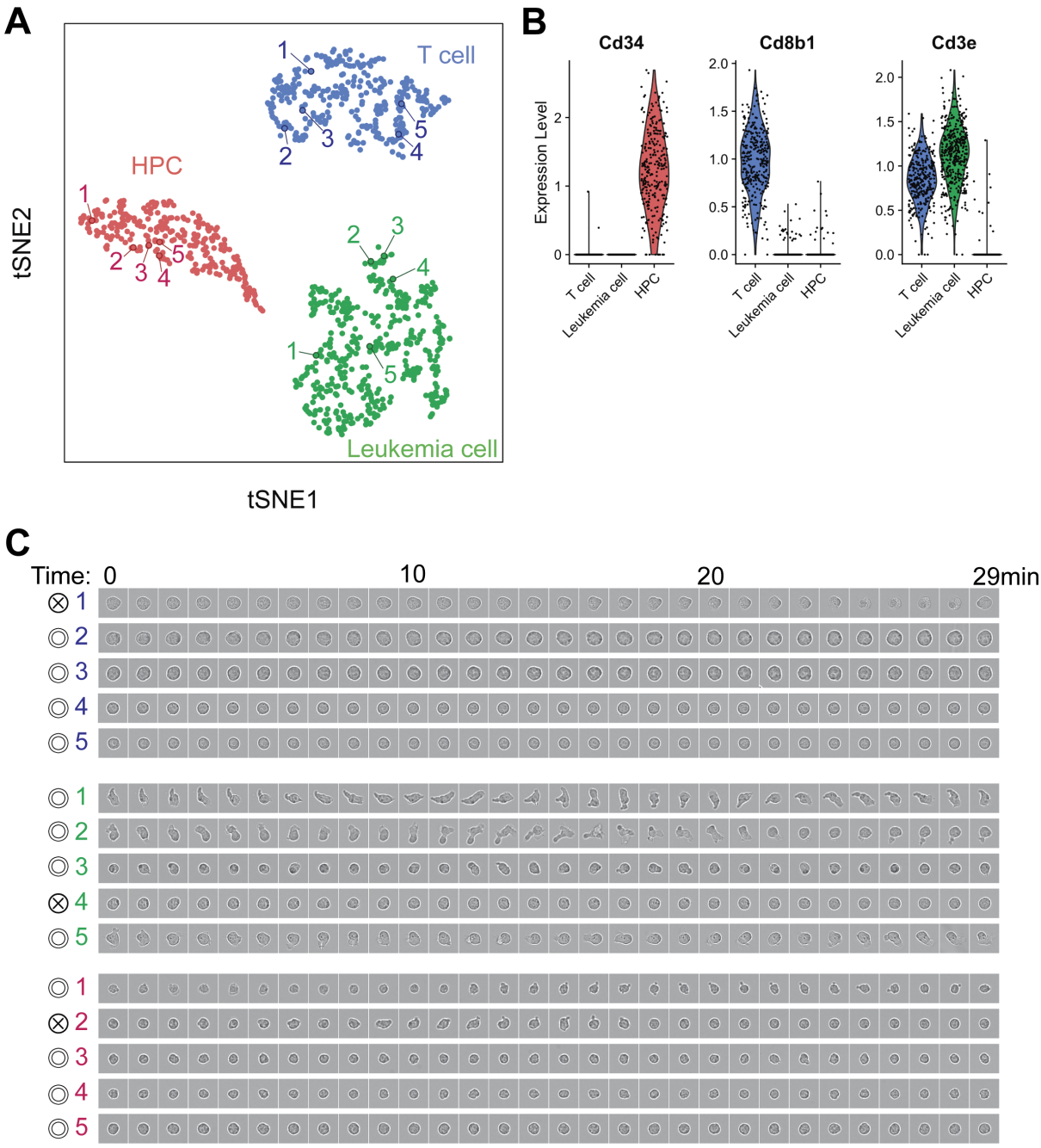


Figure S7: Whole-genome expression and time-lapse images linked for the same cell from 3mix-ALPS-timelapse.

(A) t-SNE visualization of the cells from 3mix-ALPS-timelapse. Colors represent cell types determined by gene expression. Numbers correspond to the cells shown in (C). **(B)** Violin plots showing the expression distributions of three cell marker genes in the three clusters of the cells from 3mix-ALPS-timelapse. Cell types for the three clusters were determined as in SI Appendix Fig. S5. **(C)** Examples of bright-field time-lapse images of the cells from 3mix-ALPS-timelapse. ◎ indicates that the cell type was correctly predicted in deep learning classification, and ⊗ indicates that it was not. Images of four successfully predicted cells and one unsuccessfully predicted cell from the 1st well of the 3mix-timelapse-plate1 for each type are shown (one cell was skipped because another cell was present in some frames, though it was successfully predicted). Image size, 32 μm × 32 μm. The contrast of the shown images was changed linearly.

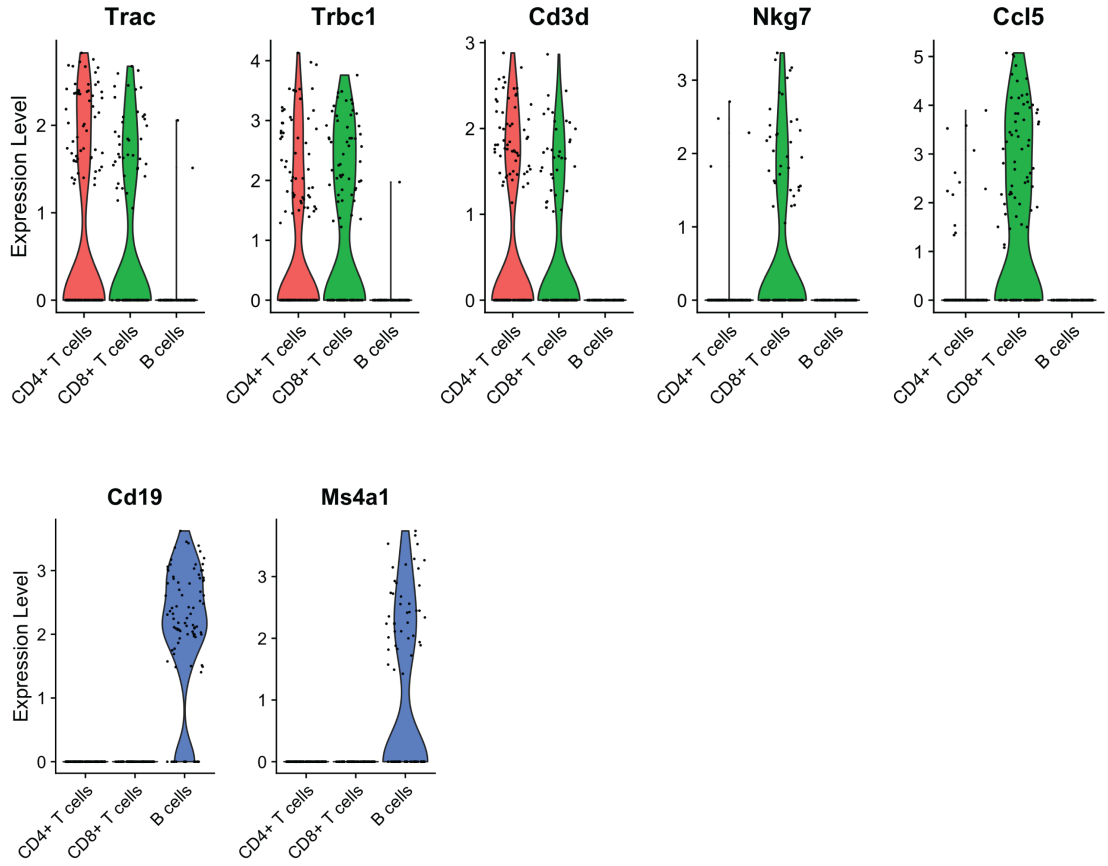


Figure S8: Violin plots showing the expression distributions of selected cell marker genes in PBMC subpopulations obtained by clustering.

The cell types for these cell subpopulations were determined as follows (28–30): $CD4^+$ T cells were identified by the expression of Trac, Trbc1 and Cd3d and without the pronounced expression of Nkg7 and Ccl5; $CD8^+$ T cells were identified by the expression of Trac, Trbc1, Cd3d, Nkg7, and Ccl5; B cells were identified by the expression of Cd19 and Ms4a1.

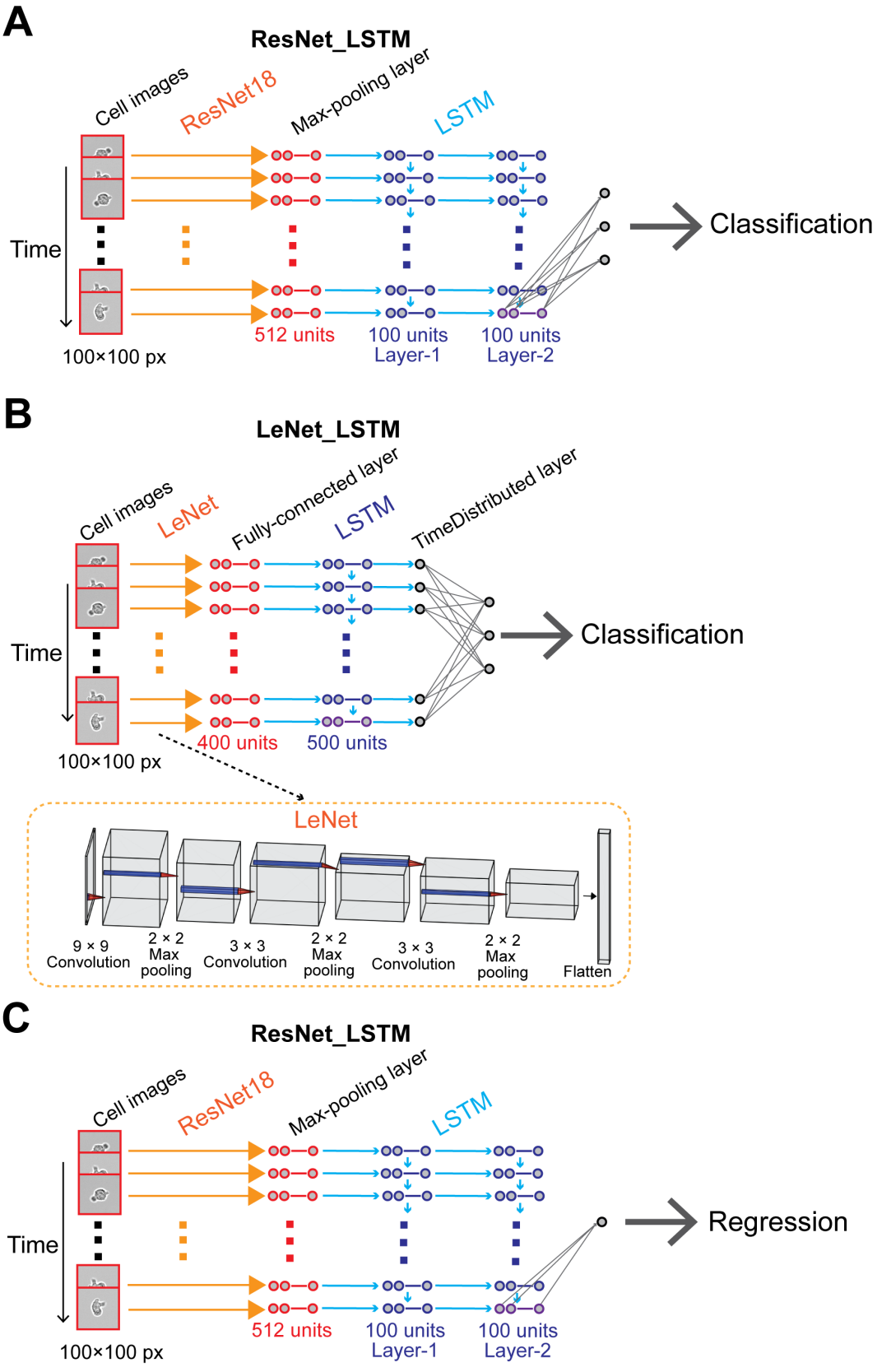


Figure S9: Schematic of the deep learning architectures (details in SI Appendix Text 9) used for classification and regression.

The LeNet architecture was drawn using NN-SVG (31).

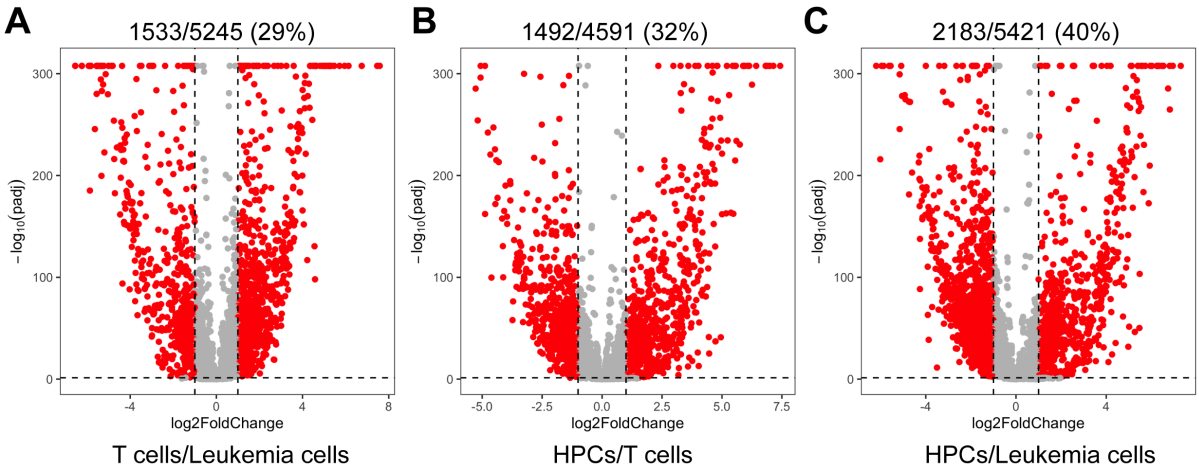


Figure S10: Volcano plot showing the differential gene expression between the cell types of 3mix-ALPS-timelapse cells.

The padj and log2FoldChange were calculated using DESeq2 (32); the genes whose detected number of molecules was ≥10 in at least one cell were considered as detected genes. Horizontal dotted line, padj = 0.05; vertical dotted line, log2FoldChange = -1 or 1. The number of differentially expressed genes (red dots)/the number of detected genes (all dots) and the proportion of differentially expressed genes (in the parentheses) are shown.

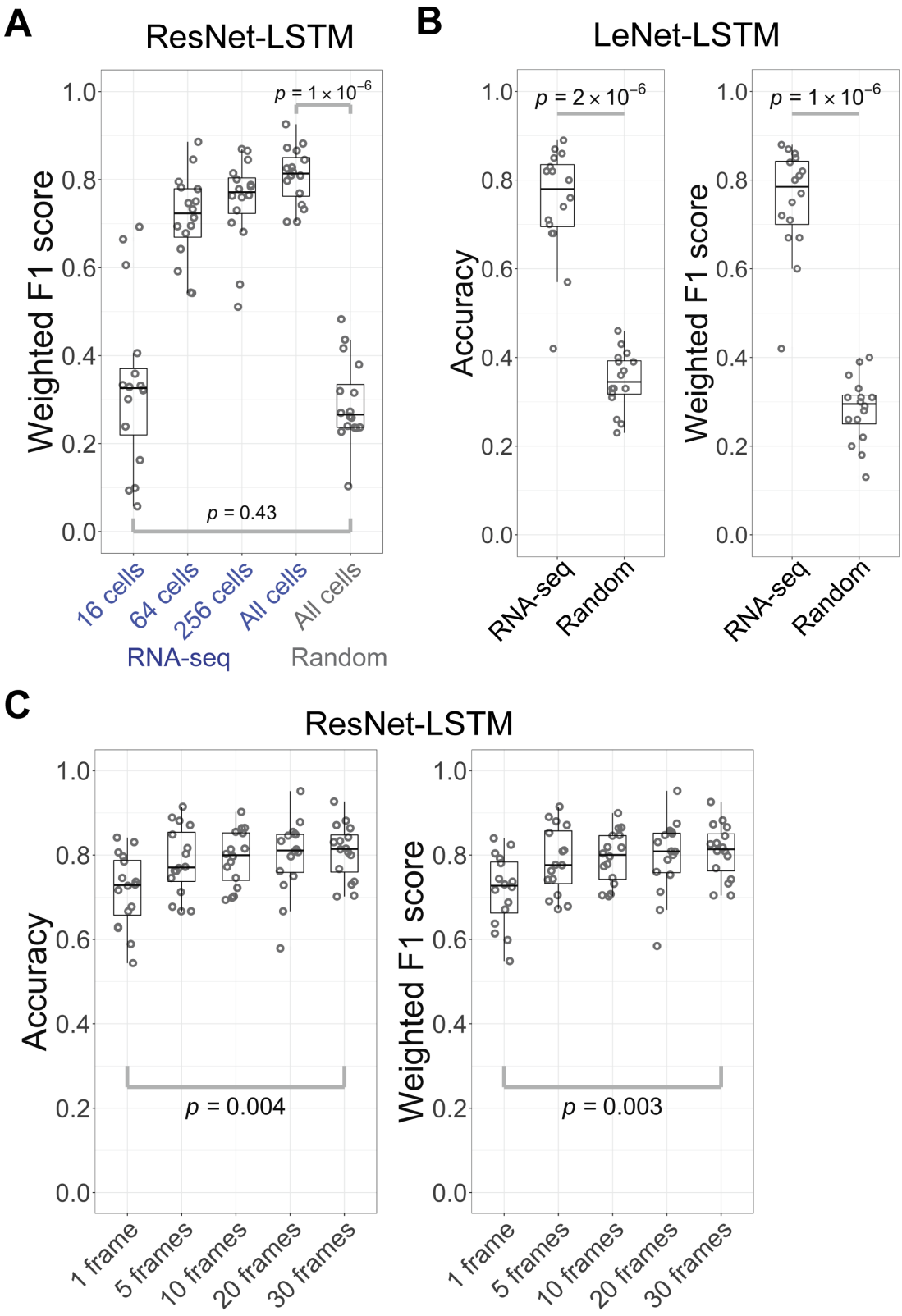
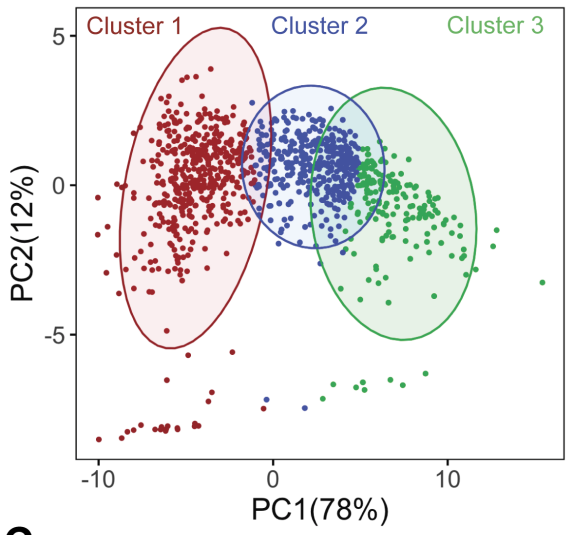


Figure S11: Deep learning-based classification for time-lapse imaged cells from 3mix-ALPS-timelapse.
(A) Weighted F1 score for predicting scRNA-seq-determined cell types using ResNet-LSTM. RNA-seq, cell types were determined by scRNA-seq. Random, cell types were labeled randomly ((**B**) as well). 16, 64, 256, and All

cells, 16, 64, 256, and all cells were randomly selected from 15 plates for training (cells in one plate were used for testing). The *p* value was calculated by the Kruskal-Wallis rank sum test ((**B** and **C**) as well). **(B)** Accuracy and weighted F1 score for predicting scRNA-seq-determined cell types using LeNet-LSTM. **(C)** Accuracy and weighted F1 score for predicting scRNA-seq-determined cell types by ResNet-LSTM using different numbers of frames.

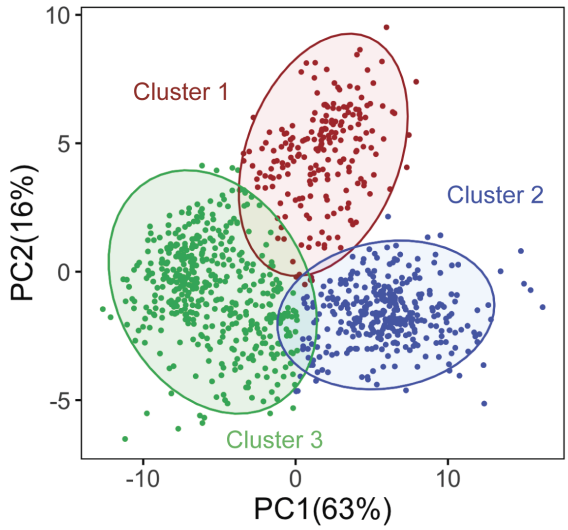
A



B

Feature-determined cluster	Cluster 1	Cluster 2	Cluster 3	Accuracy
Assigned Cell type	HPC	Leukemia cell	T cell	0.42
	HPC	T cell	Leukemia cell	0.47
	T cell	Leukemia cell	HPC	0.18
	T cell	HPC	Leukemia cell	0.15
	Leukemia cell	HPC	T cell	0.35
	Leukemia cell	T cell	HPC	0.43

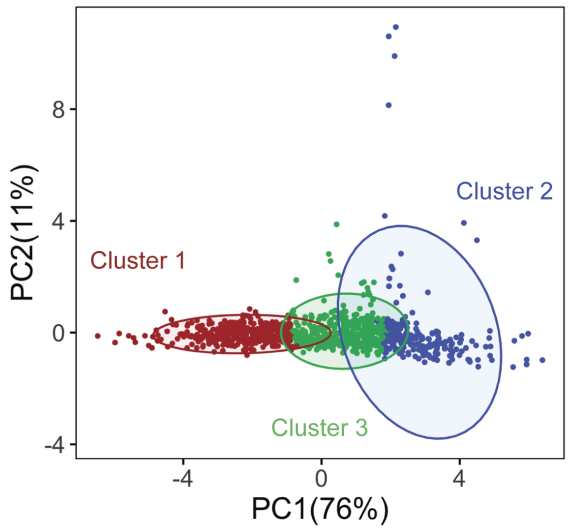
C



D

Feature-determined cluster	Cluster 1	Cluster 2	Cluster 3	Accuracy
Assigned Cell type	HPC	Leukemia cell	T cell	0.30
	HPC	T cell	Leukemia cell	0.70
	T cell	Leukemia cell	HPC	0.19
	T cell	HPC	Leukemia cell	0.40
	Leukemia cell	HPC	T cell	0.11
	Leukemia cell	T cell	HPC	0.29

E



F

Feature-determined cluster	Cluster 1	Cluster 2	Cluster 3	Accuracy
Assigned Cell type	HPC	Leukemia cell	T cell	0.45
	HPC	T cell	Leukemia cell	0.48
	T cell	Leukemia cell	HPC	0.16
	T cell	HPC	Leukemia cell	0.20
	Leukemia cell	HPC	T cell	0.36
	Leukemia cell	T cell	HPC	0.34

Figure S12: Conventional image analysis of cells from 3mix-ALPS-timelapse.

(A) PCA of the cells with 24 morphological and dynamical features (Dataset S7) obtained by NIS-Elements. For each cell, the average value of the 30 images for each feature was used. For each feature, the average value of each cell was divided by the median of the average values of all cells and was transformed to the log2 scale. For the "Tangential acceleration" feature, the negative average values (2.8% of the cells) were set as the minimum of the positive average values before division. Colors correspond to the three clusters obtained by Kmeans clustering. 95% confidence ellipse for each cluster is shown. **(B)** The feature-determined clusters in (A) were assigned as the three cell types respectively in any possible one-to-one correspondence. The accuracy of each correspondence was calculated as the number of cells matched with their scRNA-seq-determined cell types divided by the total number of cells. **(C)** PCA of the cells with 82 morphological and dynamical features (Dataset S8) obtained by CellProfiler 4. The features were averaged and normalized using the same strategy as (A). Before averaging, all values of the "Orientation" feature were added by 90^o to rescale the range from $-90^o \sim 90^o$ to $0^o \sim 180^o$, and for each of the "TrajectoryX","TrajectoryY", "Central Moment", "Hu Moment", "Normalized Moment" and "Inertia Tensor" features, the minimum value was subtracted from all values. Colors and ellipses are the same as (A). **(D)** The accuracies for the clustering in (C) calculated using the same strategy as (B). **(E)** PCA of the cells with 24 morphological and dynamical features (Dataset S9) obtained by TrackMate 7 in Fiji(ImageJ). The features were averaged and normalized using the same strategy as (A). Before averaging, all values of the "Ellipse Theta" feature were added by π to rescale the range from $-\pi \sim \pi$ to $0 \sim 2\pi$, and the minimum value of the "Ellipse X0" or "Ellipse Y0" feature was subtracted from all values of the "Ellipse X0" or "Ellipse Y0" feature. Colors and ellipses are the same as (A). **(F)** The accuracies for the clustering in (E) calculated using the same strategy as (B).

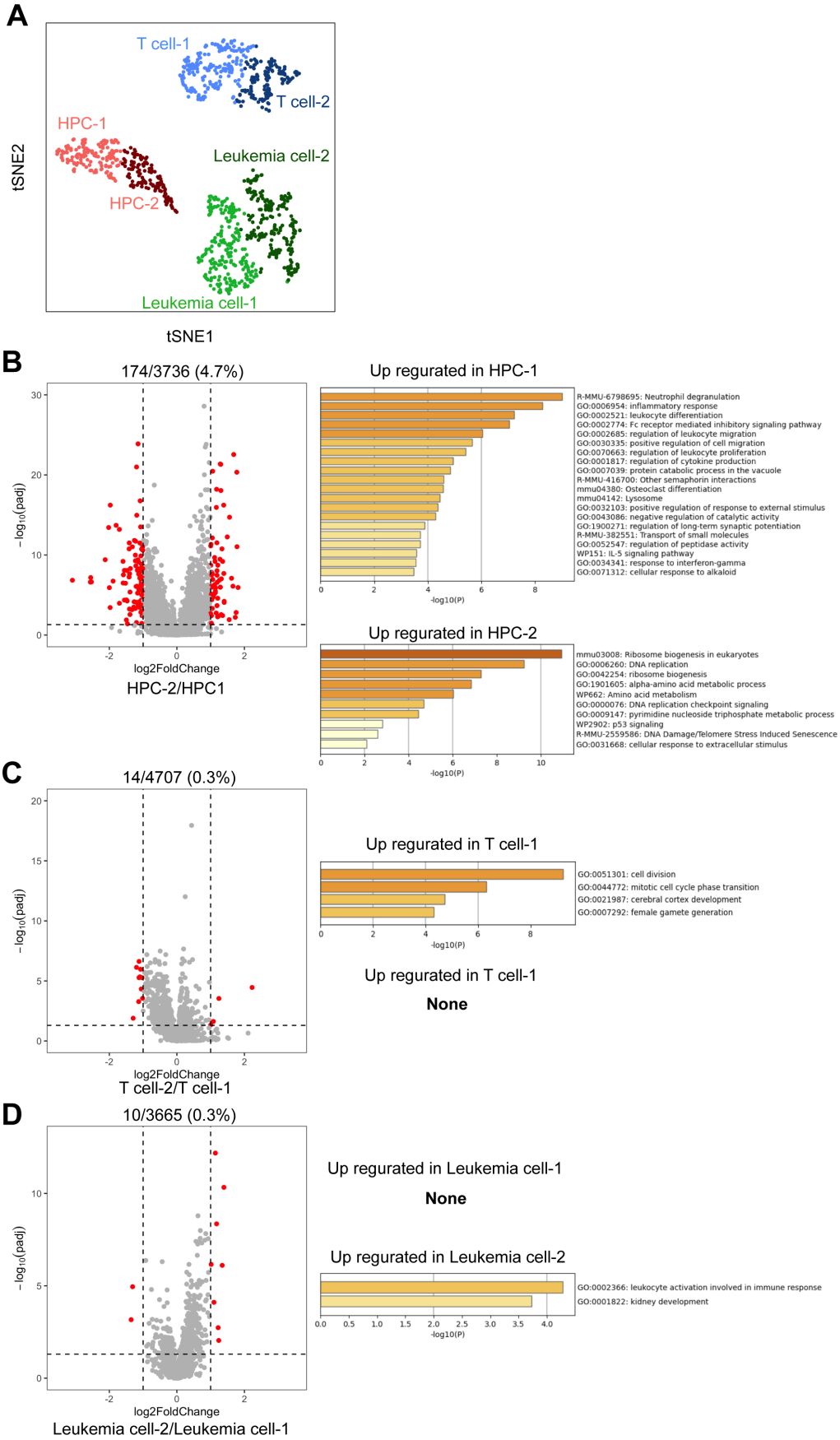


Figure S13: Sub-clusters of 3mix-ALPS-timelapse cells represent biological states within each cell type.

(A) t-SNE visualization of the cells from 3mix-ALPS-timelapse which were clustered into six sub-clusters. Each cell type contains two sub-clusters (T cell-1, T cell-2, Leukemia cell-1, Leukemia cell-2, HPC-1, and HPC-2). **(B)** Volcano plot (in the same format as Fig. S10) showing the differential gene expression between the sub-clusters, HPC-1 and HPC-2, and the Metascape bar graph for viewing top non-redundant enrichment clusters with statistical significances. **(C)** Same as (B) for sub-clusters, T cell-1 and T cell-2. **(D)** Same as (B) for sub-clusters, Leukemia cell-1 and Leukemia cell-2.

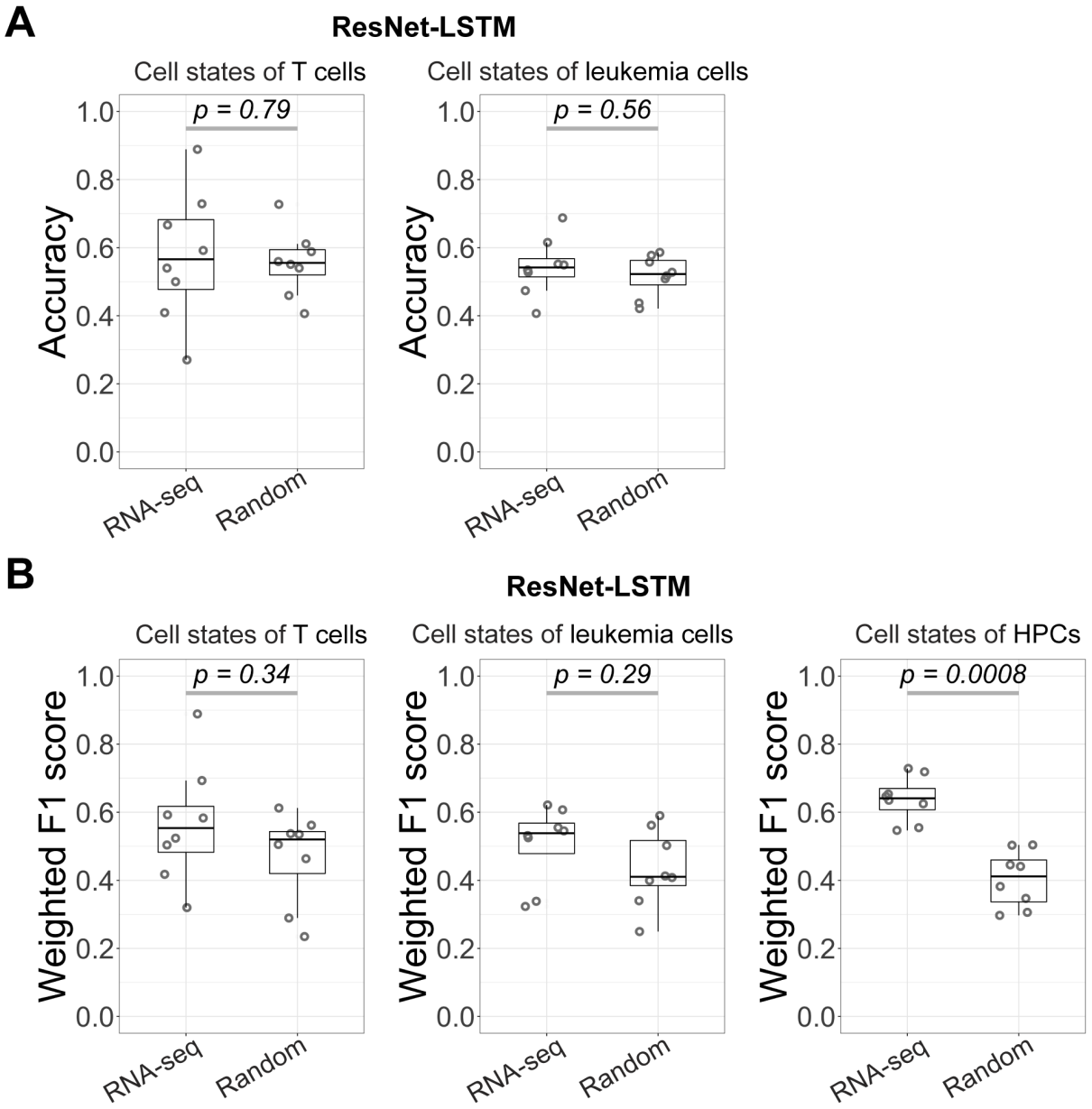


Figure S14: Deep learning-based classification using ResNet-LSTM for the different biological states of 3mix-ALPS-timelapse cells.

(A) Accuracy for predicting scRNA-seq-determined cell states for T cells and leukemia cells respectively using ResNet-LSTM. **(B)** Weighted F1 score for predicting scRNA-seq-determined cell states for T cells, leukemia cells, and HPCs respectively using ResNet-LSTM. For all figures, RNA-seq, cell states were determined by scRNA-seq; Random, cell states were labeled randomly; the *p* value was calculated by the Kruskal-Wallis rank sum test.

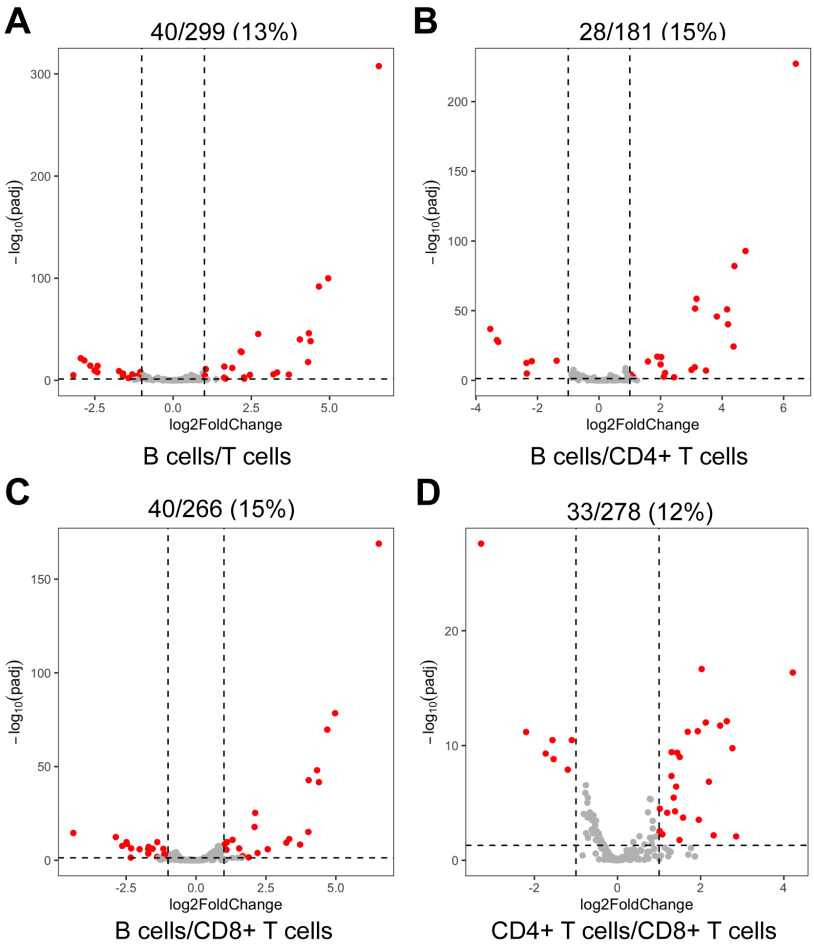


Figure S15: Volcano plot showing the differential gene expression between the cell types of PBMCs.

Shown in the same format as Fig. S10.

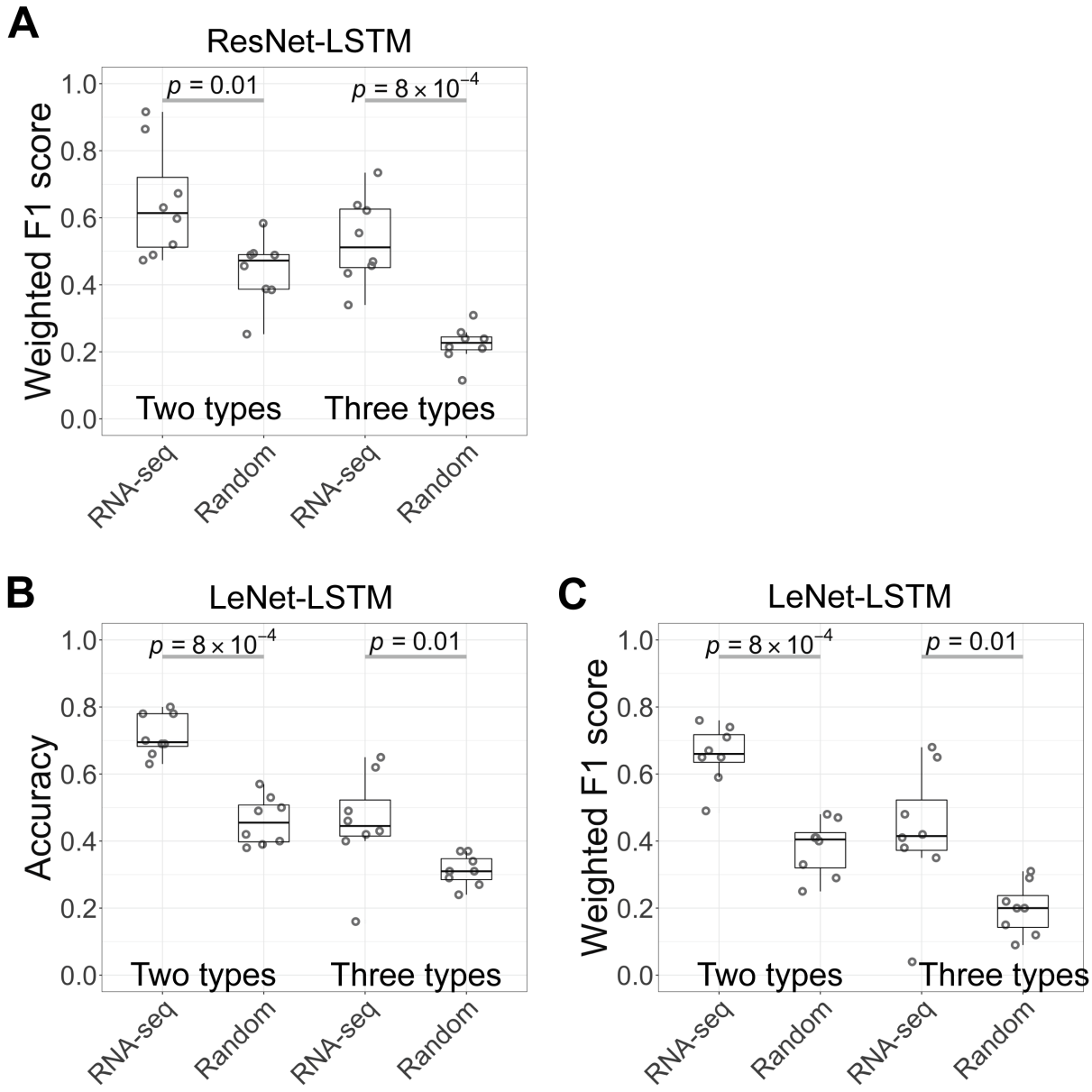
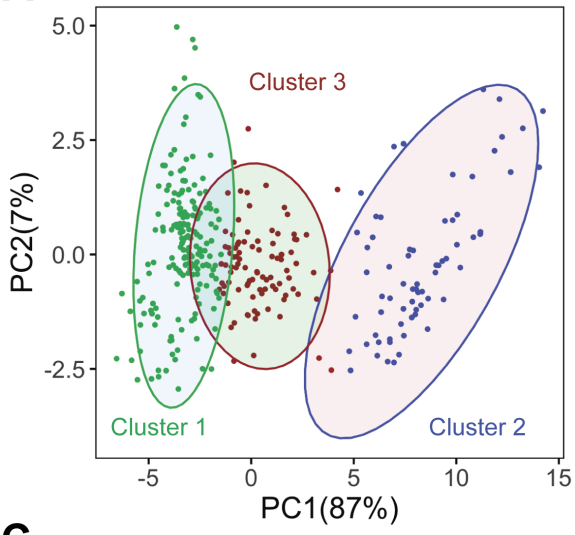


Figure S16: Deep learning-based classification of time-lapse imaged PBMCs.

(A) Weighted F1 score for predicting scRNA-seq-determined cell types using ResNet-LSTM. RNA-seq, cell types were determined by scRNA-seq. Random, cell types were labeled randomly. Two types, PBMCs were clustered into T cells and B cells. Three types, PBMCs were clustered into $CD4^+$ T cells, $CD8^+$ T cells, and B cells. The *p* value was calculated by the Kruskal-Wallis rank sum test. **(B** and **C)** Accuracy and weighted F1 score for predicting scRNA-seq-determined cell types using LeNet-LSTM.

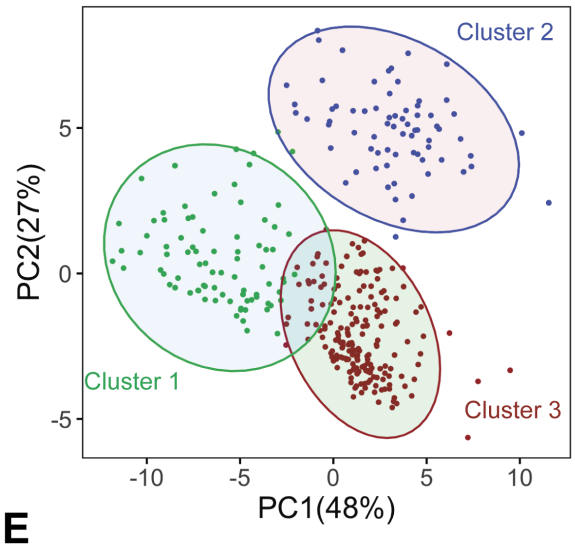
A



B

Feature-determined cluster	Cluster 1	Cluster 2	Cluster 3	Accuracy
Assigned Cell type	CD4⁺ T cell	CD8⁺ T cell	B cell	0.20
	CD4⁺ T cell	B cell	CD8⁺ T cell	0.22
	B cell	CD4⁺ T cell	CD8⁺ T cell	0.14
	CD8⁺ T cell	B cell	CD4⁺ T cell	0.40
	B cell	CD8⁺ T cell	CD4⁺ T cell	0.23
	CD8⁺ T cell	CD4⁺ T cell	B cell	0.29

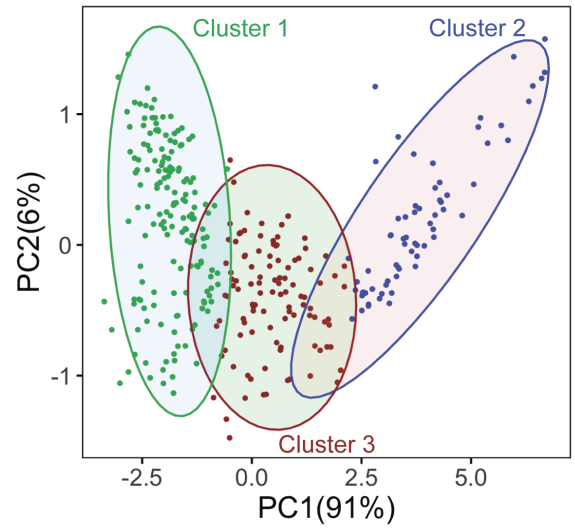
C



D

Feature-determined cluster	Cluster 1	Cluster 2	Cluster 3	Accuracy
Assigned Cell type	CD4⁺ T cell	CD8⁺ T cell	B cell	0.26
	CD4⁺ T cell	B cell	CD8⁺ T cell	0.23
	B cell	CD4⁺ T cell	CD8⁺ T cell	0.21
	CD8⁺ T cell	B cell	CD4⁺ T cell	0.53
	B cell	CD8⁺ T cell	CD4⁺ T cell	0.36
	CD8⁺ T cell	CD4⁺ T cell	B cell	0.41

E



F

Feature-determined cluster	Cluster 1	Cluster 2	Cluster 3	Accuracy
Assigned Cell type	CD4⁺ T cell	CD8⁺ T cell	B cell	0.24
	CD4⁺ T cell	B cell	CD8⁺ T cell	0.29
	B cell	CD4⁺ T cell	CD8⁺ T cell	0.30
	CD8⁺ T cell	B cell	CD4⁺ T cell	0.46
	B cell	CD8⁺ T cell	CD4⁺ T cell	0.39
	CD8⁺ T cell	CD4⁺ T cell	B cell	0.33

Figure S17: Conventional image analysis of PBMCs.

(A) PCA of the cells with 24 morphological and dynamical features (Dataset S10). For each cell, the average value of the 30 images for each feature was used. For each feature, the average value of each cell was divided by the median of the average value of all cells and was transformed to the log2 scale. For the "Tangential acceleration"

feature, the negative average values (0.58% of the cells) were set as the minimum of the positive average values before division. Colors correspond to the three clusters obtained by Kmeans clustering. 95% confidence ellipse for each cluster is shown. **(B)** The feature-determined clusters were assigned as the three cell types respectively in any possible one-to-one correspondence. The accuracy of each correspondence was calculated as the number of cells matched with their scRNA-seq-determined cell types divided by the total number of cells. **(C)** PCA of the cells with 82 morphological and dynamical features (Dataset S11) obtained by CellProfiler 4. The features were averaged and normalized using the same strategy as (A). Before averaging, all values of the "Orientation" feature were added by 90^o to rescale the range from $-90^o \sim 90^o$ to $0^o \sim 180^o$, and for each of the "TrajectoryX","TrajectoryY", "Central Moment", "Hu Moment", "Normalized Moment" and "Inertia Tensor" features, the minimum value was subtracted from all values. Colors and ellipses are the same as (A). **(D)** The accuracies for the clustering in (C) calculated using the same strategy as (B). **(E)** PCA of the cells with 24 morphological and dynamical features (Dataset S12) obtained by TrackMate 7 in Fiji(ImageJ). The features were averaged and normalized using the same strategy as (A). Before averaging, all values of the "Ellipse Theta" feature were added by π to rescale the range from $-\pi \sim \pi$ to $0 \sim 2\pi$, and the minimum value of the "Ellipse X0" or "Ellipse Y0" feature was subtracted from all values of the "Ellipse X0" or "Ellipse Y0" feature. Colors and ellipses are the same as (A). **(F)** The accuracies for the clustering in (E) calculated using the same strategy as (B).

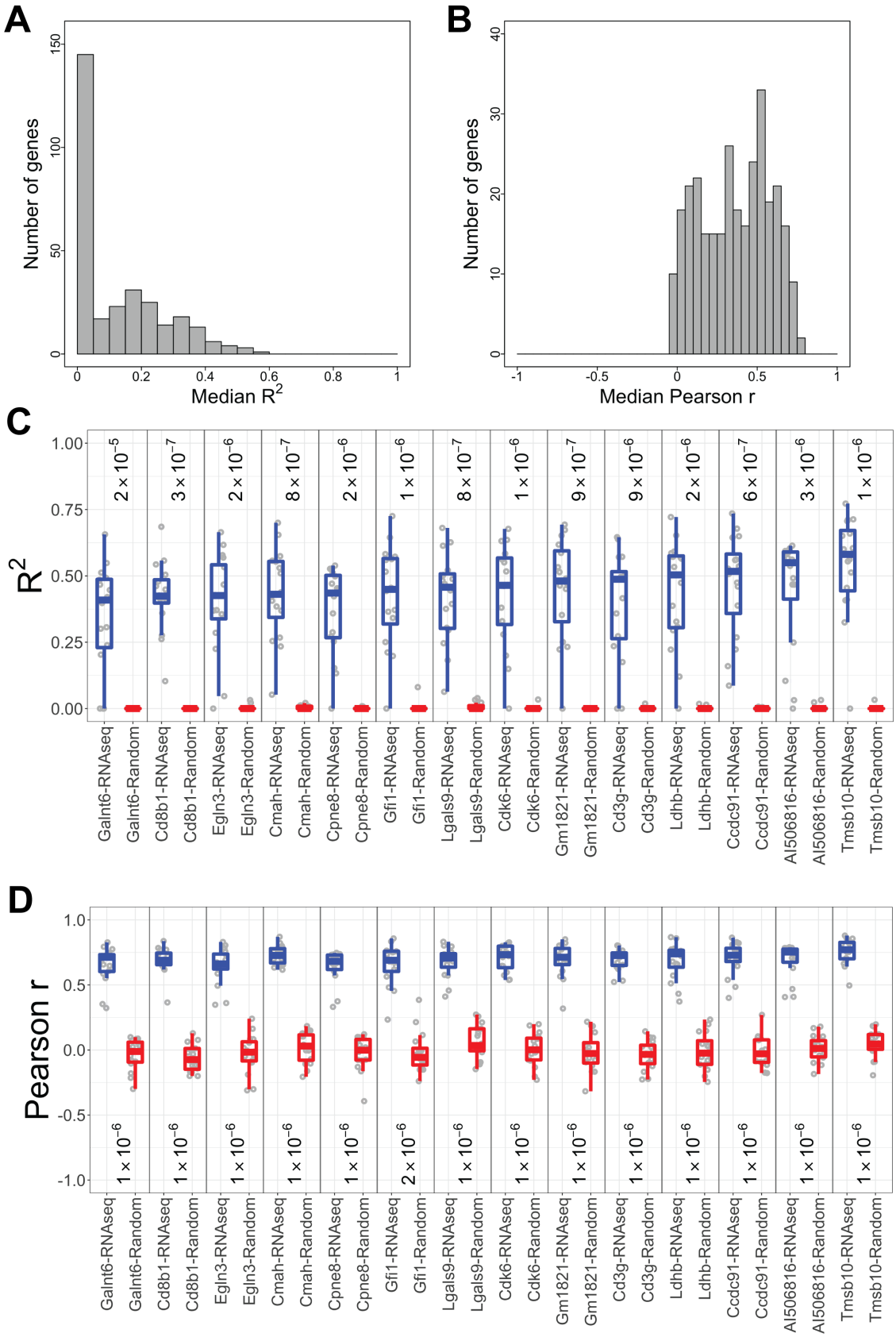


Figure S18: Out-of-sample R^2 and Pearson's r for measured versus predicted gene expression for each of 300 variant genes of the cells from 3mix-ALPS-timelapse.

(A) Distribution of the median out-of-sample R^2 from the 16 predictions. **(B)** Distribution of the median Pearson's r from the 16 predictions. **(C)** Out-of-sample R^2 of the 14 genes that showed the highest median out-of-sample R^2. RNAseq, the gene expression levels were determined by scRNA-seq; Random, the gene expression levels were out of order from scRNA-seq results ((**D**) as well). The *p* values calculated by the Kruskal-Wallis rank sum test are shown ((**D**) as well). **(D)** Pearson's r of the 14 genes as c. Both the out-of-sample R^2 and Pearson's r obtained by the correctly linked datasets of cell images and gene expression levels were significantly higher than those obtained by incorrectly linked datasets, indicating that the deep learning model used here for predicting the gene expression levels of these 14 genes from the cell images achieved robust generalization.

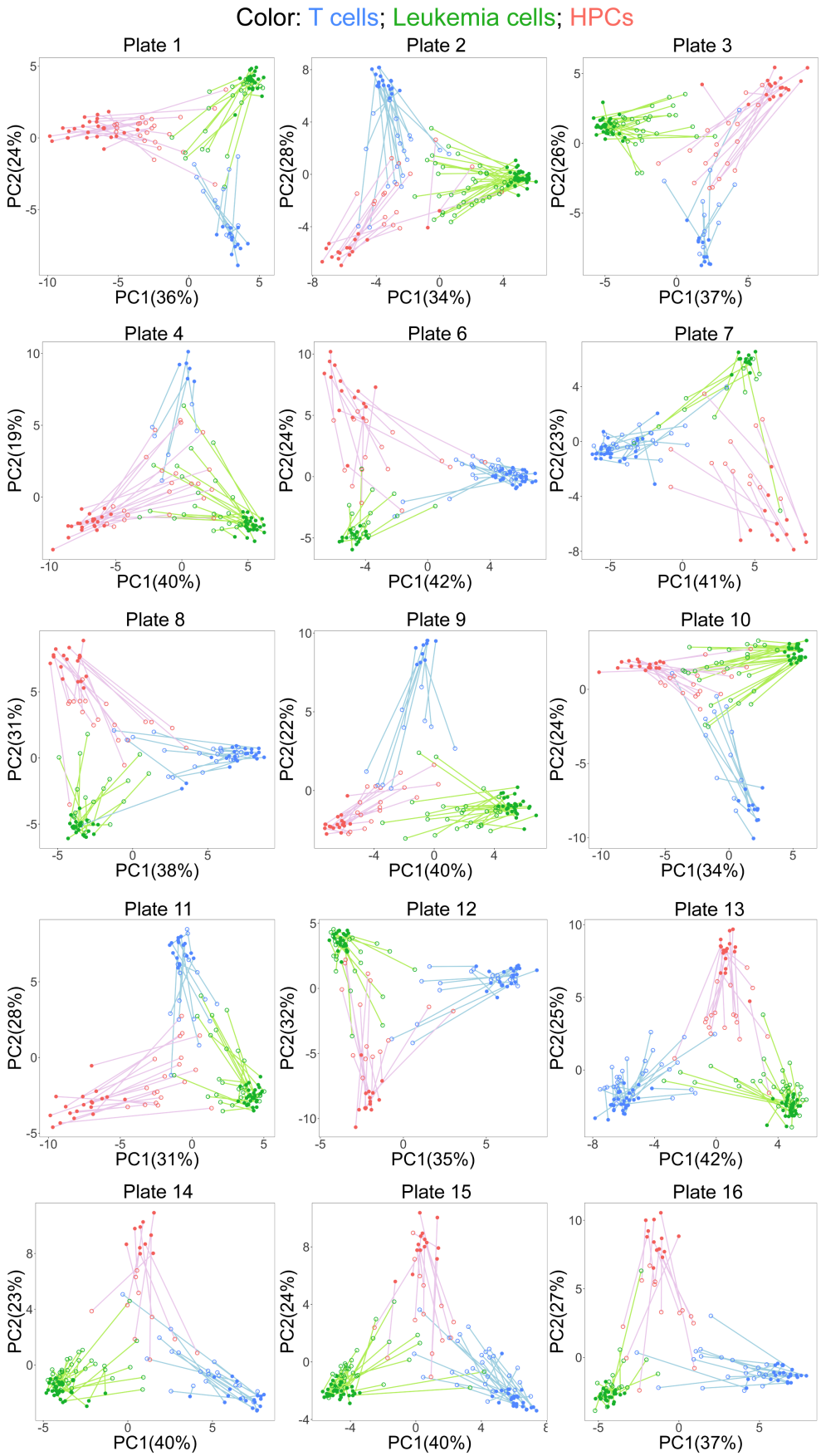


Figure S19: PCA of the cells from each plate of 3mix-ALPS-timelapse with predicted (open circle) and measured (closed circle) expression levels of 300 genes.

Colors correspond to the three different cell types determined by the measured whole transcriptome. Lines link the predicted and measured expressions of the same cell. Plate 1-16 represent 3mix-timelapse-plate1-16. 3mix-timelapse-plate5 is shown in Fig. 4C.

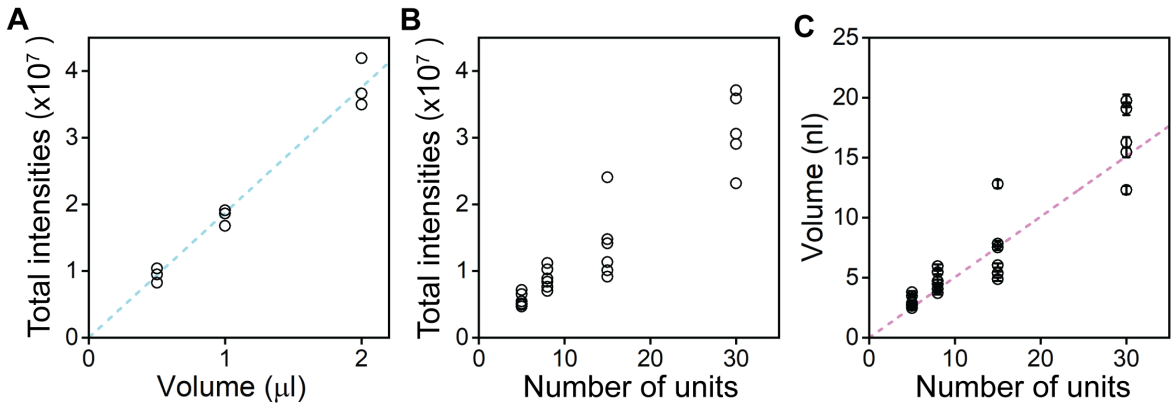


Figure S20: Calibration of the aspiration volume of one volume unit (i.e., a minimum handling volume) of the automated syringe pump in ALPS.

(A) The total fluorescence intensities of fluorescent beads in a solution depending on the solution volume. The fluorescent bead solution (F8823, Thermo Fisher Scientific) was diluted 100,000 times with Milli-Q water and vortexed at 3,200 rpm for 10 s (Vortex Genie 2, Scientific Industries). Then, 0.5 μl, 1 μl, and 2 μl of the diluted solution was aspirated and dispensed by a P-2 Pipetman (Gilson) at the center of a well in a 96-well glass bottom plate (Greiner Bio-One), which covered a circular area with a 1 - 2.2 mm radius. The dispensed solution was dried at room temperature for 3 h. The dried fluorescent beads in each well were imaged using a Nikon microscope with GFP illumination (1% of the CoolLED) (Materials and Methods), 1 ms exposure time, and a 20× objective. The sum of the fluorescence intensities of the pixels in the circle area (2.5 mm radius) after background subtraction was calculated. The background was determined as the average intensity of a ring area (2.5 – 3.5 mm radius) in the well. Each volume was measured three times. After fitting using $y = a_1 x$ (cyan line), the total intensity of beads in 1 μl of diluted solution was determined: $a_1 = 1.9 \pm 0.2 \times 10^7$ / μl. **(B)** The total fluorescent bead intensities in a solution depending on the number of volume units used in aspiration by the syringe pump. The original bead solution was diluted 1,000 times with Milli-Q water. The diluted bead solution was aspirated by the pump using 5, 8, 15, and 30 units from a glass bottom dish (D35-27-1.5-U, Matsunami) and deposited using 8, 12, 23, and 45 units, respectively, at the center of a well in a 96-well glass bottom plate, where 2 μl of Milli-Q water had been added in advance. The beads solution was dried and imaged as in (A). Each number of units was measured six times. **(C)** The volume of aspirated solution depending on the volume units. The volume of aspirated solution was calculated based on the parameter a_1. After fitting using $y = a_2 x$ (magenta line), one aspiration volume unit of the pump was determined: $a_2 = 0.57 \pm 0.05$ nl/unit.

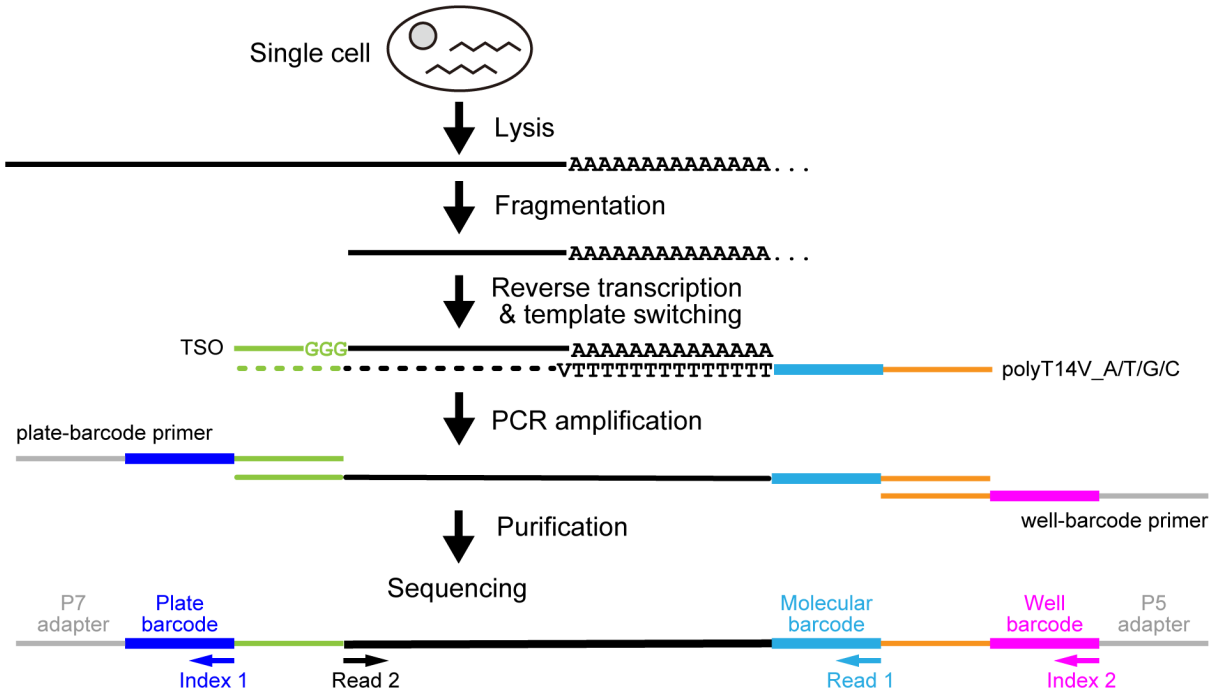


Figure S21: Schematic for library generation, purification, and sequencing in ALPS&RNA-seq (SI Appendix Text 4).

Sequences of polyT14V_A/T/G/C, TSO, well-position barcode primers, and plate barcode primers are listed in Dataset S13. Read 1, Read 2, Index 1, and Index 1 were sequenced in an Illumina platform; for Index 1 and Read 2, the customized sequencing primers I1_primer and R2_primer were used (Dataset S13).

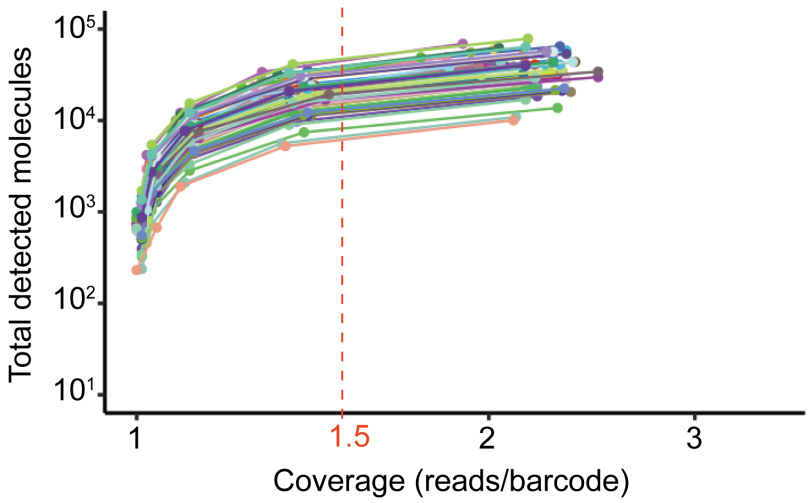


Figure S22: The total detected RNA molecules depending on the coverage (the average number of reads per unique molecular barcode).

To confirm that the number of reads per unique barcode was sufficient, we randomly sampled a fraction (1%, 3.3%, 10%, 33%, and 100%) of the reads in the results of the sequencing run for the cells from 3mix-ALPS-random&target-multimode and three-type-mixed cells isolated by FACS and counted the total number of detected molecules for each cell (each line). The results showed that the total number of detected molecules of each cell was basically saturated when the average number of reads per unique molecular barcode was more than 1.5. In our measurement, the average number of reads per unique barcode in each cell was mostly (99.6%) > 1.5. The results from 3mix-random-plate1 are shown here, and coverages for all cells are presented in Datasets S3 and S5.

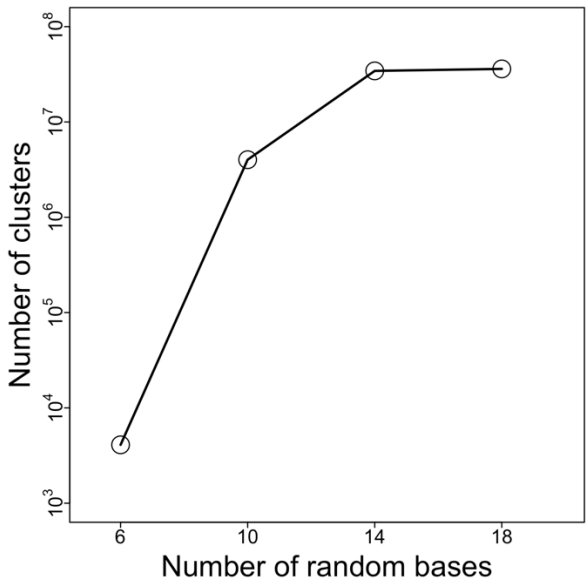


Figure S23: The number of clusters (unique molecular barcodes) depending on the number of random bases designed in the molecular barcodes.

We trimmed random bases in the sequenced barcodes of the HiSeq sequencing run (for the cells from 3mix-ALPS-random&target-multimode and three-type-mixed cells isolated by FACS) from the 3' end and counted the number of clusters as a function of the number of remaining random bases. The results showed that having more than 14 random bases did not further increase the number of clusters, indicating that the 18 random bases we used were sufficient for measuring approximately 4×10^7 molecules in a single HiSeq sequencing run.

Tables

Table S1: Comparison of the methods which performed whole transcriptome analysis on microscopically observed single live cells

(1) Methods with cell imaging in microwells

Literature	Method name	Live-cell label-free imaging with quantitative analysis	Single-cell whole transcriptome analysis	High-throughput data acquisition	Image-based deep learning based on live-cell images and single-cell whole transcriptome
K. Lane, et al., Cell Syst. 2017 (33)		∆ Eclipse Ti fluorescence microscope (Nikon) was used Cells were isolated (captured*) and imaged in microwells** (Fluidigm C1 chip)	○ Library preparation was performed using Fluidigm C1 chip	○ 96 cells can be measured in one Fluidigm C1 chip	Not performed
J. Yuan, et.al., Genome Biol 2018 (34)	SCOPE-Seq	∆ Eclipse Ti-U microscope (Nikon) was used Cells were isolated and imaged in PDMS*** microwells**(chip)	○ Library preparation was performed using chip and beads	○ 2,352 RNA-Seq expression profiles were measured using one microwell array device and 1,133 of them were linked to imaging data	Not performed
Z. Liu, et al., Sci. Rep. 2020 (35)	SCOPE-seq2	∆ Eclipse Ti2 microscope (Nikon) was used Cells were isolated and imaged in PDMS*** microwells**(chip)	○ Library preparation was performed using chip and beads	○ >9,000 cells were profiled in one microwell array device and > 4,100 of them were linked to imaging data	Not performed

○: Enough, ∆: Not enough

* The capture of Fluidigm C1 chip is cell size- and shape-dependent (36).

** It is often difficult to culture an isolated single cell in a microwell and to quantitatively analyze a bright-field cell image taken from a microwell. For example, the Fluidigm C1 chip paper listed in the table mentioned that Fluidigm C1 chip is not suitable for cells that cannot be cultured in isolation. Moreover, all three listed papers on SCOPE-Seq, SCOPE-seq2, and Fluidigm C1 chip did not show any quantitative analysis for bright-field cell images. The Fluidigm C1 chip paper also mentioned that they were unable to perform a fold-change detection of fluorescence in cell images.

*** Some studies mentioned that PDMS microwell prevents high resolution imaging (37, 38).

(2) Methods with cell imaging in dish

Literature	Method name	Live-cell label-free imaging with quantitative analysis	Single-cell whole transcriptome analysis	High-throughput data acquisition	Image-based deep learning based on live-cell images and single-cell whole transcriptome
M. Saint, et al., Nat. Microbiol. 2019 (39)		○ MSM-400 tetrad-dissection microscope (Singer Instrument) was used Cells were cultured and imaged in dish	○ Library preparation was performed in tube	△ Manual cell imaging and isolation were performed	Not performed
This study	ALPS& RNA-seq	○ Eclipse Ti2 microscope (Nikon) was used Cells were cultured and imaged in dish	○ Library preparation was performed in tube	○ Automated cell imaging and isolation were performed	Performed

Datasets

Dataset S1: The detailed failure cases for T cell isolations by ALPS with different cell densities.

Dataset S2: The success rate of each plate by APLS&RNA-seq for all three-type-mixed cells and PBMCs experiments.

Dataset S3: Total detected RNA molecules, total detected genes, sequencing coverage, cell type in three-types mixed cells experiments.
Cell_ID corresponds to the Cell_ID in Dataset S4.

Dataset S4: The number of detected RNA molecules for each gene of each cell in three-types mixed cells experiments.

Dataset S5: Total detected RNA molecules, total detected genes, sequencing coverage, cell type in PBMC experiments.
Cell_ID corresponds to the Cell_ID in Dataset S6.

Dataset S6: The number of detected RNA molecules for each gene of each cell in PBMC experiments.

Dataset S7: Morphological and dynamical features of the cells from 3mix-ALPS-timelapse obtained by NIS-elements.
Plate_ID and Well_ID_in_Plate correspond to the Plate_ID and Well_ID_in_Plate in Dataset S3.

Dataset S8: Morphological and dynamical features of the cells from 3mix-ALPS-timelapse obtained by CellProfiler 4.
Plate_ID and Well_ID_in_Plate correspond to the Plate_ID and Well_ID_in_Plate in Dataset S3.

Dataset S9: Morphological and dynamical features of the cells from 3mix-ALPS-timelapse obtained by TrackMate 7.
Plate_ID and Well_ID_in_Plate correspond to the Plate_ID and Well_ID_in_Plate in Dataset S3.

Dataset S10: Morphological and dynamical features of the PBMCs obtained by NIS-elements.
Plate_ID and Well_ID_in_Plate correspond to the Plate_ID and Well_ID_in_Plate in Dataset S5.

Dataset S11: Morphological and dynamical features of the PBMCs obtained by CellProfiler 4.
Plate_ID and Well_ID_in_Plate correspond to the Plate_ID and Well_ID_in_Plate in Dataset S5.

Dataset S12: Morphological and dynamical features of the PBMCs obtained by TrackMate 7.
Plate_ID and Well_ID_in_Plate correspond to the Plate_ID and Well_ID_in_Plate in Dataset S5.

Dataset S13: Sequences of primers used in all experiments.

Movies

Movie S1: A demonstration movie of automated single cell isolation using ALPS from top view.

Left side, the Nikon microscope; right side, the motorized stage for the displacement of a 96-well plate (Materials and Methods). Weak green light of the microscope: bright field illumination; blue light: GFP fluorescent illumination; strong green light of the microscope: PE fluorescent illumination; red light: Alx648 fluorescent illumination; green light on the stage, bright field illumination by LED. The movie was recorded from the step-(auto-2) to the step-(auto-7) with two rounds of isolation using the program used for multi-mode imaging and random selection (Materials and Methods) (acceleration and deceleration time of the arm was from the time-lapse experiment).

Movie S2: A movie of the picking and depositing by ALPS.

Left side, live imaging under the microscope; middle, current state (e.g., preview of captured image, current stage position) of the running program in NIS-Elements; right side, live imaging of a well of the 96-well plate. The movie was a clip from the step-(auto-2) of the 7^{th} to the step-(auto-7) of the 8^{th} of the T cell isolation (T-cell-plate2, Dataset S1). Red square, 139 μm × 99 μm.

Movie S3: A movie of picking and depositing a 1 μm bead using ALPS.

Left side, live imaging under the microscope; right side, live imaging of a well of the 96-well plate. Needle size, 40 μm. The deposited bead was around "2 o'clock".

Movie S4: A movie of picking and depositing a 3 μm bead using ALPS.

Left side, live imaging under the microscope; right side, live imaging of a well of the 96-well plate. Needle size, 40 μm.

Movie S5: A movie of picking and depositing a 45 μm bead using ALPS.

Left side, live imaging under the microscope; right side, live imaging of a well of the 96-well plate. Needle size, 75 μm.

Movie S6: A movie of picking and depositing a 350 μm bead using ALPS.

Left side, live imaging under the microscope; right side, live imaging of a well of the 96-well plate. Needle size, ~400 μm.

Movie S7: A movie of picking and depositing an intestinal organoid using ALPS.

Left side, live imaging under the microscope; right side, live imaging of a well of the 96-well plate. Needle size, ~400 μm.

Movie S8: A movie of picking and depositing a crypt using ALPS.

Left side, live imaging under the microscope; right side, live imaging of a well of the 96-well plate. Needle size, 40 μm.

SI References

1. P. D. Holler, L. K. Chlewicki, D. M. Kranz, TCRs with high affinity for foreign pMHC show self-reactivity. *Nat. Immunol.* **4**, 55–62 (2003).
2. K. Shimizu, *et al.*, PD-1 Imposes Qualitative Control of Cellular Transcriptomes in Response to T Cell Activation. *Mol. Cell* **77**, 937-950.e6 (2020).
3. T. Ikawa, H. Kawamoto, L. Y. T. Wright, C. Murre, Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. *Immunity* **20**, 349–360 (2004).
4. S. Sugimoto, T. Sato, Establishment of 3D Intestinal Organoid Cultures from Intestinal Stem Cells. *Methods Mol. Biol.* **1612**, 97–105 (2017).
5. Y. Li, *et al.*, A growth factor-free culture system underscores the coordination between Wnt and BMP signaling in Lgr5+ intestinal stem cell maintenance. *Cell Discov.* **4**, 49 (2018).
6. T. Ogawa, K. Kryukov, T. Imanishi, K. Shiroguchi, The efficacy and further functional advantages of random-base molecular barcodes for absolute and digital quantification of nucleic acid molecules. *Sci. Rep.* **7**, 13576 (2017).
7. M. C. Frith, A new repeat-masking method enables specific detection of homologous sequences. *Nucleic Acids Res.* **39**, e23 (2011).
8. A. Dobin, *et al.*, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
9. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
10. T. Stuart, *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
11. D. R. Stirling, *et al.*, CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 1–11 (2021).
12. D. Ershov, *et al.*, TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat. Methods* **19** (2022).
13. J. Schindelin, *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
14. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2016), pp. 770–778.
15. S. Hochreiter, J. Schmidhuber, Long Short-Term Memory. *Neural Comput.* **9**, 1735–1780 (1997).
16. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
17. G. I. Webb, *et al.*, "Leave-One-Out Cross-Validation" in *Encyclopedia of Machine Learning*, (Springer US, 2011), pp. 600–601.
18. J. L. Cardona, M. F. Howland, J. O. Dabiri, Seeing the wind: Visual wind speed prediction with a coupled convolutional and recurrent neural network. *Adv. Neural Inf. Process. Syst.* **32** (2019).
19. A. Paszke, *et al.*, "PyTorch: An Imperative Style, High-Performance Deep Learning Library" in *Advances in Neural Information Processing Systems 32*, H. Wallach, *et al.*, Eds. (Curran Associates, Inc., 2019), pp. 8024–8035.
20. S. J. Pan, Q. Yang, A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359 (2009).

21. M. Ranzato, F. J. Huang, Y.-L. Boureau, Y. LeCun, Unsupervised learning of invariant feature hierarchies with applications to object recognition in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, (2007), pp. 1–8.
22. F. Buggenthin, *et al.*, Prospective identification of hematopoietic lineage choice by deep learning. *Nat. Methods* **14**, 403–406 (2017).
23. F. Chollet, Keras. GitHub. https://github.com/fchollet/keras (2015).
24. V. Nair, G. E. Hinton, Rectified linear units improve restricted boltzmann machines in *Icml*, (2010).
25. S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift in *International Conference on Machine Learning*, (2015), pp. 448–456.
26. A. Mopin, V. Driss, C. Brinster, A detailed protocol for characterizing the murine C1498 cell line and its associated leukemia mouse model. *J. Vis. Exp.* **2016**, 1–12 (2016).
27. L. E. Sidney, M. J. Branch, S. E. Dunphy, H. S. Dua, A. Hopkinson, Concise review: Evidence for CD34 as a common marker for diverse progenitors. *Stem Cells* **32**, 1380–1389 (2014).
28. T. F. Tedder, C. M. Isaacs, Isolation of cDNAs encoding the CD19 antigen of human and mouse B lymphocytes. A new member of the immunoglobulin superfamily. *J. Immunol.* **143**, 712–7 (1989).
29. P. A. Szabo, *et al.*, Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10** (2019).
30. T. M. Gierahn, *et al.*, Seq-Well: Portable, low-cost rna sequencing of single cells at high throughput. *Nat. Methods* **14**, 395–398 (2017).
31. A. LeNail, NN-SVG: Publication-Ready Neural Network Architecture Schematics. *J. Open Source Softw.* **4**, 747 (2019).
32. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
33. K. Lane, *et al.*, Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-κB Activation. *Cell Syst.* **4**, 458-469.e5 (2017).
34. J. Yuan, J. Sheng, P. A. Sims, SCOPE-Seq: a scalable technology for linking live cell imaging and single-cell RNA sequencing. *Genome Biol* **19**, 227 (2018).
35. Z. Liu, *et al.*, Integrating single-cell RNA-seq and imaging with SCOPE-seq2. *Sci. Rep.* **10**, 1–15 (2020).
36. D. M. DeLaughter, The Use of the Fluidigm C1 for RNA Expression Analyses of Single Cells. *Curr. Protoc. Mol. Biol.* **122**, 139–148 (2018).
37. T. W. Frisk, M. A. Khorshidi, K. Guldevall, B. Vanherberghen, B. Önfelt, A silicon-glass microwell platform for high-resolution imaging and high-content screening with single cell resolution. *Biomed. Microdevices* **13**, 683–693 (2011).
38. M. Tonin, N. Descharmes, R. Houdré, Hybrid PDMS/glass microfluidics for high resolution imaging and application to sub-wavelength particle trapping. *Lab Chip* **16**, 465–470 (2016).
39. M. Saint, *et al.*, Single-cell imaging and RNA sequencing reveal patterns of gene expression heterogeneity during fission yeast growth and adaptation. *Nat. Microbiol.* **4**, 480–491 (2019).