Multi-dimensional morphology analysis enables identification and label-free enrichment of heterogeneous tumor cell populations

Andreja Jovic[1], Kiran Saini[1], Ryan Carelli[1], Tiffine Pham[1], Christian Corona[1], Jeanette Mei[1], Michael Phelan[1], Stephane C. Boutet[1], Kevin B Jacobs[1], Julie Kim[1], Manisha Ray[1], Chassidy Johnson[1], Nianzhen Li[1], Mahyar Salek[1], Matt Barnes[2], Cyril Y Ramathal[*2], Maddison Masaeli[1]

[1]Deepcell Inc. Menlo Park, CA; [2]Abbvie Inc, North Chicago, IL

KEY POINTS

- The Deepcell platform characterizes & sorts cells based on multi-dimensional morphology analysis without labels which alleviate the restriction and the limitation of using specific biomarkers.
- We demonstrate that multi-dimensional morphology can distinguish 1) parent vs chemo-resistant cell lines; 2) lung adenocarcinoma (LA) vs squamous cell carcinoma (SCC); and 3) heterogeneity within dissociated tumor cell (DTC) samples.
- Single cell analysis demonstrated that sorting based on morphology enriches for a transcriptional subpopulation of cancer cells.

ABSTRACT

- We performed multi-dimensional morphology analysis on cell images to phenotype and sort cells in real-time using sorting without labels.
- We trained a combined deep learning and morphometrics (computer vision) model using pure cell populations from dissociated tumor cells (DTC) prepared from NSCLC biopsies.
- Cell populations of interest were selected and sorted resulting in viable cells with minimal perturbation and compatible with downstream analysis.
- Our approach eliminates the need for cell staining and biomarker-specified isolation of tumor cells, and adds quantitative morphology as a biological readout.
- Future work: evaluate link between morphology, molecular, and functional traits of subpopulations.

METHODS

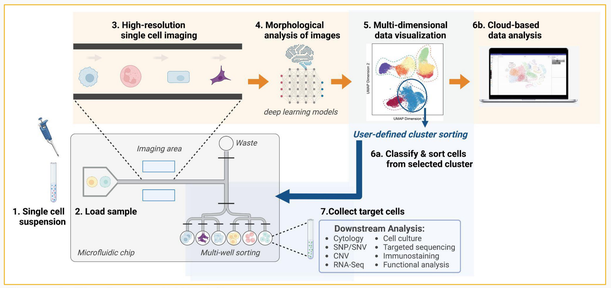


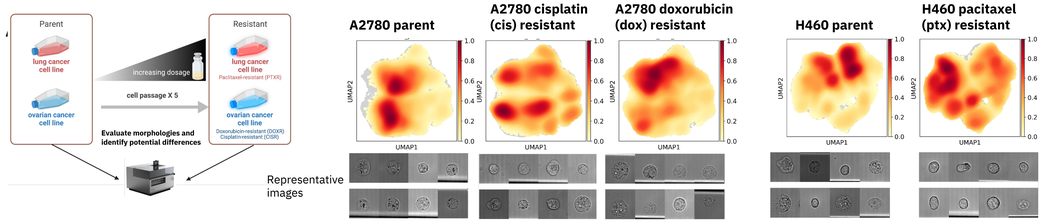
Figure 1. The Deepcell Workflow

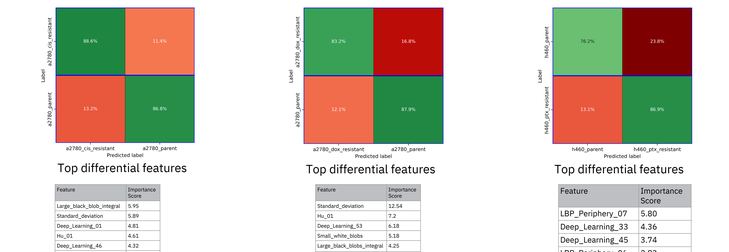
We analyzed each of the sample types using the Deepcell Workflow (**Figure 1**). **1)** Single cell suspension is **(2)** loaded onto a microfluidic chip. **(3)** Images of single cells are captured and analyzed in real-time by **(4)** deep learning and morphometric (computer vision) models to generate **(5)** multi-dimensional quantitative morphological profiles. User-defined cell clusters can be **(6a)** sorted for **(7)** downstream functional or molecular analysis. **(6b)** Morphology descriptions (embeddings) are extracted for data analysis/exploration and custom model training. The embeddings were used to profile the morphological heterogeneity of the cells, identify which morphology features differentiate each condition, and train a random forest classifier to predict each class of cells.

Learn more at Deepcell.com

RESULTS

Multi-dimensional morphology can distinguish between parent and chemo-resistant cell lines





Top differential features

Feature	Importance Score
Large_black_blob_integral	5.95
Standard_deviation	5.89
Deep_Learning_01	4.81
Hu_01	4.61
Deep_Learning_46	4.32
Deep_Learning_25	2.81
Deep_Learning_16	2.69
Deep_Learning_53	2.56
Small_white_blobs_integral	2.24
Deep_Learning_0	2.23
Deep_Learning_50	1.82
Small_white_blobs	1.76
Area	1.59

Top differential features

Feature	Importance Score
Standard_deviation	12.54
Hu_01	7.2
Deep_Learning_53	6.18
Small_white_blobs	5.18
Large_black_blobs_integral	4.25
Negative_fraction	3.51
Deep_Learning_16	3.42
Deep_Learning_40	3.05
LBP_center_07	2.59
Solidity	2.14
Percentile_25	1.84
LBP_center_06	1.70
Roundness	1.62

Top differential features

Feature	Importance Score
LBP_Periphery_07	5.80
Deep_Learning_33	4.36
Deep_Learning_45	3.74
LBP_Periphery_06	2.83
Deep_Learning_43	2.64
LBP_Periphery_03	2.39

Definition of morphometric features

Features describing **shape**
- **Area**: Cell area in microns
- **Roundness**: Similarity to a circle
- **Solidity**: ratio of the area to convex hull?

Features describing **pixel intensity**
- **Hu_01**: measure of pixel intensity
- **Large black blobs**: count of large black blobs
- **Large black blobs integrat:** integral over large black blob area
- **Negative fraction:** percent of pixels with a negative grayscale value
- **Percentile 75**: 75th percentile of grayscale values
- **Standard deviation**: standard deviation of grayscale values
- **Small white blobs integrat:** integral over white blobs
- **Small white blobs**: number of small white blobs

Features describing **texture**
- **LBP center**: describes texture in center of cell
- **LBP Periphery**: describes texture at edges of cell

Figure 3. Parent cell lines can be distinguished from drug-resistant cell lines based on morphology alone. (A) Chemical resistance was induced in two cell lines. Each cell line was imaged and analyzed on the Deepcell platform then projected onto a UMAP. **(B)** The density plot of each cell line on the UMAP projection shows distinct patterns for the parent and drug resistant cells for both A2780 and H460. Representative images for each cell line are shown. **(C)** A random forest classifier was trained to predict the cell type of each parent vs resistant combination, and the results are shown in the confusion matrices. The top differential morphology features separating each cell type show a mix of morphometric (computer vision) and Deep Learning features, including features observed by visual assessment such as the "black blobs" seen in A2780 parent cells and much less visible in the resistant cells.

Multi-dimensional morphology can distinguish lung adenocarcinoma from squamous cell carcinoma

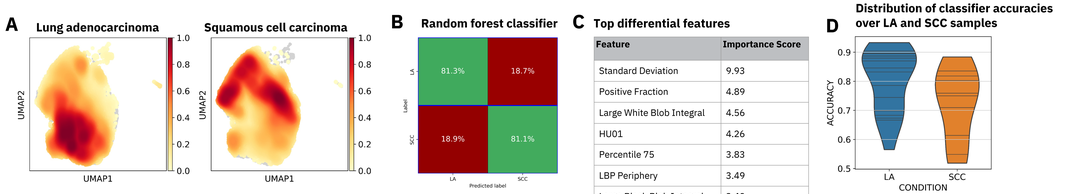
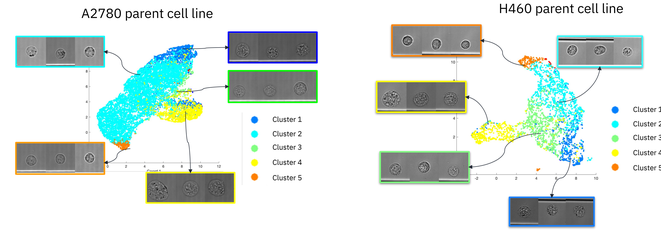


Figure 4. Multi-dimensional morphology profiling distinguishes different types of lung carcinomas. (A) Dissociated tumor biopsies from 17 lung adenocarcinomas (LA) and 12 squamous cell carcinoma (SCC) cells were run on the Deepcell platform, and UMAPs of the resulting analysis show distinct clusters for LA vs SCC morphologies. Representative images of cells from each cluster are shown. **(B)** A random forest classifier can predict LA vs SCC types with >80% accuracy, as shown by the confusion matrix. **(C)** The top morphological features distinguishing the two carcinomas are shown in the table. **(D)** The distribution of accuracies for each sample is shown in the violin plot, demonstrating that the 29 samples had a range of accuracies, but the majority are greater than 50% accurate to the predicted cell type.

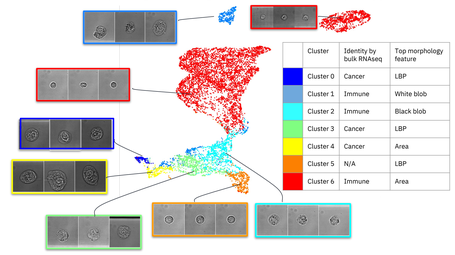
RESULTS cont.

Morphological profiling reveals heterogeneity within the cancer population of lung DTCs

A At least 5 distinct morphologies were observed in "pure" cell lines, revealing previously uncharacterized heterogeneity

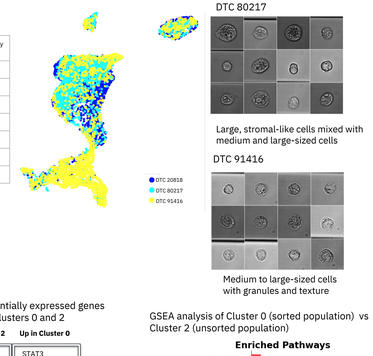


B UMAP of morphologies in lung DTCs colored by morphotype



Cluster	Identity by bulk RNAseq	Top morphology feature
Cluster 0	Cancer	LBP
Cluster 1	Immune	White blob
Cluster 2	Immune	Black blob
Cluster 3	Cancer	LBP
Cluster 4	Cancer	Area
Cluster 5	N/A	LBP
Cluster 6	Immune	Area

C UMAP of morphologies in lung DTCs colored by DTC sample



D Single-cell RNAseq identification of the sorted cell types in DTCs

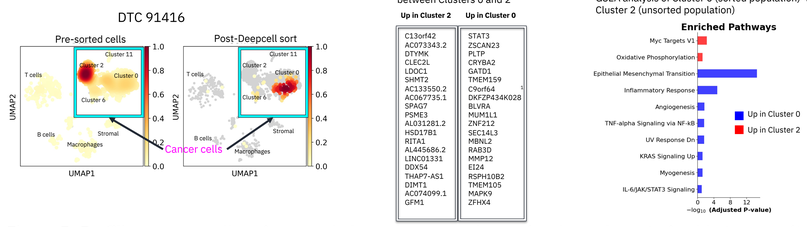


Figure 5. Profiling cancer cells using morphology alone reveals previously uncharacterized heterogeneity within lung DTC samples (A) Clustering using morphology alone demonstrated significant heterogeneity within the "pure" cell lines used for the drug resistance study, with at least 5 distinct morphologies observed in both A2780 and H460. We imaged and sorted three lung DTCs, covering a range of cancer cells (5%, 10%, and 90% cancer content) and included both LA and SCC cancers, and identified six morphologically distinct clusters **(B).** The six clusters are colored on the UMAP and representative images for each cluster are shown. **(C)** Morphology UMAP, colored by DTC sample, shows that samples contain different compositions of the same cell types. Examining images from the cluster containing cancer cells showed significant morphological heterogeneity. **(D)** Single-cell transcriptional analysis showed AI predicted cancer cells represent a subpopulation (Cluster 0) within the cancer cells present in the DTC (Cluster 2). **(E)** The top differential genes between Clusters 0 and 2 were assessed and used for GSEA. The enriched pathways demonstrate different cancer functions for each cluster.

CONCLUSIONS

- The Deepcell platform can distinguish between resistant and parent cell lines using morphology alone with 76-89% accuracy
- Morphology analysis uncovers previously unquantified heterogeneity within theoretically homogenous cell lines
- These results suggest that the Deepcell platform has the potential to sort resistant cells out of a mixed population or examine effects of drug treatment on specific subpopulations
- The method can accurately distinguish between lung adenocarcinoma and squamous cell carcinoma cells with 81% accuracy.
- We demonstrated the enrichment of malignant cells from NSCLC tumor tissue using morphology alone.
- Single-cell gene expression analysis showed distinct transcriptional profiles in cancer-related pathways in the subpopulation of Deepcell sorted cancer cells
- These results warrant further investigation of the potential link between morphology and functional phenotype uncovered in Deepcell sorting studies