Multi-modal Single-Cell Analysis: Self-supervised Vision Models Reveal Hidden Morphology-Transcriptome Relationships Across Diverse Biological Systems

Shreya S. Deshmukh[1], Mehdi Mohseni[1], Shan Sabri[1], Richard Yau[1], Yining Chen[2], Makenzie Sacca[1], Nirvan Rouzbeh[1], Olaia F. Villa[1], Gaetano Viscido[2], James Walker[3], Joseph Ecker[3], Andrea Califano[2], Pier Federico Gherardini[1], Gary P. Schroth[1]

[1]Cellanome, Foster City, CA, [2]CZ Biohub New York, New York, NY, [3]Salk Institute for Biological Sciences, Department of Plant Biology, La Jolla, CA, USA

INTRODUCTION Understanding cellular heterogeneity requires capturing not only gene expression but also the morphological context that reflects a cell's functional state. While single-cell transcriptomics has transformed our ability to profile gene expression at scale, it remains largely disconnected from the visual phenotype of individual cells. Cellanome's CellCage™ enclosure technology uniquely enables the simultaneous capture of high-resolution cell images and matched single-cell transcriptomes, providing a foundation for linking live cell behavior to function. Leveraging recent advances in computer vision, we applied DINOv2[1], a self-supervised vision transformer, to extract rich morphological embeddings from CellCage™ enclosure images. By integrating these features with single-cell transcriptomic data, we uncover hidden phenotypic states and gene expression programs that are invisible to either modality alone — revealing a new dimension of cellular heterogeneity across diverse biological systems.

Cellanome R3200: A Micro-3D Printing Platform That Unites Morphology and Transcriptomics

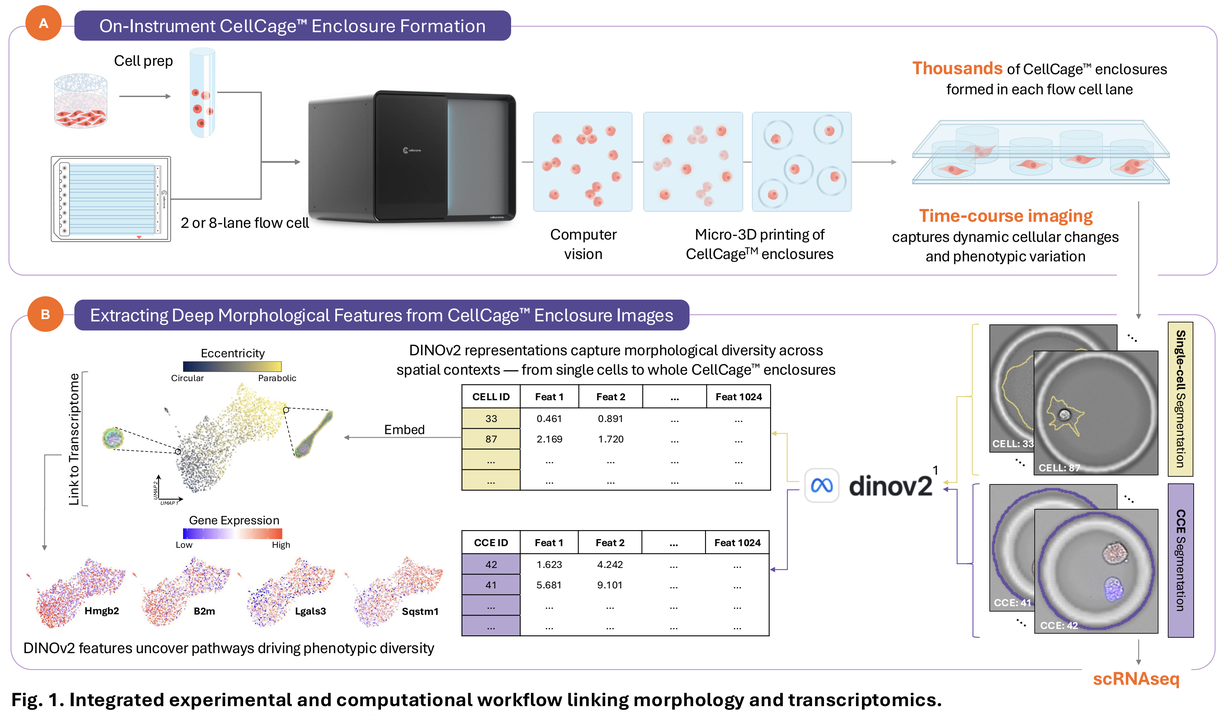

Fig. 1. Integrated experimental and computational workflow linking morphology and transcriptomics.
(A) Schematic of Cellanome's CellCage™ enclosure technology, which enables multi-modal phenotypic and functional profiling of the same single cells. Tens of thousands of suspended cells are mixed with a hydrogel precursor and loaded into a multi-lane flow cell. Individual cells are located, and CellCage™ enclosures (CCEs) are automatically generated around them via light-guided polymerization. The CCEs are permeable to small molecules and antibodies, allowing long-term culture and time-course imaging of the same cells. Following imaging-based assays, cells are lysed *in situ*, and mRNA is captured for cDNA synthesis, library preparation, and sequencing. **(B)** DINOv2-based computational workflow for extracting high-dimensional morphological embeddings from CCE images. These features capture phenotypic variation across scales — from single cells to entire CCEs — and can be integrated with single-cell transcriptomic data to reveal morphology–gene expression relationships.

DINOv2: A Self-Supervised Model for Extracting Morphological Features from Images

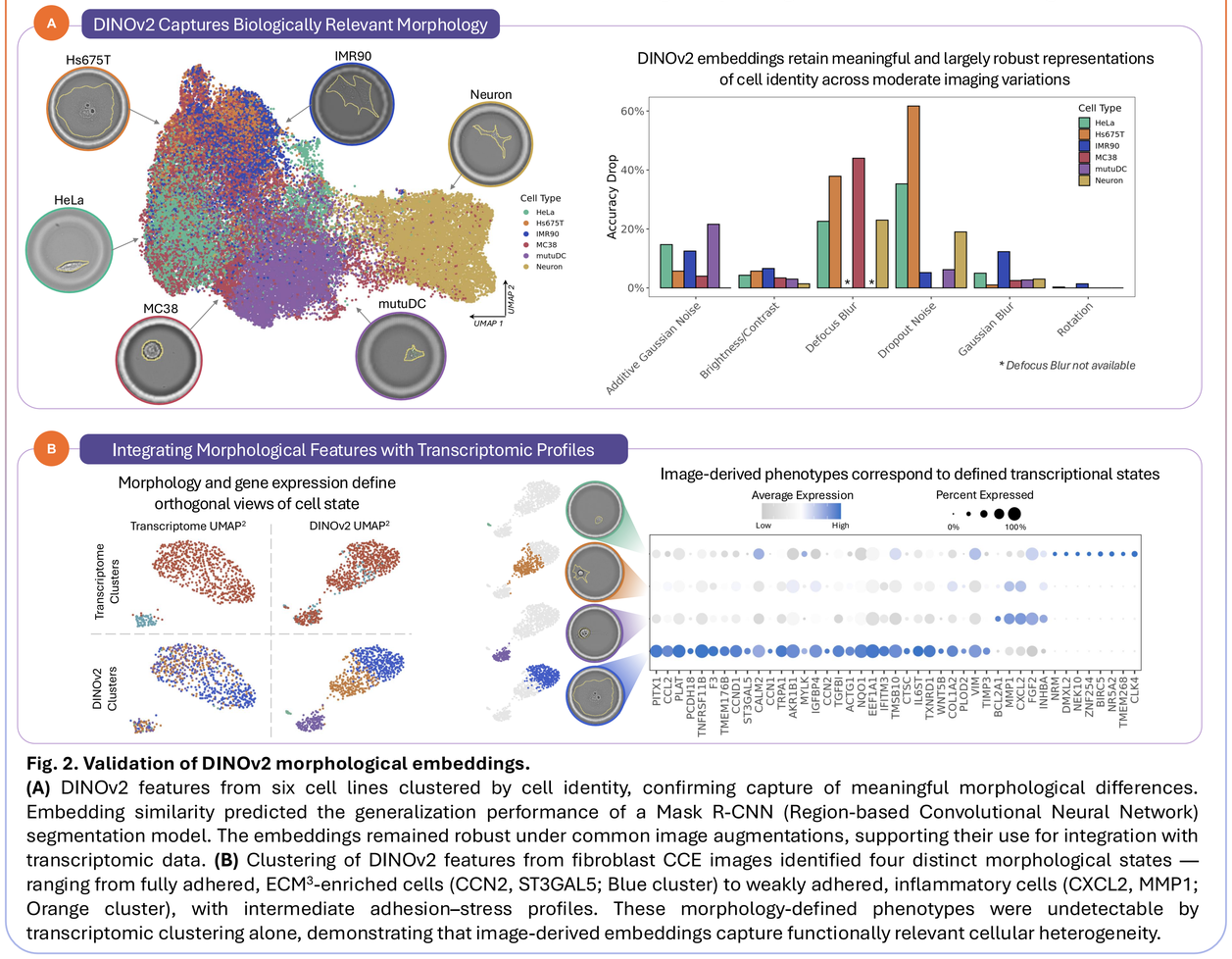

Fig. 2. Validation of DINOv2 morphological embeddings.
(A) DINOv2 features from six cell lines clustered by cell identity, confirming capture of meaningful morphological differences. Embedding similarity predicted the generalization performance of a Mask R-CNN (Region-based Convolutional Neural Network) segmentation model. The embeddings remained robust under common image augmentations, supporting their use for integration with transcriptomic data. **(B)** Clustering of DINOv2 features from fibroblast CCE images identified four distinct morphological states — ranging from fully adhered, ECM[3]-enriched cells (CCN2, ST3GAL5; Blue cluster) to weakly adhered, inflammatory cells (CXCL2, MMP1; Orange cluster), with intermediate adhesion–stress profiles. These morphology-defined phenotypes were undetectable by transcriptomic clustering alone, demonstrating that image-derived embeddings capture functionally relevant cellular heterogeneity.

Uncovering Morphology-Function Relationships Across Plant Protoplast Phenotypes

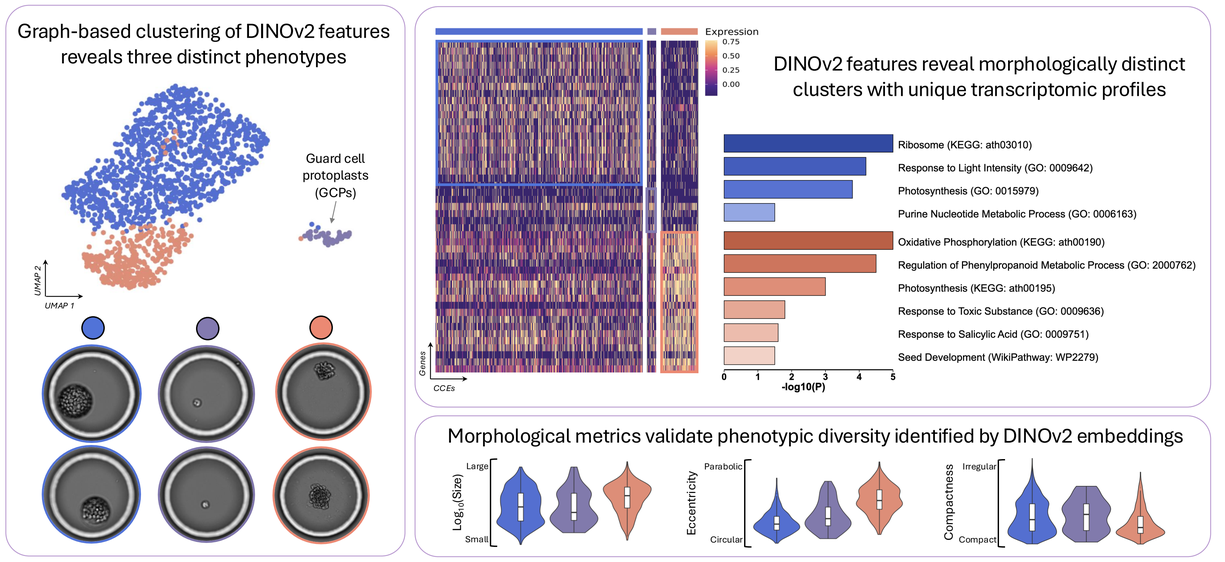

Fig. 3. DINOv2 embeddings reveal distinct morphological and functional states in plant protoplasts.
Clustering of DINOv2 features identified three populations: healthy, round cells enriched for photosynthesis genes; small guard cells associated with abscisic acid signaling; and stressed, parabolic-shaped cells with elevated oxidative phosphorylation.

Macrophage Polarization Spans a Transcriptomic and Morphological Continuum

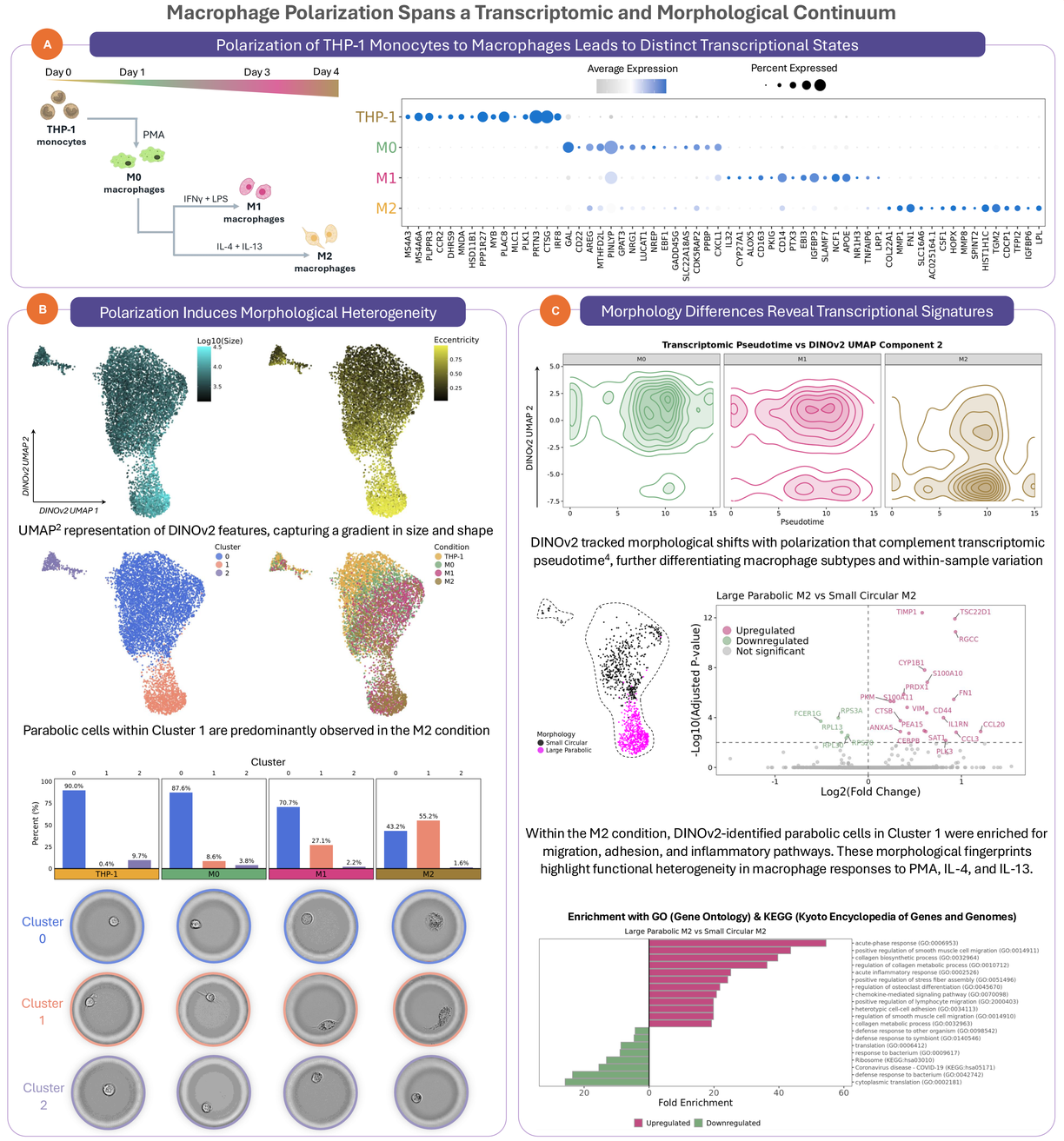

Fig 4. Morphological phenotypes add another dimension to transcriptomic profiling of macrophages.
DINOv2 detected multiple phenotypes that correlated with expected polarization pathways[5,6], revealing diverse cellular responses to the same treatment. Within the M2 condition, parabolic cells had a distinct transcriptional signature from circular cells. Thus, our DINOv2-enabled approach captures both overlap and divergence between morphology and transcriptomic features, combining both modalities to paint a fuller picture of cell biology.

QUESTIONS?

[1]**DINOv2**: Oquab *et al.* (2023)
[2]**UMAP**: Uniform Manifold Approximation & Projection; McInnes, Healy, Melville (2018)
[3]**ECM**: Extracellular Matrix
[4]**Monocle**: Cao *et al.* (2019)
[5]**Macrophage morphology**: Selig *et al.* (2024)
[6]**Macrophage morphology**: McWhorter *et al.* (2013)