

IMMUNOLOGY

Multiplexed high-throughput immune cell imaging reveals molecular health-associated phenotypes

Yannik Severin¹, Benjamin D. Hale¹, Julien Mena¹, David Goslings², Beat M. Frey², Berend Snijder^{1*}

Phenotypic plasticity is essential to the immune system, yet the factors that shape it are not fully understood. Here, we comprehensively analyze immune cell phenotypes including morphology across human cohorts by single-round multiplexed immunofluorescence, automated microscopy, and deep learning. Using the uncertainty of convolutional neural networks to cluster the phenotypes of eight distinct immune cell subsets, we find that the resulting maps are influenced by donor age, gender, and blood pressure, revealing distinct polarization and activation-associated phenotypes across immune cell classes. We further associate T cell morphology to transcriptional state based on their joint donor variability and validate an inflammation-associated polarized T cell morphology and an age-associated loss of mitochondria in CD4⁺ T cells. Together, we show that immune cell phenotypes reflect both molecular and personal health information, opening new perspectives into the deep immune phenotyping of individual people in health and disease.

INTRODUCTION

The morphology of a cell closely reflects its state, as it adapts to dynamic functional requirements and thereby constrains future behavior (1–4). This feedback mechanism has been shown to influence many cellular events, including cell differentiation (5, 6), cell division (2, 7, 8), adaptation to the microenvironment (9–11), and malignant transformation (12, 13). Few differentiated healthy human cells change their phenotype as markedly as immune cells: a plasticity that is critical to the correct function of the immune system as a whole (14–16). As a consequence, studying immune cellular heterogeneity at the molecular level has been transformative for our understanding of the immune system, measured, for example, by flow cytometry (17, 18), single-cell mass cytometry (19, 20), and single-cell RNA sequencing (RNA-seq) (21–25). Complementary to these molecular measurements, microscopy has shown the importance of immune cell morphology in multiple settings: Distinct cellular morphologies are associated with, and influence the outcome of, monocyte polarization (26, 27) and T and B cell activation (28–33), and label-free imaging of hematopoietic cells has enabled predicting the outcome of future lineage choices (34). In addition, a recent study, using organelle marker abundance as a proxy for cell morphology, found extensive evidence for morphological heterogeneity in both healthy and diseased immune cells (35). Because of their mixed adherent nature, however, primary immune cells such as peripheral blood mononuclear cells (PBMCs) were long considered incompatible with automated fluorescence microscopy, the tool of choice to characterize cellular morphology with spatial resolution across millions of cells (4, 9, 11, 36–39). This has hampered the comprehensive measurement and study of morphological heterogeneity present in the immune system and thus has left unanswered the question of which molecular and health factors globally shape the compendium of human immune cell morphologies.

¹Department of Biology, Institute of Molecular Systems Biology, ETH Zürich, 8049 Zürich, Switzerland. ²Blood Transfusion Service Zürich, SRC, 8952 Schlieren, Switzerland.

*Corresponding author. Email: snijder@immsb.biol.ethz.ch

RESULTS

Deep learning enables accurate eight-class cell type classification on multiplexed immunofluorescence and automated microscopy of human immune cells

To be able to comprehensively measure immune cell phenotypes, we developed a multiplexed immunofluorescence approach for PBMCs that extends our previously developed protocol for high-throughput image-based screening in human biopsies compatible with mixed nonadherent cells (Fig. 1) (40–42). In contrast to previously reported cyclical multiplexed immunofluorescence protocols (43–45), we stain once with a comprehensive immune cell marker panel that multiplexes eight surface markers and a nuclear dye, which is imaged by automated confocal microscopy and bright-field imaging in a single run (Fig. 1i and table S1). A deep convolutional neural network (CNN) (46) with custom architecture (fig. S1A) was subsequently used to classify each cell, making use of distinct marker expression patterns, lineage-specific labeling encoded by the staining panel, and likely differences in immune cell morphology (Fig. 1i). The CNN was trained across eight immune cell classes, using 89,483 manually curated five-channel subimages (four fluorescent channels and one bright-field channel) centered on individual cells sampled from 15 healthy donors (available at <https://doi.org/10.3929/ethz-b-000343106>). The eight immune classes capture the predominant immune lineages present in PBMCs, including three distinct T cell subsets (CD4⁺, CD8⁺, and CD4⁻CD8⁻), monocytes, dendritic cells, natural killer (NK) cells, B cells, and nucleated immune cells negative for all eight surface markers (Fig. 1).

CNN performance was stable across retraining, showed no sign of overfitting, and was 97% accurate for unseen donors systematically left out of the training data (fig. S1, B and C). The network further achieved 97.7% classification accuracy (Fig. 1ii) on a previously unseen test dataset of 24,000 curated cells comprising PBMCs from the same 15 healthy donors (fig. S1D). The classification efficiently demultiplexed mixed marker signals in the same fluorescence channel (fig. S1E), such that the resulting abundances of each subpopulation matched our expectations (fig. S1F). Both the class fractions (fig. S1G) and class probabilities (fig. S1H) showed good reproducibility over different experimental replicates [median correlation

Copyright © 2022
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
License 4.0 (CC BY).

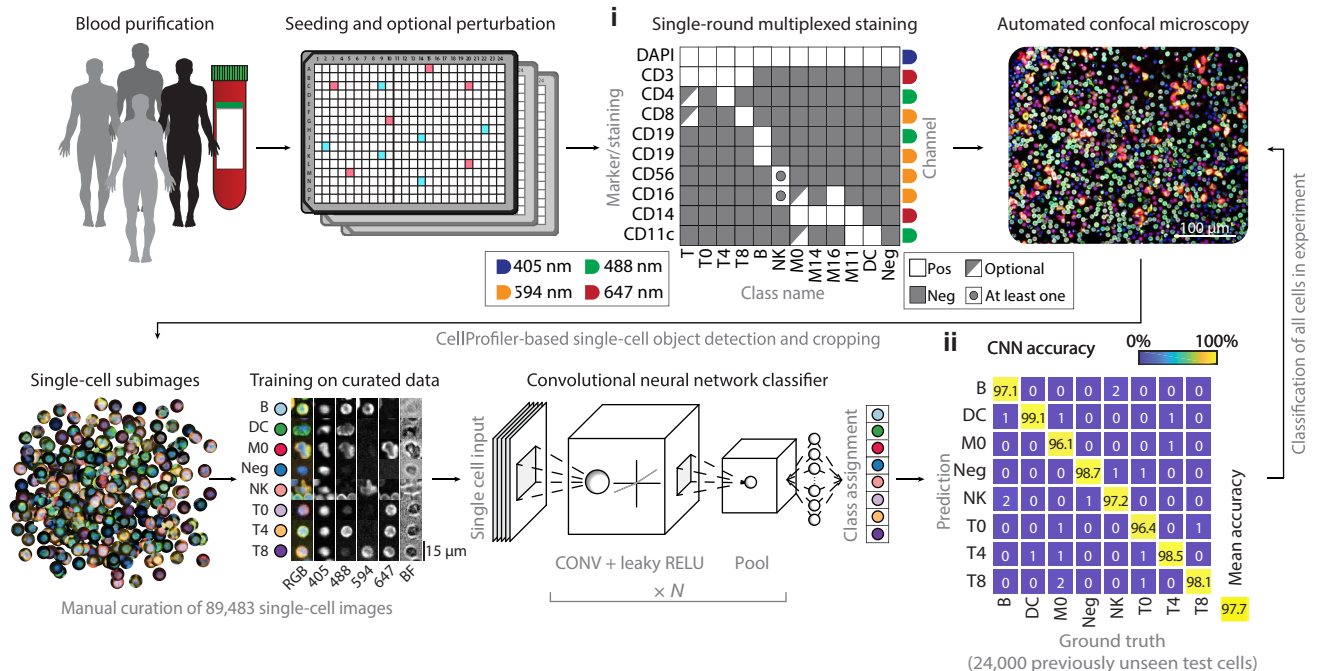


Fig. 1. Deep learning-enabled multiplexed immunofluorescence and microscopy of human immune cells. Workflow for the single-round multiplexed immunofluorescence, image-based screening, and associated deep learning-based classification of human PBMCs. PBMCs of healthy human donors are seeded in 384-well plates, optionally containing drugs or immune stimuli. Cells are fixed and stained with a comprehensive antibody panel (i) and imaged by automated confocal microscopy. A CNN is trained on 89,483 manually curated subimages to distinguish eight different immune cell classes and subsequently classifies all cells in the experiment. The curated test set contains 100 cells per class per donor per staining condition. (ii) Confusion matrix of CNN performance across all 24,000 cells that the CNN did not see before. CONV, convolution; RELU, rectified linear unit.

coefficient (r) = 0.90 and 0.95, respectively]. While marker expression likely contributed toward the accurate classification of morphologically similar classes (such as $CD4^+$ and $CD8^+$ T cells), cell morphology likely contributed to the separation of distinct cell types whose markers were multiplexed in the same channel. For example, $CD14^+$ monocytes and $CD3^+$ T cells, both stained in the allophycocyanin (APC) channel (fig. S1E), showed separation in cell size and staining pattern (fig. S1I). Supporting this interpretation, retraining the eight-class CNN without 4',6-diamidino-2-phenylindole (DAPI) and bright-field channels significantly reduced the classification accuracy ($P < 0.05$; fig. S1J); in addition, inversely, a two-class CNN could separate T cells and monocytes with 95% accuracy based on the DAPI and bright-field channels alone (fig. S1K). Last, we evaluated withholding all morphological information from a neural network classifier, by training an eight-class fully connected neural network classifier on the mean channel intensities per cell. This resulted in a classification accuracy of 86.2%, considerably lower than that of the CNN (fig. S1L). Thus, the eight-class CNN learned to generalize and leverage immune phenotypes across individual donors and experiments, presenting a robust, efficient, and data-rich high-throughput screening strategy with broad applicability.

Neural network uncertainty clusters immune cell phenotypes

Both supervised and unsupervised deep learning algorithms are increasingly used for image clustering (47, 48), which we here explored for the purpose of clustering immune cell phenotypes. The CNN returns a confidence vector for each cell that creates an eight-dimensional feature space, which we visualized by t -distributed

stochastic neighbor embedding (t -SNE) (Fig. 2A) (49). To minimize possible batch effects and confounding factors from ex vivo culturing, we analyzed a subset of 10 of the 15 donors on which the CNN was trained, whose blood had been simultaneously processed and incubated for just 1 hour before fixing and imaging across replicate wells and plates. Visualization of unperturbed immune cells from these 10 donors suggested considerable cell-to-cell variability, particularly among monocytes, even just within the cells classified with high CNN confidence (>0.7 class probability) (Fig. 2A). Projecting molecular and morphological cell features measured by conventional image analysis on the t -SNE revealed that the CNN had separated monocytes on the basis of their CD16 and CD11c expression levels, although it was not trained explicitly to do so (Fig. 2A, inset). Moreover, this showed that, even for high-confidence cells, the CNN class probabilities reflected marker expression and morphological heterogeneity for all eight immune cell classes, with nuclear size and bright-field intensity differences observed within each class (Fig. 2, A and B). Thus, while the eight-class CNN was strictly trained in a supervised manner, its neural network uncertainty additionally allowed further grouping of previously unannotated cellular phenotypes, capturing recurrent phenotypes present in primary human immune cells.

We next tested whether this deep learning uncertainty could also be used to quantify and categorize extrinsically induced changes in immune cell phenotypes. To this end, we stimulated PBMCs with nine immune modulators and three distinct controls ex vivo across concentrations and replicates, measuring 5 million multiplexed stained and imaged PBMCs (table S2). To avoid a possible bias from functional differences between immune cells from different donors, all cells were sampled from a single additional donor.

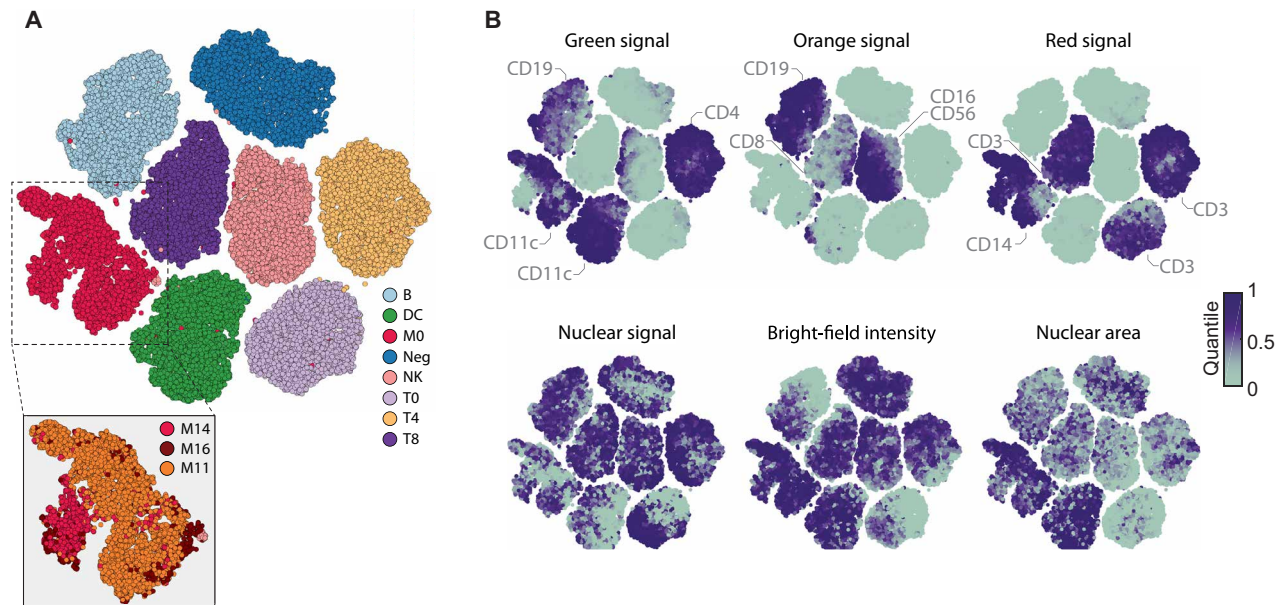


Fig. 2. Phenotypic landscape of the unperturbed immune system across 10 healthy donors. (A) *t*-SNE of the eight-class CNN probabilities of up to 1000 randomly subsampled high-confidence multiplexed cells per class and per donor (class probability > 0.7). Monocytes are further divided into three subpopulations by thresholding the immunofluorescence (IF) intensity of CD16 and CD11c stainings, respectively (inset). Figure depicts a total of 78,850 cells, randomly sampled from 40 wells for each of the 10 donors. All donors were processed and measured together in a single experiment, across 40 replicate wells per donor distributed over two 384-well plates. (B) Selected single-cell features projected onto the *t*-SNE shown in (A). Median value of overlapping data points is calculated, and color is assigned accordingly. Points are plotted in order of intensity, with the lowest intensity on top.

First, we visualized the structure in the CNN confidence by *t*-SNE (Fig. 3A), equally sampling cells from across all eight classes and 12 conditions. This revealed monocytes to be divided into three clusters associated with distinct CNN confidence profiles, not trivially explained by marker expression differences (fig. S2, A to C). To identify the contribution of distinct immune modulators to the morphological landscape of immune cells, we developed a method called *k*-nearest neighbor (KNN) local enrichment analysis (LEA) by hypergeometric testing (Fig. 3B and Material and Methods). For each cell, LEA identifies the nearest neighbors in the original eight-class probability space and calculates the hypergeometric significance of enrichment for cells with a certain property in this neighborhood. LEA next assigns this significance back to the original starting cell (see data file S1 for an example of this analysis). Projecting the LEA results back on the *t*-SNE revealed that the monocyte subcluster with the lowest CNN confidence was enriched for monocytes exposed to M1-type-inducing agents, *Escherichia coli* lipopolysaccharides (LPSs) and granulocyte-macrophage colony-stimulating factor (GM-CSF) (Fig. 3C and fig. S2D) (50), or to cytotoxic agents, causing the release of danger-associated molecular patterns. The second monocyte cluster was strongly enriched for cells exposed to M2-type-associated dexamethasone or interleukin-4 (IL-4), while the third and highest-confidence monocyte cluster was not enriched for most perturbations, thus likely reflecting unperturbed monocyte phenotypes (Fig. 3D).

Stimulation with microbial compounds such as LPS can selectively alter immune cell cross-talk, for example, through the induction of cell-cell contacts. We therefore suspected that phenotypes in the M1-type cluster could, in part, reflect changes in the multicellular context. To verify this, we performed spatially resolved single-cell analysis across the eight classified immune cell types, allowing the

high-throughput screening of 36 distinct immune cell-cell interactions simultaneously, a significant increase compared to our previous nonmultiplexed efforts (fig. S3, A and B) (40). Analysis of all 43 million cell-cell interactions measured in this experiment (fig. S3A) confirmed the M1-like monocyte cluster to be enriched for monocyte-to-monocyte interactions (fig. S3C). Thus, LPS-mediated monocyte activation led to distinct M1-like monocyte phenotypes, defined, in part, by an altered multicellular context. Collectively, LEA revealed that the uncertainty of the deep neural network reconstituted previously established monocyte M1/M2-type polarization phenotypes in a fully unsupervised manner (Fig. 3), while exposing considerably phenotypic complexity, with most immunomodulatory perturbations simultaneously affecting the phenotype of multiple immune cell class (Fig. 3A and fig. S2D).

Immune cell phenotypes associate with personal health information

The phenotypic heterogeneity of circulating immune cells captured by our image-based measurements could reflect both genetic and nongenetic influences (15, 51). To explore this, we analyzed commonalities and differences in the unperturbed immune phenotypes across the discovery cohort of the 10 donors shown in Fig. 2. We first used LEA to measure enrichment of cells from the same donor in the nearest-neighborhood in the eight-dimensional CNN class probability space. This identified distinct cellular phenotype regions significantly enriched for each of the 10 donors across several immune cell classes (Fig. 4A). As these enriched phenotypes were measured across technical repeats, they not only potentially indicated donor-individual characteristics of immune cell morphologies but could also reflect batch effects acting upstream of our sample processing and imaging. Repeating the analysis with randomized

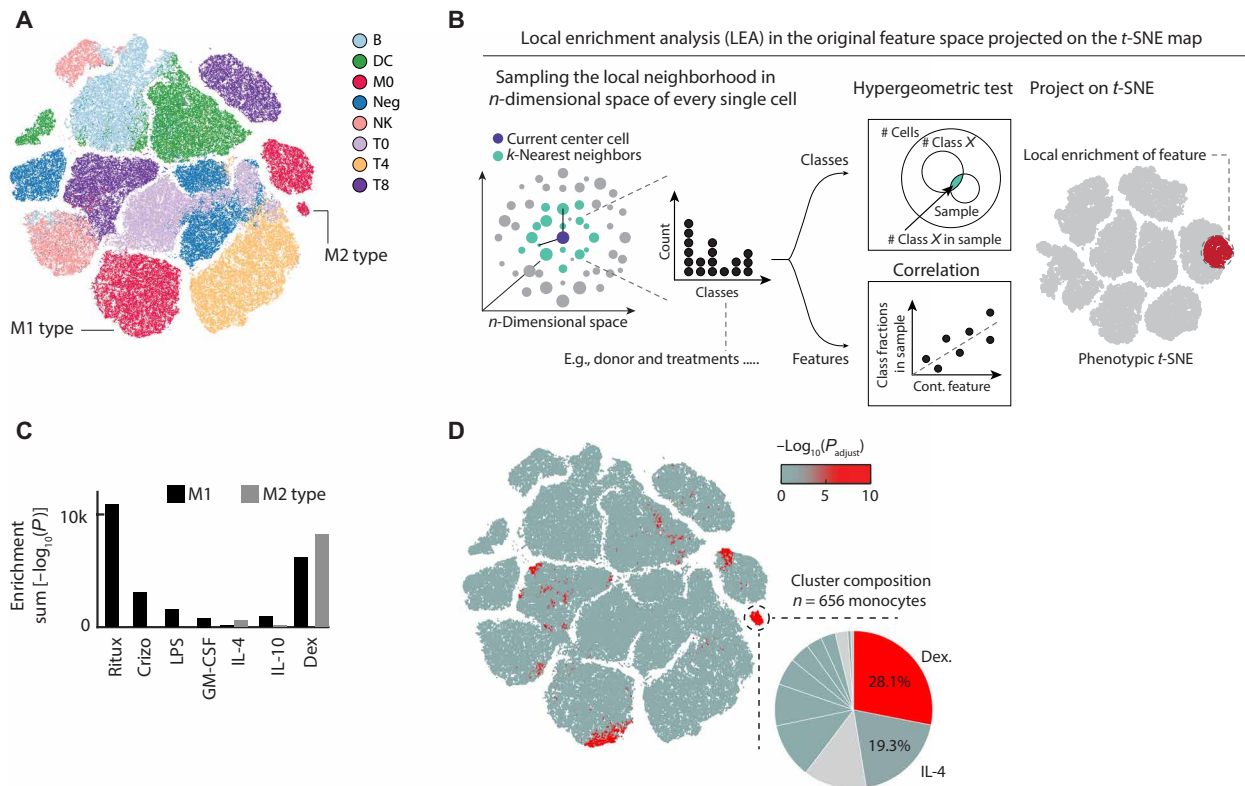


Fig. 3. Phenotypic landscape of a perturbed immune system and LEA analysis. (A) The CNN class probability *t*-SNE on the left shows 600 randomly chosen cells per class and condition; colored by class. (B) Overview of the LEA workflow. LEA probes the KNNs of each single cell in a multidimensional space for enrichment of either continuous or discrete features. For discrete features, the probability of finding *n* cells of condition *X* in the probed neighborhood follows a hypergeometric distribution, from which an enrichment *P* value is calculated. For continuous features, the relative fraction of cells of each donor in the probed local neighborhood is calculated. These fractions are then rank-correlated with a continuous feature that was measured across donors, and the *P* value of the correlation is calculated. In both cases, the enrichment probability is assigned to the center cell, and the approach is iterated for each single cell in the analysis. (C) Bar graph depicting the sum total $\log_{10}(\text{LEA } P \text{ values})$ for selected perturbations in the M1 (black) and M2 (gray) monocyte clusters. (D) LEA reveals regions in the phenotypic space that are significantly enriched for dexamethasone-treated PBMCs. Cells in the *t*-SNE are colored by their enrichment significance of the LEA run $[-\log_{10}(P_{\text{adjusted}})]$; see color bar. Inset highlights the contribution of different perturbations to the selected M2-type monocyte cluster. Figure depicts a total of 199,375 cells, randomly sampled from across 240 wells for a single donor.

donor labels and comparing the sum of enrichments showed that the actual donor enrichment in nearest neighbors of the latent space was well above what would be expected by random ($P < 1.1 \times 10^{-308}$; Fig. 4A, inset). We next looked for phenotypes that were enriched in donors with the same biological gender, with the 10 donors including 4 women and 6 men. This revealed strong gender associations with various immune cell morphologies ($P < 1.1 \times 10^{-308}$; Fig. 4B), with NK and negative cell class phenotypes particularly enriched in female donors and not explained by enrichment in any individual female donor (fig. S4A).

We next explored immune phenotype associations with continuous health parameters such as donor age, which has been described to markedly alter the immune phenotypic landscape (Figs. 3B and 4C and Material and Methods) (52, 53). A modification of LEA for continuous variables calculates the significance of the rank correlation between the fraction of cells per donor in the nearest neighborhood and any continuous variable of each donor (Fig. 3B). As before, the LEA analysis was run in the eight-dimensional CNN class probability space. To correct for spurious associations, we compare the association strength with those observed in many repeats with the same health parameter randomized across the donors. Testing donor age, height, weight, body mass index, blood pressure, and hemoglobin

levels revealed significant associations with donor age ($P < 1.3 \times 10^{-9}$) and systolic blood pressure ($P < 4.5 \times 10^{-4}$; Fig. 4C) but not to any of the other measured health parameters. The age-associated phenotype map revealed bimodal age associations for several immune subpopulations, particularly notable for CD4^+ T cells (Fig. 4D). Across the cells that make up the phenotype map, the age associations were mutually exclusive of the single-donor enrichments ($r = -0.002$; fig. S4B).

Donor variability allows to link T cell phenotypes with bulk gene expression data

To investigate the above identified phenotypic and health associations, we next used LEA to associate molecular pathway expression as measured by transcriptomics with immune cell phenotypes. Focusing on T cells, we performed bulk RNA-seq of CD3^+ positive cells isolated from the same 10 healthy donor blood samples, detecting on average around 15,000 expressed transcripts (fig. S5A). To associate bulk transcript measurements with single-cell imaging data, we first randomly subsampled 10,000 imaged T cells per donor, irrespective of their subpopulations (T0, T4, and T8). This random subsampling was performed to reflect the composition of isolated bulk T cells on which RNA-seq was performed. We then correlated

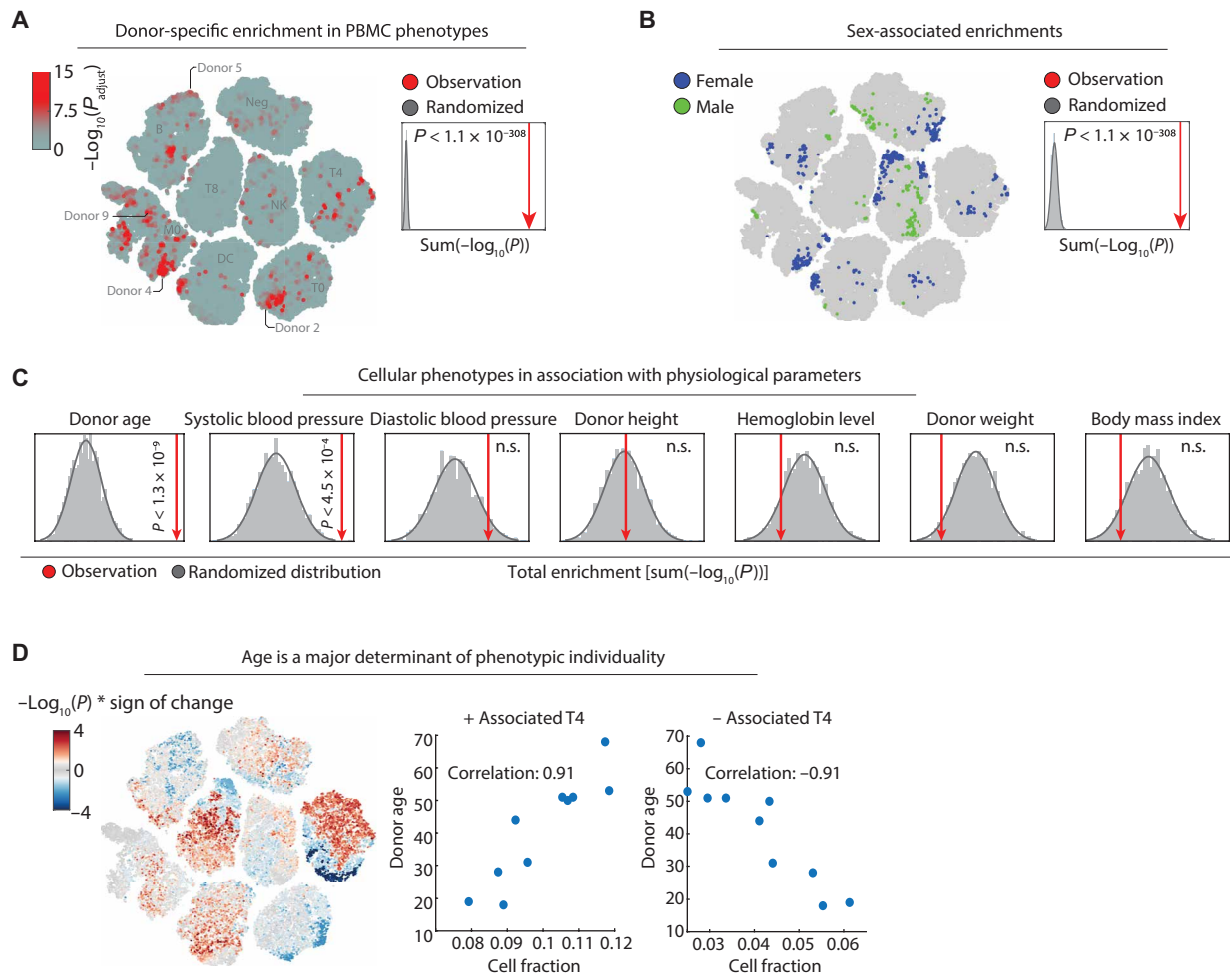


Fig. 4. LEA associates immune cell phenotypes with personal health information. (A) LEA of donor-specific cells across 10 donors, visualized on the *t*-SNE of Fig. 2A. The cells are colored by their maximum LEA significance across the 10 donors $[-\log_{10}(P_{\text{adjusted}})]$; see color bar. Inset: A randomized null distribution of donor enrichments was generated by randomizing the donor labels 2000 times and summing up all single-cell enrichments calculated by LEA per randomized run (gray bars). Sum enrichment of the actual data is shown in red, and the significance compared to the randomized runs is calculated by a one-sided *t* test. (B) LEA of biological gender-specific phenotypes projected onto the *t*-SNE. Cells are colored by their significant enrichment in female (blue) or male (green) specific phenotypes. Inset: A null distribution of random gender enrichment (gray) was generated by randomizing the donor labels 2000 times and summing up all single-cell enrichments calculated by LEA. Sum enrichment of the actual data is shown in red, as in Fig. 3A. (C) Association analysis of various health parameters with cellular phenotypes calculated by LEA. Null distributions of random correlation significance (gray) were generated by randomizing the donor labels 2000 times and summing up the all single-cell enrichments calculated by LEA per randomized run. Enrichment of the actual data is shown in red (one-sided *t* test). n.s., not significant. (D) LEA age associations projected onto the *t*-SNE. Single cells are colored by their signed significance of correlation $[-\log_{10}(P) \times \text{sign of the correlation}]$; see color bar. Inset: Fraction of all significantly positive and negative age-associated CD4⁺ T cells with donor age ($P < 0.05$).

the local phenotype abundance in the eight-dimensional CNN class probability space with the bulk transcript measurement using LEA across donors (Fig. 5A). To benchmark these phenotype-to-transcriptome associations, we first compared the LEA associations of *CD4* and *CD8A* transcript abundance (Fig. 5A) with the CD4 and CD8 protein expression levels explicitly measured by immunofluorescence for each T cell (fig. S5B). Validating the approach, LEA achieved excellent results for these proof-of-concept benchmarks, with areas under the receiver operating curve of 0.93 and 0.89 for CD4- and CD8-positive cells, respectively (Fig. 5B).

We next sought to validate these pathway-phenotype associations by querying the associations the other way round: starting from well-known pathways and seeing what phenotypes are associated with it. To this end, we inspected the associations with the T cell

receptor (TCR) signaling pathway as proxy for T cell activation. TCR signaling was strongly associated with distinct subregions of the phenotype map, including the cluster periphery of CD8⁺ T cells (Fig. 5C). This pattern was recapitulated by the LEA associations with *MAPK1* (*ERK2*), part of the TCR-induced signaling cascade, which largely, but not exclusively, overlapped with regions enriched for cells from donor 2 (Fig. 5C). Visual inspection of cells residing in TCR signaling and *MAPK1*-associated phenotypic regions revealed a notable polarized and activated T cell morphology, henceforth referred to as T_{ACT} cells. In contrast, randomly sampled cells from adjacent and nonenriched regions contained conventional small and round T cell morphologies, which we refer to as T_{CON} cells (Fig. 5D). To robustly quantify the T_{ACT} morphology further, we trained a dedicated CNN on manually curated T_{ACT} and T_{CON} phenotypes,

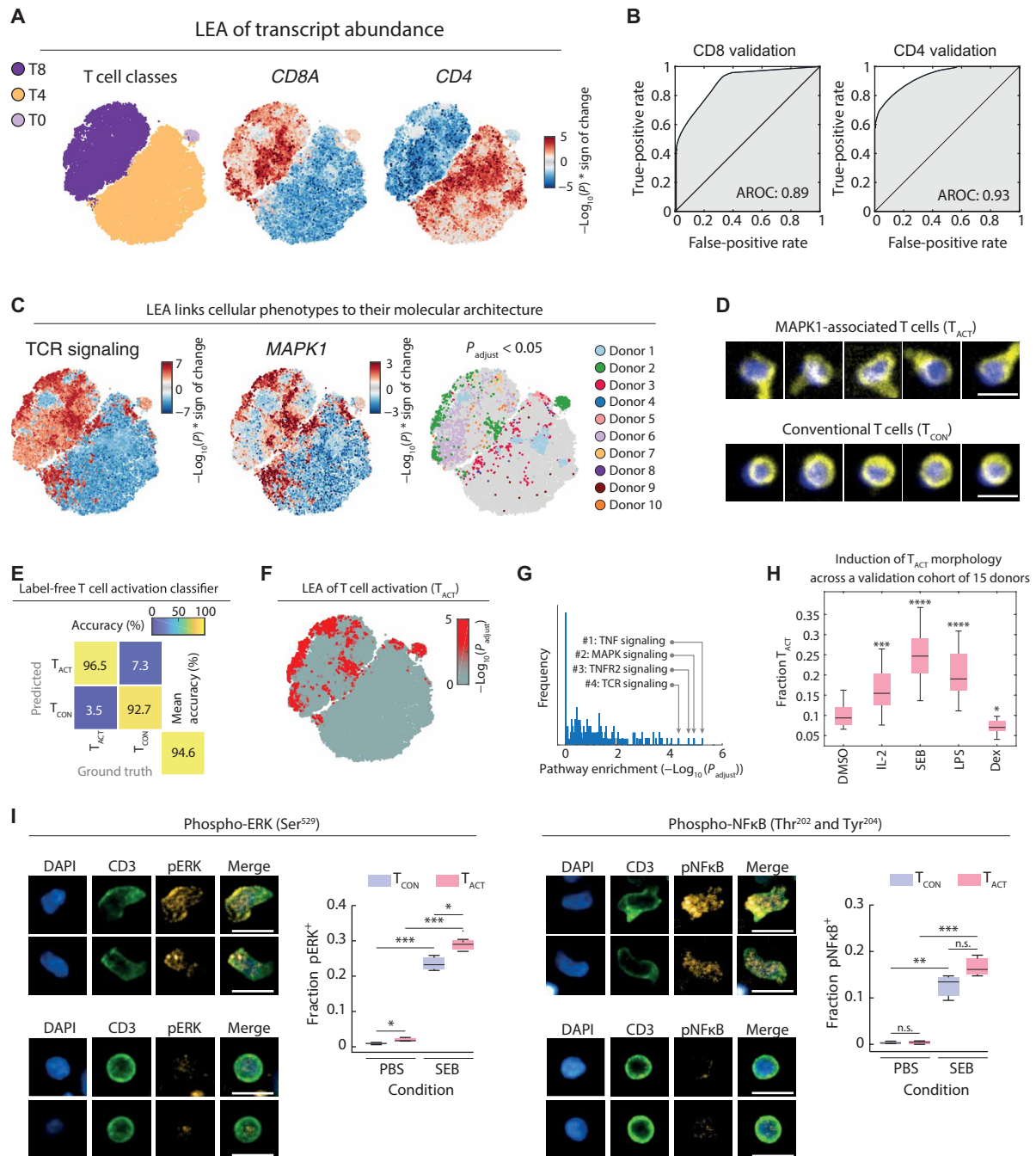


Fig. 5. Discovery and validation of an activated T cell morphology. (A) *t*-SNE of the CNN class probabilities of T cells from 10 healthy donors. A total of 10,000 T cells were randomly selected across all T cell populations per donor (without confidence threshold) to reflect their original abundance in the respective sample. *t*-SNE colored per T cell class (left), by LEA-based associations with *CD8A* transcript abundance (middle), or *CD4* transcript abundance (right) [$-\log_{10}(P) * \text{sign of change}$]. (B) Area under the receiver operating characteristic curve (AROC) for the consistency between the *CD8A* LEA associations with CD8 levels by immunofluorescence (left) and the *CD4* LEA associations with CD4 levels by immunofluorescence (right). (C) LEA associations of TCR signaling (left), *MAPK1* transcript abundance (middle), and donor enriched regions (right; $P_{\text{adjust}} < 0.05$) projected on the *t*-SNE map. (D) Representative CD8⁺ T cells from the positively associated *MAPK1*-associated regions (T_{ACT}) and conventional CD8⁺ T cells of other regions (T_{CON}). Yellow, CD3; blue, DAPI. (E) Confusion matrix of the T_{ACT} classifier. Test set comprises 369 random T_{ACT} cells and 738 random T_{CON} cells across multiple donors (including the 10 depicted donors). (F) LEA of the T_{ACT} phenotype projected on the *t*-SNE map. (G) Distribution of pathway significance across all retroactively classified T_{ACT} morphologies. Pathway enrichments were calculated using a hypergeometric test on positively associated genes (0.95 percentile), and *P* values were corrected for multiple testing. (H) Induction and suppression of the T_{ACT} phenotype with immunomodulatory agents across an independent validation cohort of 15 individual donors. Compounds were screened at 100 ng/ml. Box plots show the mean relative fraction of T_{ACT} of all T cells across all wells of each condition per donor. (I) Immunofluorescence quantification of phospho-nuclear factor κB (pNF κB) and phospho-extracellular signal-regulated kinase (pERK) levels in T_{ACT} and T_{CON} cells. Box plots show the fraction of phospho-signaling positive T_{ACT} (red) and T_{CON} (blue) cells after 48 hours of incubation with SEB or control ($n = 3$). Representative T_{ACT} and T_{CON} cell morphologies shown. Scale bars, 10 μm . * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

which achieved 94.6% validation accuracy on images from donors and experiments that it was not trained on (Fig. 5E and fig. S5C). This allowed us to retroactively detect the T_{ACT} morphology for all imaged T cells, which confirmed that the phenotype was present in all donors and most enriched in the cells of donor 2 (Fig. 5F and fig. S5D). Coming full circle, the T_{ACT} enriched regions associated with tumor necrosis factor (TNF) and mitogen-activated protein kinase (MAPK) signaling as most enriched pathways after multiple testing correction (Fig. 5G).

Validation of the inflammation-associated T cell morphology (T_{ACT}) in an independent donor cohort

To confirm that the T_{ACT} morphology is associated with inflammation and T cell activation in an independent validation cohort, we stimulated PBMCs derived from 15 additional healthy donors with proinflammatory cytokine IL-2, superantigen *Staphylococcus aureus* enterotoxin B (SEB), or LPS, which all led to significant increases in the fraction of T cells adopting a T_{ACT} morphology (Fig. 5H and fig. S5E). Exposure to the anti-inflammatory synthetic glucocorticoid dexamethasone, in contrast, reduced the relative abundance of T_{ACT} cells across the 15 donors (Fig. 5H and fig. S5E). To rule out the possibility that the T_{ACT} morphology was induced by cellular fixation before imaging, we further conducted live cell imaging of SEB-stimulated PBMCs and visually confirmed the induction of the T_{ACT} cell phenotype (fig. S5F). We next measured by immunofluorescence the levels of phosphorylated nuclear factor κ B (NF κ B) (Ser⁵²⁹) and extracellular signal-regulated kinase (ERK) (Thr²⁰² and Tyr²⁰⁴) as a function of T cell morphology, at baseline and upon SEB stimulation in PBMCs. At baseline, T_{ACT} cells showed slightly but significantly higher levels of phosphorylated ERK. SEB stimulation increased phosphorylated levels of ERK significantly higher in T_{ACT} than T_{CON} cells. Together, these results experimentally validated the LEA-based pathway enrichment analysis with the polarized T_{ACT} morphology. Thus, part of the donor unique fingerprints that we previously observed had resulted from differences in T cell activation between the donors, with 15% of T cells from donor 2 adopting the T_{ACT} morphology, predominantly in CD8⁺ T cell compartment, while on the other end of the spectrum, only 7% of donor 1 T cells were T_{ACT} cells, here mostly in CD4⁺ T cells (fig. S5D).

Deep learning uncertainty reveals an age-associated mitochondrial decline in CD4⁺ T cells

Having validated the phenotype-to-pathway association approach and its ability to find and correctly describe new cellular phenotypes, we explored the pathway enrichments for age-associated T cell phenotypes (Fig. 6A and fig. S6A). Pathways enriched in phenotypes that were reduced with age included nucleotide excision repair, telomere maintenance (54), cilia assembly (55), and propanoate metabolism (fig. S6A). In contrast, pathways associated with T cell phenotypes that increased with age included inflammation- and stress-related pathways, particularly for the CD8⁺ compartment, and lysosome- and vesicle-associated pathways in CD4⁺ T cells (Fig. 6A, right). Inflammation is a well-described risk factor for age-associated diseases (56), and, consistently, the age-associated phenotypes overlapped partially with the above validated phenotype for activated CD8⁺ T cells (Fig. 6A, right). Furthermore, impaired organelle and lysosome homeostasis in aged CD4⁺ T cells has been previously described as a relevant process in aging of T cells (57).

Pathway enrichments for oxidative phosphorylation and mitochondrial respiration in age-associated T cell phenotypes were in line with reports of defective respiration in CD4⁺ T cells of aged mice (58, 59) and suggested that the neural network might have identified a phenotypic T cell signature associated with both donor age and mitochondrial abundance. The CD4⁺ T cells showed strong bright-field intensity differences, a measure of intracellular granularity (Fig. 2, A and B, and fig. S5B). This bright-field trend followed the age associations that we observed, with CD4⁺ T cells enriched in younger people measured to be more granular (referred to as T_{4BFD} for “bright-field dark” CD4⁺ T cells; Figs. 2, A and B, and 4C). Quantifying this association across all subpopulations, CD4⁺ T cells showed the most significant age-associated bright-field intensity differences ($P < 10^{-70}$), followed by the CD8⁺ T cells ($P < 10^{-40}$), and less for the other immune cell classes (Fig. 6B).

To reproduce this association, we sampled an additional validation cohort of 15 healthy donors (Fig. 6C) and trained a different neural network architecture on a new set of images generated only from this validation cohort (fig. S6B). This independent repetition of the workflow revealed that the age-associated T_{4BFD} phenotype was independent of the donor cohort and neural network and experimental batch (Fig. 6C and fig. S6B). The age-associated bright-field intensity differences and mitochondrial pathway association might reflect loss of mitochondrial abundance in age in CD4⁺ T cells (60). To support this interpretation, we analyzed whether bright-field intensity reflects mitochondrial abundance using the natural heterogeneity observed within CD4⁺ T cells of a single donor (Fig. 6D). Those cells that were darkest by bright-field imaging displayed significantly higher mitochondrial abundance as measured by image-based quantification of the MitoTracker dye (Fig. 6D). The deep learning uncertainty thus had revealed a label-free phenotype reflecting an age-associated mitochondrial decline in CD4⁺ T cells, explaining, in part, how immune cell phenotypes measured by our high-throughput single-cell imaging pipeline capture donor information such as age.

DISCUSSION

We here explore the molecular health determinants of human immune cell phenotypes using a workflow that combines automated high-throughput microscopy, single-round multiplexed immunofluorescence, and deep learning-based phenotypic analysis. The presented method for phenotyping of immune cells distinguishes itself for its ability to integrate cell morphology, protein levels and localization, and multicellular context into a quantitative metric across eight major immune cell classes, hundreds of conditions, and millions of cells. The resulting single-cell phenotype space, derived from the CNN's uncertainty, reflected both genetic and nongenetic donor health information. We find age, gender, blood pressure, and inflammatory state to be significantly associated with human immune cell phenotypes, yet many more influences likely exist and more phenotype associations captured by our approach remain unexplored.

Our workflow is tailored to make use of two large sources of biological heterogeneity: the heterogeneity observed between individuals, and heterogeneity observed within cells of the same class and donor. That dependency, however, is at the same time its limitation: The single-round multiplexed staining strategy benefits from the presence of multiple cell types with variable cell morphologies and

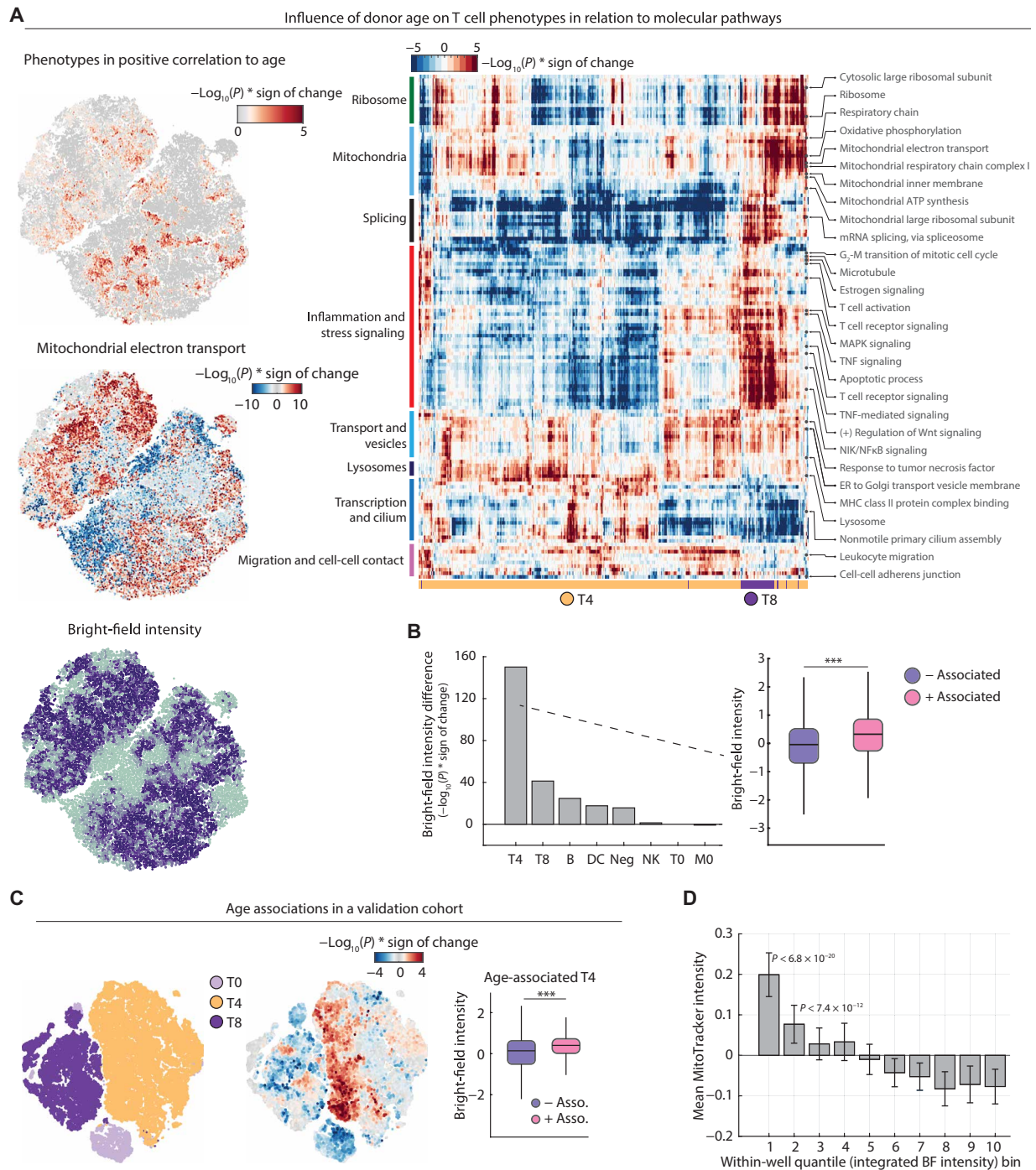


Fig. 6. Discovery and validation of an age-associated mitochondrial decline in CD4⁺ T cells. (A) Top left: Positive LEA associations with donor age projected on the *t*-SNE as in Fig. 5A [colored by $-\log_{10}(P)$]. Middle left: LEA associations of mitochondrial electron transport projected on the *t*-SNE. Bottom left: Bright-field single-cell intensity projected onto the *t*-SNE. Median value of overlapping data points is calculated, and color is assigned accordingly. Points are plotted in order of intensity, with the lowest intensity on top. Right: Heatmap overview of all significantly enriched pathways in positive age-associated T cells [$-\log_{10}(P) > 5$]. ATP, adenosine 5'-triphosphate; NIK, NFκB-inducing kinase; ER, endoplasmic reticulum; MHC, major histocompatibility complex. (B) Comparison of the difference in bright-field intensity of positively versus negatively associated immune cells per population [with an association cutoff of $-\log_{10}(P) > 1.3$]. Bar plots show the $-\log_{10}(P)$ * sign of the change $\{1 - [\text{median}(\text{positive enrichment}) / \text{median}(\text{negative enrichment})]\}$. (C) Negative and positive age associations with cellular T cell phenotypes and donor age in an independent validation cohort of 15 healthy individuals calculated by LEA (left and middle). *t*-SNE depicts a total of 5000 cells per donor. Right: Comparison of differences in bright-field intensity of positively versus negatively age-associated CD4⁺ T cells [with an association cutoff of $-\log_{10}(P) > 1.3$]. (D) Mitochondrial content (as measured by MitoTracker) of CD4⁺ T cells decreases with increased bright-field (BF) intensity. Bar plots display the mean MitoTracker intensity of CD4⁺ T cells per well per 10-percentile bins of bright-field intensity within each well. Mean and SDs across 10 repeat wells with a combined total of $n = 78,095$ CD4⁺ T cells are shown. *P* values are from a two-tailed *t* test of all replicate wells per bin against those of the brightest bright-field (right most) bin. ****P* < 0.001.

marker profiles, and LEA requires donor or condition heterogeneity to power its associations. Furthermore, while the marker panel shown here reliably captures the predominant immune cell classes present in PBMCs, it does not resolve certain smaller subpopulations, such as NK T cells (61). However, the approach is flexible as the panel composition can readily be tailored to the identification of additional subpopulations or adapted to different tissues, building on the same logic developed here.

This is not the first work that deploys CNN-based cell classification (62–70) and feature extraction (63, 65, 69, 71). Here, we apply deep learning in high-throughput screening and phenotypic analyses of primary human PBMCs. By training the CNN on curated cells from across independent experiments, multiple donors, and conventional and multiplexed staining panels, we could prevent overfitting on phenotypes of single donors and technical bias stemming from experimental conditions. However, the CNN class probability space, which we here successfully use as a phenotype discovery tool, is sensitive to different phenotypes resulting from different experimental conditions. Hence, while CNN classification can be trained to be robust, experimental care needs to be taken when interpreting the CNN class probability space.

Once new phenotypes are found, as we demonstrate for the inflammation-associated T_{ACT} cell morphology, the ability to retroactively reclassify cells based on their morphology with dedicated CNNs allows robust morphological subclassification of previously imaged cells even in the absence of tailored marker panels. Attesting to the robustness of the found phenotypes, the inflammation-associated T_{ACT} and age-associated T_{4BFD} phenotypes could be validated in independent experiments, in an independent validation cohort, using distinct neural network architectures, and, for the T_{ACT} morphology, in both live-cell and fixed sample imaging.

In the future, repeated profiling of individual donors will allow to further stratify temporally stable from dynamic immune cell phenotypes. Furthermore, comparative studies across larger patient and donor cohorts, as well as identifying clinically relevant cell morphologies in the context of personalized treatment identification for hematological malignancies (41, 42), will be additionally attractive avenues of study. This will inevitably define the boundaries of the personal health information reflected by immune cell phenotypes. Given that the workflow allows simultaneous phenotype discovery combined with the molecular and personal health associations, it is well positioned to lead to the discovery of more as yet undescribed and clinically relevant immune cell phenotypes.

MATERIALS AND METHODS

Experimental model

Buffy coats or whole blood tubes were obtained from coded healthy donors provided by the Blutspende Zürich, under a study protocol approved by the Cantonal Ethics Committee, Zürich (KEK Zürich, BASEC-Nr 2019-01579). Detailed donor information can be found in table S3.

Experimental details

Collection and purification of human PBMCs

Healthy donor buffy coats or blood samples were diluted 1:1 in phosphate-buffered saline (PBS; Gibco), and PBMCs were isolated with a Histopaque-1077 density gradient (Sigma-Aldrich) according to the manufacturer's instructions. PBMCs at the interface

were collected, washed once in PBS, and resuspended in medium. In all experiments, immune cells were cultured in RPMI 1640 and GlutaMAX medium (Gibco) supplemented with 10% fetal bovine serum (FBS; Gibco) and incubated at 37°C with 5% CO₂. Cell number and viability were determined using a Countess II Cell Counter from Thermo Fisher Scientific according to the manufacturer's instructions.

Nonadherent PBMC monolayer formation and drug screening and cell fixation

In the proof-of-concept drug screen, 5 μ l of a selected screening compounds (10 \times stock) and all respective controls (as outlined in table S2) were transferred to CellCarrier 384 Ultra, clear-bottom, tissue culture-treated plates (PerkinElmer) with five replicates per condition. All conditions were screened in four concentrations: cytokines (0.1, 1, 10, and 100 ng/ml), rituximab (0.05, 0.1, 0.5, and 1 μ g/ml), LPS (0.1, 1, 10, and 100 ng/ml), dexamethasone (0.4, 4, 40, and 400 ng/ml), and crizotinib (0.01, 0.1, 1, and 10 μ M). Fifty microliters of medium containing approximately 4×10^5 cells/ml was pipetted into each well of a 384-well compound plate, and cells were allowed to settle to the bottom. The whole blood samples of the discovery cohort (shown in Figs. 2, A and B, and 4 to 6) were incubated for 1 hour, whereas all buffy coat samples, including all samples from the validation cohort (Figs. 5H and 6C), were incubated for 24 hours. All assays were terminated by fixing and permeabilizing the cells with 20 μ l of a solution containing 0.5% (w/v) formaldehyde (Sigma-Aldrich), 0.05% (v/v) Triton X-100 (Sigma-Aldrich), 10 mM sodium(meta)periodate (Sigma-Aldrich), and 75 mM L-lysine monohydrochloride (Sigma-Aldrich), for 20 min at room temperature. For MitoTracker staining (Thermo Fisher Scientific), cells were stained live with 500 nM MitoTracker Red, before fixation. Fixative-containing medium was subsequently removed, and cells were blocked and photobleached in 5% FBS/PBS overnight at 4°C. Photobleaching was used to reduce background fluorescence and was performed by illuminating the fixed cells with conventional white light light-emitting diode panels.

Immunostaining and imaging

All fluorescent primary antibodies used in this work (outlined in table S1) were used at a 1:300 dilution in PBS. All antibody cocktails for immunohistochemistry (IHC) contained 6 μ M DAPI (Sigma-Aldrich) for nuclear detection. Before IHC staining, the blocking solution was removed, and 20 μ l of the antibody cocktail was added per well and incubated for 1 hour at room temperature. Besides fully multiplexed wells, each plate additionally contained several staining control wells with a reduced number of antibodies (table S1). The staining control wells served for evaluating antibody functionality and the generation of the CNN training data (see below). For imaging, a PerkinElmer Opera Phenix automated spinning disk confocal microscope was used. Each well of a 384-well plate was imaged at $\times 20$ magnification with 5×5 nonoverlapping images, covering the whole well surface. The images were taken sequentially from the bright-field (650 to 760 nm), DAPI/nuclear signal (435 to 480 nm), green fluorescent protein/green signal (500 to 550 nm), phycoerythrin/orange signal (570 to 630 nm), and APC/red signal (650 to 760 nm) channels. Subsequently, the raw .tiff images were transferred from the microscope for further analysis.

Conventional image analysis and quality filtering

Cell detection and single-cell image analysis were performed using CellProfiler v2 (72). Nuclear segmentation was performed via thresholding on DAPI intensity. Cellular outlines were estimated by a

circular expansion from the outlines of the nucleus. In addition, a second and larger expansion from the nuclei was performed to measure the local area around each single cell (local cellular background). Standard CellProfiler-based intensity, shape, and texture features of the nucleus and cytoplasm and the local cell proximity were extracted for each measured channel. Raw fluorescent intensities were \log_{10} -transformed and normalized toward the local cellular background as described by Vladimer *et al.* (40).

Convolutional neural networks

CNNs used in this work were implemented using MATLAB's Neural Network Toolbox version R2020a. The curated dataset used in training, validation, and testing of the CNN framework contains images of cells from fully multiplexed stainings and images from staining controls. Staining controls were designed to contain only a subset of the antibodies used in the multiplexed setting (table S1). This reduced complexity first enables to evaluate the functionality of the selected antibody and the presence of the targeted antigen in each sample. Furthermore, antibody combinations in the staining controls were picked to mirror the staining of the selected subpopulation in the multiplexed setting (e.g., staining control 1 only contained antibodies marking T cell-specific antigens; T cells in the multiplexed setting will have the same staining pattern). The same staining patterns in the controls and the mostly nonoverlapping emission spectra of the chosen antibodies allow an easy, marker intensity-based identification of subpopulations. This facilitates a fast and unbiased selection of training examples. For the generation of single-cell images, the center of each cell was determined by its nuclear staining via the software CellProfiler (see above). Around each nuclei center, a 50-pixel by 50-pixel (or 39.5- μm by 39.5- μm) wide subimage was generated across all five measured channels. Single-cell subimages were then manually annotated and sorted for their respective class using custom MATLAB scripts. For training and validation of the discovery cohort CNN, a dataset of 89,483 cells was manually annotated (containing both multiplexed and control staining cells). In the separate test datasets, each donor-associated set is independently split in multiplexed and control staining cells, resulting in a total of 30 independent test datasets with each 100 cells per class. This test setup allows inferring the network performance toward each donor, experiment, and staining type independently.

In discovery cohort (10 donors), a 17-layer-deep CNN with an adapted "AlexNet" architecture (73) with 50-pixel by 50-pixel and five-channel input images was used. Before training, the labeled eight-class dataset was randomly split in a training set containing 90% and a validation set with the remaining 10% of all images. Network layer weights and biases were initialized randomly before the CNN network was trained. Networks were trained up to 20 epochs with a mini batch size of 512 images. The learning rate was fixed to 0.0001. To avoid overfitting, L2 regularization with 0.005 was applied. Furthermore, in each iteration, input images were randomly rotated in 45° steps with an additional possibility to be also flipped vertically or horizontally. Performance of the trained networked was tested on the separate test sets of staining control and multiplexed images of all 15 donors. Stochastic gradient descent with momentum of 0.9 is defined as the optimization algorithm. Last, we trained 20 differently initialized networks with differently split training and validation sets. For the final classification of the complete unlabeled dataset, the best performing network was used. As in the generation of the labeled dataset, 50-pixel by 50-pixel subimages around each nuclei

center were generated. Cells closer than 25 pixels to the border of an image were excluded from classification.

In validation cohort (15 donors), a 71-layer-deep CNN with an adapted ResNet architecture (74) with 48-pixel by 48-pixel and five-channel input images was used. Before classification and training, all intensity values were first \log_{10} -transformed and then channel wise-normalized to a range of 0 to 1. The eight-class CNN was trained using randomly initialized weights and biases and the adaptive learning rate optimization "ADAM." The network was trained for 20 epochs with an initial learning rate of 0.001, which was dropped every 5 epochs with a factor of 0.1. Furthermore, a mini batch size of 512 images and L2 regularization with 0.001 was applied. To further strengthen generalization, input images were augmented in each iteration. Here, images were randomly rotated in 45° steps with an additional possibility to be also flipped vertically or horizontally. To block an overreliance on absolute intensity values, channel intensity shifts were simulated via a multiplication with a random fixed factor. This used factor was randomly drawn out of a normal distribution with a mean of 1 and an SD of 0.2. Furthermore, images were augmented with random noise (specifically salt and pepper noise, speckle noise, Gaussian noise, or image blurring). In all CNN classifications, 48-pixel by 48-pixel subimages around each nuclei center were generated. Cells closer than 24 pixels to the border of an image were excluded from all classifications.

In label-free T_{ACT} classifier, CNNs and single-cell images were generated as described above. The labeled training and validation dataset comprised a total of 8862 cells (1:2 T_{ACT}:T_{CON} ratio). CNNs were trained with a mini-batch size of 200 images to a maximum of 100 epochs, which could be terminated if validation loss was greater than the previous smallest loss for five consecutive times. In addition, the images were randomly rotated by 45° and mirrored vertically or horizontally per iteration to limit orientation bias toward polarized T_{ACT} cells. The CNN performance was assessed by classifying 1107 test cells (1:2 T_{ACT}:T_{CON} ratio) that had neither been used in CNN training nor in validation.

RNA sequencing

In T cell isolation and RNA extraction, T cells were isolated from fresh PBMCs directly after obtaining them via density centrifugation, as described above. Isolation was performed via a column-based extraction method with CD3 Microbeads as described in the manufacturer's instructions (Miltenyi Biotec). RNA extraction of the isolated cells was performed with a Quick-RNA MiniPrep Kit by Zymo according to manufacturer's instructions.

RNA-seq was performed by the Functional Genomics Center Zürich. Briefly, cDNA libraries were obtained according to protocols published by Picelli *et al.* (75). Illumina library was obtained via tagmentation using the Illumina Nextera Kit. All samples were sequenced in a single run on a NovaSeq 6000 (single read, 100 base pairs, depth of 20 Mio reads per sample).

In data processing and normalization, Illumina adapters, sequences of poor quality, as well as polyadenine (polyA) and polythymine (polyT) sequences were removed from the raw reads using TrimGalore v.0.6.0 with cutadapt v.2.0 before alignment. Reads were then aligned to the human reference genome GRCh38, v93 (Ensembl) using STAR v. 2.5.3a. Reads per gene were counted using the `-quantMode GeneCounts` flag in STAR. Gene counts below a threshold of 20 raw counts were filtered, and raw counts were normalized [DESeq2 (76)]. Only transcripts annotated as "protein coding" or "long noncoding RNA" were considered in the subsequent analysis.

Statistical analysis

If not stated otherwise, all significance scores were calculated on the basis of a two-tailed Student's *t* test with mean of 0. Asterisks in figures indicate significance per condition, compared with controls. *P* value less than 0.05 is flagged with one asterisk (*), *P* value less than 0.01 is flagged with two asterisks (**), and all *P* values with less than 0.001 are flagged with three asterisks (***).

For cell-cell interaction analysis, a simplified version of the interaction method by Vladimer *et al.* (40) was used. Here, cell-cell interaction analysis was conducted over all different image sites within the same well. Cells were scored as interacting if their nuclear centroids were within a Euclidean distance of 40 pixels. To calculate the interaction score of a cell with type A interacting with a cell of type B, we first calculated specific interactions and total interactions per well. We define specific interactions, as the total count of "B" cells within the defined radius around a cell of type "A." Total interactions are considered as the total count of all interacting cells in that well. To calculate the final interaction score, specific interactions were divided by the product of (the fraction of type A cells of all cells) × (the fraction of type B cells of all cells) × total interactions. In contrast to the previously published method, this approach is simplified as the interactions scores are nondirected, which reduces the number of edges from 72 to 36. Mean interaction score over all replicates was calculated, log₂-transformed, and normalized toward its respective control (see table S2).

All *t*-SNE visualizations were calculated on the $-\log_{10}(\text{class probability matrices})$. In the *t*-SNE calculation, a Mahalanobis distance metric, a perplexity of 30, and an exaggeration parameter of 4 were applied. To reduce the calculation time, the Barnes-Hut algorithm with a θ value of 0.5 was used.

To calculate whether a certain condition displays local enrichment in the eight-dimensional class probability space, we developed the KNN LEA by hypergeometric testing or rank-based correlation. Here, we probe the local neighborhood around each single cell, which is defined as the KNNs in the original CNN class probability space. For discrete variables (such as donor identity), we calculate the probability to randomly find at least *n* cells of condition *X* in a certain neighborhood using a hypergeometric cumulative distribution function. This takes into account the total number of cells in the probed neighborhood, the total number of cells in the tested class probability space, and the total number of cells of condition *X*. In case of continuous variables (such as donor age or gene transcript counts), the relative fraction of cells of each donor in the probed local neighborhood is calculated. The fractions are then correlated (Spearman's rank correlation) with a continuous variable, and the significance of the correlation is calculated. In both cases, the enrichment probability is assigned to the center cell of the probed region, and the approach is iterated for each single cell in the selected *n*-dimensional space. If not stated otherwise, neighborhoods were defined as $k = 400$ nearest neighbors for Figs 2 to 4 and figs. S2 to S4 and as $k = 200$ for the T cell for Figs. 5 and 6 and figs. S5 and S6. In general, KNN-based algorithms, such as KNN-based classifiers, can overfit when *K* is set too small or underfit when *K* is set too big. As a general rule of thumb, the square root of the number of data points in the dataset is a reasonable starting value for *K*. Please see data file S1 for a full example workflow of LEA. *P* values were corrected for multiple testing, i.e., by the number of total cells (i.e., tests) in the analysis.

In pathway enrichment analysis, pathway annotations were obtained using the DAVID database (77). Gene enrichments per single

cell were calculated via LEA (see above). To calculate pathway enrichments per single cell, the LEA gene enrichments of all genes belonging to a certain pathway annotation were compared against the enrichment of all other genes. Significance scores were calculated on the basis of a two-tailed Student's *t* test, and directionality was calculated by the difference of the means of both populations.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abn5631>

[View/request a protocol for this paper from Bio-protocol.](#)

REFERENCES AND NOTES

1. C. Bakal, J. Aach, G. Church, N. Perrimon, Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science* **316**, 1753–1756 (2007).
2. J. Folkman, A. Moscona, Role of cell shape in growth control. *Nature* **273**, 345–349 (1978).
3. T. Lecuit, P.-F. Lenne, Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat. Rev. Mol. Cell Biol.* **8**, 633–644 (2007).
4. M. Boutros, F. Heigwer, C. Laufer, Microscopy-based high-content screening. *Cell* **163**, 1314–1325 (2015).
5. R. McBeath, D. M. Pirone, C. M. Nelson, K. Bhadriraju, C. S. Chen, Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. *Dev. Cell* **6**, 483–495 (2004).
6. D. E. Discher, D. J. Mooney, P. W. Zandstra, Growth factors, matrices, and forces combine and control stem cells. *Science* **324**, 1673–1677 (2009).
7. J. G. Carlton, H. Jones, U. S. Eggert, Membrane and organelle dynamics during cell division. *Nat. Rev. Mol. Cell Biol.* **21**, 151–166 (2020).
8. N. Ramkumar, B. Baum, Coupling changes in cell shape to chromosome segregation. *Nat. Rev. Mol. Cell Biol.* **17**, 511–521 (2016).
9. B. Snijder, R. Sacher, P. Rämö, E.-M. Damm, P. Liberali, L. Pelkmans, Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* **461**, 520–523 (2009).
10. B. Snijder, L. Pelkmans, Origins of regulated cell-to-cell variability. *Nat. Rev. Mol. Cell Biol.* **12**, 119–125 (2011).
11. P. Liberali, B. Snijder, L. Pelkmans, A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell* **157**, 1473–1487 (2014).
12. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
13. P.-H. Wu, D. M. Gilkes, J. M. Phillip, A. Narkar, T. W.-T. Cheng, J. Marchand, M.-H. Lee, R. Li, D. Wirtz, Single-cell morphology encodes metastatic potential. *Sci. Adv.* **6**, eaaw6938 (2020).
14. L. Zhou, M. M. W. Chong, D. R. Littman, Plasticity of CD4⁺ T cell lineage differentiation. *Immunity* **30**, 646–655 (2009).
15. S. J. Galli, N. Borregaard, T. A. Wynn, Phenotypic and functional plasticity of cells of innate immunity: Macrophages, mast cells and neutrophils. *Nat. Immunol.* **12**, 1035–1044 (2011).
16. A. Sica, A. Mantovani, Macrophage plasticity and polarization: In vivo veritas. *J. Clin. Invest.* **122**, 787–795 (2012).
17. H. T. Maecker, J. P. McCoy, R. Nussenblatt, Standardizing immunophenotyping for the Human Immunology Project. *Nat. Rev. Immunol.* **12**, 191–200 (2012).
18. F. E. Craig, K. A. Foon, Flow cytometric immunophenotyping for hematologic neoplasms. *Blood* **111**, 3941–3967 (2008).
19. M. H. Spitzer, G. P. Nolan, Mass cytometry: Single cells, many features. *Cell* **165**, 780–791 (2016).
20. S. C. Bendall, E. F. Simonds, P. Qiu, E.-a. D. Amir, P. O. Krutzik, R. Finck, R. V. Bruggner, R. Melamed, A. Trejo, O. I. Ornatsky, R. S. Balderas, S. K. Plevritis, K. Sachs, D. Pe'er, S. D. Tanner, G. P. Nolan, Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011).
21. E. Papalexli, R. Satija, Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* **18**, 35–45 (2018).
22. D. A. Jaitin, E. Kenigsberg, H. Keren-Shaul, N. Elefant, F. Paul, I. Zaretsky, A. Mildner, N. Cohen, S. Jung, A. Tanay, I. Amit, Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).
23. A. K. Shalek, R. Satija, X. Adiconis, R. S. Gertner, J. T. Gaublot, R. Raychowdhury, S. Schwartz, N. Yosef, C. Malboeuf, D. Lu, J. J. Trombetta, D. Gennert, A. Gnirke, A. Goren, N. Hacohen, J. Z. Levin, H. Park, A. Regev, Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498**, 236–240 (2013).
24. A.-C. Villani, R. Satija, G. Reynolds, S. Sarkizova, K. Shekhar, J. Fletcher, M. Griesbeck, A. Butler, S. Zheng, S. Lazo, L. Jardine, D. Dixon, E. Stephenson, E. Nilsson, I. Grundberg, D. McDonald, A. Filby, W. Li, P. L. De Jager, O. Rozenblatt-Rosen, A. A. Lane, M. Haniffa,

- A. Regev, N. Hacohen, Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
25. A. Giladi, I. Amit, Single-cell genomics: A stepping stone for future immunology discoveries. *Cell* **172**, 14–21 (2018).
 26. F. R. Bertani, P. Mozetic, M. Fioramonti, M. Iuliani, G. Ribelli, F. Pantano, D. Santini, G. Tonini, M. Trombetta, L. Businaro, S. Selci, A. Rainer, Classification of M1/M2-polarized human macrophages by label-free hyperspectral reflectance confocal microscopy and multivariate analysis. *Sci. Rep.* **7**, 8965 (2017).
 27. F. Y. McWhorter, T. Wang, P. Nguyen, T. Chung, W. F. Liu, Modulation of macrophage phenotype by cell shape. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17253–17258 (2013).
 28. C. Gómez-Moutón, J. L. Abad, E. Mira, R. A. Lacalle, E. Gallardo, S. Jiménez-Baranda, I. Illa, A. Bernad, S. Mañes, C. Martínez-A. Segregation of leading-edge and uropod components into specific lipid rafts during T cell polarization. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9642–9647 (2001).
 29. K. B. L. Lin, S. A. Freeman, S. Zabetian, H. Brugger, M. Weber, V. Lei, M. Dang-Lawson, K. W. K. Tse, R. Santamaria, F. D. Batista, M. R. Gold, The rap GTPases regulate B cell morphology, immune-synapse formation, and signaling by particulate B cell receptor ligands. *Immunity* **28**, 75–87 (2008).
 30. N. van Panhuys, F. Klauschen, R. N. Germain, T-cell-receptor-dependent signal intensity dominantly controls CD4⁺ T cell polarization in vivo. *Immunity* **41**, 63–74 (2014).
 31. S. Russell, How polarity shapes the destiny of T cells. *J. Cell Sci.* **121**, 131–136 (2008).
 32. S. Faure, L. I. Salazar-Fontana, M. Semichon, V. L. J. Tybulewicz, G. Bismuth, A. Trautmann, R. N. Germain, J. Delon, ERM proteins regulate cytoskeleton relaxation promoting T cell-APC conjugation. *Nat. Immunol.* **5**, 272–279 (2004).
 33. W. Lin, Y. Suo, Y. Deng, Z. Fan, Y. Zheng, X. Wei, Y. Chu, Morphological change of CD4⁺ T cell during contact with DC modulates T-cell activation by accumulation of F-actin in the immunology synapse. *BMC Immunol.* **16**, 49 (2015).
 34. F. Buggenthin, F. Buettner, P. S. Hoppe, M. Endele, M. Kroiss, M. Strasser, M. Schwarzfischer, D. Loeffler, K. D. Kokkariaris, O. Hilsenbeck, T. Schroeder, F. J. Theis, C. Marr, Prospective identification of hematopoietic lineage choice by deep learning. *Nat. Methods* **14**, 403–406 (2017).
 35. A. G. Tsai, D. R. Glass, M. Juntilla, F. J. Hartmann, J. S. Oak, S. Fernandez-Pol, R. S. Ohgami, S. C. Bendall, Multiplexed single-cell morphometry for hematopathology diagnostics. *Nat. Med.* **26**, 408–417 (2020).
 36. Z. E. Perlman, M. D. Slack, Y. Feng, T. J. Mitchison, L. F. Wu, S. J. Altschuler, Multidimensional drug profiling by automated microscopy. *Science* **306**, 1194–1198 (2004).
 37. M. J. Wawer, K. Li, S. M. Gustafsdottir, V. Ljosa, N. E. Bodycombe, M. A. Marton, K. L. Sokolnicki, M.-A. Bray, M. M. Kemp, E. Winchester, B. Taylor, G. B. Grant, C. S.-Y. Hon, J. R. Duvall, J. A. Wilson, J. A. Bittker, V. Dančík, R. Narayan, A. Subramanian, W. Winckler, T. R. Golub, A. E. Carpenter, A. F. Shamji, S. L. Schreiber, P. A. Clemons, Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10911–10916 (2014).
 38. D. W. Young, A. Bender, J. Hoyt, E. McWhinnie, G.-W. Chirn, C. Y. Tao, J. A. Tallarico, M. Labow, J. L. Jenkins, T. J. Mitchison, Y. Feng, Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nat. Chem. Biol.* **4**, 59–68 (2008).
 39. J. C. Caicedo, S. Cooper, F. Heigwer, S. Warchal, P. Qiu, C. Molnar, A. S. Vasilevich, J. Barry, H. S. Bansal, O. Kraus, M. Wawer, L. Paavolainen, M. D. Herrmann, M. Rohban, J. Hung, H. Hennig, J. Concannon, I. Smith, P. A. Clemons, S. Singh, P. Rees, P. Horvath, R. G. Linington, A. E. Carpenter, Data-analysis strategies for image-based cell profiling. *Nat. Methods* **14**, 849–863 (2017).
 40. G. I. Vladimer, B. Snijder, N. Krall, J. W. Bigenzahn, K. V. M. Huber, C.-H. Lardeau, K. Sanjiv, A. Ringler, U. W. Berglund, M. Sabler, O. L. de la Fuente, P. Knöbl, S. Kubicek, T. Helleday, U. Jäger, G. Superti-Furga, Global survey of the immunomodulatory potential of common drugs. *Nat. Chem. Biol.* **13**, 681–690 (2017).
 41. B. Snijder, G. I. Vladimer, N. Krall, K. Miura, A.-S. Schmolke, C. Kornauth, O. Lopez de la Fuente, H.-S. Choi, E. van der Kouwe, S. Gültelkin, L. Kazianka, J. W. Bigenzahn, G. Hoermann, N. Prutsch, O. Merkel, A. Ringler, M. Sabler, G. Jerczynski, M. E. Mayerhoefer, I. Simonitsch-Klupp, K. Ocko, F. Felberbauer, L. Müllauer, G. W. Prager, B. Korkmaz, L. Kenner, W. R. Sperr, R. Kralovics, H. Gisslinger, P. Valent, S. Kubicek, U. Jäger, P. B. Staber, G. Superti-Furga, Image-based ex-vivo drug screening for patients with aggressive haematological malignancies: Interim results from a single-arm, open-label, pilot study. *Lancet Haematol.* **4**, e595–e606 (2017).
 42. C. Kornauth, T. Pemovska, G. I. Vladimer, G. Bayer, M. Bergmann, S. Eder, R. Eichner, M. Erl, H. Esterbauer, R. Exner, V. Felsleitner-Hauer, M. Forte, A. Gaiger, K. Geissler, H. T. Greinix, W. Gstöttner, M. Hacker, B. L. Hartmann, A. W. Hauswirth, T. Heinemann, D. Heintel, M. A. Hoda, G. Hopfinger, U. Jaeger, L. Kazianka, L. Kenner, B. Kiesewetter, N. Krall, G. Krajnik, S. Kubicek, T. Le, S. Lubowitzki, M. E. Mayerhoefer, E. Menschel, O. Merkel, K. Miura, L. Müllauer, P. Neumeister, T. Noesslinger, K. Ocko, L. Ohler, M. Panny, A. Pichler, E. Porpacz, G. W. Prager, M. Raderer, R. Ristl, R. Ruckser, J. Salamon, A.-I. Schiefer,
 - A.-S. Schmolke, I. Schwarzinger, E. Selzer, C. Sillaber, C. Skrabs, W. R. Sperr, I. Srdic, R. Thalhammer, P. Valent, E. van der Kouwe, K. Vanura, S. Vogt, C. Waldstein, D. Wolf, C. C. Zielinski, N. Zojer, I. Simonitsch-Klupp, G. Superti-Furga, B. Snijder, P. B. Staber, Functional precision medicine provides clinical benefit in advanced aggressive hematologic cancers and identifies exceptional responders. *Cancer Discov.* **12**, 372–387 (2022).
 43. J.-R. Lin, M. Fallahi-Sichani, P. K. Sorger, Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat. Commun.* **6**, 8390 (2015).
 44. M. J. Gerdes, C. J. Sevinsky, A. Sood, S. Adak, M. O. Bello, A. Bordwell, A. Can, A. Corwin, S. Dinn, R. J. Filkins, D. Hollman, V. Kamath, S. Kaanumalle, K. Kenny, M. Larsen, M. Lazare, Q. Li, C. Lowes, C. C. McCulloch, E. McDonough, M. C. Montalto, Z. Pang, J. Rittscher, A. Santamaria-Pang, B. D. Sarachan, M. L. Seel, A. Seppo, K. Shaikh, Y. Sui, J. Zhang, F. Ginty, Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11982–11987 (2013).
 45. G. Gut, M. D. Herrmann, L. Pelkmans, Multiplexed protein maps link subcellular organization to cellular states. *Science* **361**, eaar7042 (2018).
 46. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
 47. J. Xie, R. Girshick, A. Farhadi, in *International Conference on Machine Learning* (PMLR, 2016), pp. 478–487.
 48. E. Aljalbout, V. Golkov, Y. Siddiqui, M. Strobel, D. Cremers, Clustering with deep learning: Taxonomy and new methods. arXiv:1801.07648 [cs.LG] (23 January 2018).
 49. L. Maaten, G. Hinton, Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
 50. F. O. Martinez, S. Gordon, The M1 and M2 paradigm of macrophage activation: Time for reassessment. *F1000Prime Rep.* **6**, 13 (2014).
 51. M. Melé, P. G. Ferreira, F. Reverter, D. S. DeLuca, J. Monlong, M. Sammeth, T. R. Young, J. M. Goldmann, D. D. Pervouchine, T. J. Sullivan, R. Johnson, A. V. Segrè, S. Djebali, A. Niarchou; GTEx Consortium, F. A. Wright, T. Lappalainen, M. Calvo, G. Getz, E. T. Dermizakis, K. G. Ardlie, R. Guigó, Human genomics. The human transcriptome across tissues and individuals. *Science* **348**, 660–665 (2015).
 52. E. J. Carr, J. Dooley, J. E. Garcia-Perez, V. Lagou, J. C. Lee, C. Wouters, I. Meyts, A. Goris, G. Boeckxstaens, M. A. Linterman, A. Liston, The cellular composition of the human immune system is shaped by age and cohabitation. *Nat. Immunol.* **17**, 461–468 (2016).
 53. A. Alpert, Y. Pickman, M. Leipold, Y. Rosenberg-Hasson, X. Ji, R. Gaujoux, H. Rabani, E. Starosvetsky, K. Kveler, S. Schaffert, D. Furman, O. Caspi, U. Rosenschein, P. Khatri, C. L. Dekker, H. T. Maecker, M. M. Davis, S. S. Shen-Orr, A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. *Nat. Med.* **25**, 487–495 (2019).
 54. A. Roth, H. Yssel, J. Pene, E. A. Chavez, M. Schertzer, P. M. Lansdorp, H. Spits, R. M. Luiten, Telomerase levels control the lifespan of human T lymphocytes. *Blood* **102**, 849–857 (2003).
 55. L. A. Stephen, Y. ElMaghloob, M. J. McIlwraith, T. Yelland, P. Castro Sanchez, P. Roda-Navarro, S. Ismail, The ciliary machinery is repurposed for T cell immune synapse trafficking of LCK. *Dev. Cell* **47**, 122–132.e4 (2018).
 56. C. Franceschi, M. Bonafè, S. Valensin, Inflamm-aging: An evolutionary perspective on immunosenescence. *Ann. N. Y. Acad. Sci.* **908**, 244–254 (2000).
 57. J. Jin, X. Li, B. Hu, C. Kim, W. Cao, H. Zhang, C. M. Weyand, J. J. Goronzy, FOXO1 deficiency impairs proteostasis in aged T cells. *Sci. Adv.* **6**, eaba1808 (2020).
 58. N. Ron-Harel, G. Notarangelo, J. M. Ghergurovich, J. A. Paulo, P. T. Sage, D. Santos, F. K. Satterstrom, S. P. Gygi, J. D. Rabinowitz, A. H. Sharpe, M. C. Haigis, Defective respiration and one-carbon metabolism contribute to impaired naïve T cell activation in aged mice. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 13347–13352 (2018).
 59. A. P. Gomes, N. L. Price, A. J. Y. Ling, J. J. Moslehi, M. K. Montgomery, L. Rajman, J. P. White, J. S. Teodoro, C. D. Wrann, B. P. Hubbard, E. M. Mercken, C. M. Palmeira, R. de Cabo, A. P. Rolo, N. Turner, E. L. Bell, D. A. Sinclair, Declining NAD⁺ induces a pseudohypoxic state disrupting nuclear-mitochondrial communication during aging. *Cell* **155**, 1624–1638 (2013).
 60. D. Murera, F. Arbogast, J. Arnold, D. Bouis, S. Muller, F. Gros, CD4 T cell autophagy is integral to memory maintenance. *Sci. Rep.* **8**, 5951 (2018).
 61. A. Bendelac, P. B. Savage, L. Teyton, The biology of NKT cells. *Annu. Rev. Immunol.* **25**, 297–336 (2007).
 62. E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, D. Van Valen, Deep learning for cellular image analysis. *Nat. Methods* **16**, 1233–1246 (2019).
 63. O. Z. Kraus, B. T. Grys, J. Ba, Y. Chong, B. J. Frey, C. Boone, B. J. Andrews, Automated analysis of high-content microscopy data with deep learning. *Mol. Syst. Biol.* **13**, 924 (2017).
 64. O. Z. Kraus, J. L. Ba, B. J. Frey, Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics* **32**, i52–i59 (2016).
 65. T. Pärnamaa, L. Parts, Accurate classification of protein subcellular localization from high-throughput microscopy images using deep learning. *G3* **7**, 1385–1392 (2017).

66. O. Dürr, B. Sick, Single-cell phenotype classification using deep convolutional neural networks. *J. Biomol. Screen.* **21**, 998–1003 (2016).
67. C. Sommer, R. Hoefler, M. Samwer, D. W. Gerlich, A deep learning and novelty detection framework for rapid phenotyping in high-content screening. *Mol. Biol. Cell* **28**, 3428–3436 (2017).
68. C. Kandaswamy, L. M. Silva, L. A. Alexandre, J. M. Santos, High-content analysis of breast cancer using single-cell deep transfer learning. *J. Biomol. Screen.* **21**, 252–259 (2016).
69. W. J. Godinez, I. Hossain, S. E. Lazic, J. W. Davies, X. Zhang, A multi-scale convolutional neural network for phenotyping high-content cellular images. *Bioinformatics* **33**, 2010–2019 (2017).
70. S. Hussain, A. Das, B. P. Nguyen, M. Marzuki, S. Lin, A. Kumar, G. Wright, A. Singhal, Deep transfer learning for high-content analysis of the effects of drugs on immune cells, in *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)* (IEEE, 2019), pp. 796–801.
71. P. T. Jackson, Y. Wang, S. Knight, H. Chen, T. Dorval, M. Brown, C. Bendtsen, B. Obara, Phenotypic profiling of high throughput imaging screens with generic deep convolutional features, in *International Conference on Machine Vision Applications Tokyo, Japan, 27–31 May 2019* (IEEE, 2019), pp. 1–4.
72. A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, D. M. Sabatini, CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
73. A. Krizhevsky, I. Sutskever, G. E. Hinton, in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger, Eds. (Curran Associates Inc., 2012), pp. 84–90, vol. 25.
74. 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 (CVPR)* (IEEE, 2016), pp. 770–778.
75. S. Picelli, A. K. Björklund, B. Reinius, S. Sagasser, G. Winberg, R. Sandberg, Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).
76. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
77. D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
78. M. D. Wilkinson, M. Dumontier, I. J. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hoof, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).

Acknowledgments: We thank the anonymous blood donors from the Blutspende Zürich for contributions to this study. We thank F. Sallusto and N. de Souza, as well as all members of the Snijder lab, for constructive feedback. **Funding:** This work was supported by the Swiss National Science Foundation (project numbers PP00P3_163961 and PP00P3_194809). The funder had no role in study design, data collection and interpretation, or the decision to submit the work for publication. **Author contributions:** Y.S. designed and performed experiments and computational analysis. J.M. helped in the experimental design and advised on technical decisions. B.D.H. contributed to the morphological quantification of T cells. Buffy coats or whole blood samples were obtained by the Blutspende Zürich, under supervision of D.G. and B.M.F. B.S. supervised this study, contributed to computational analysis, and advised on experimental design and data interpretation. Y.S. and B.S. wrote the manuscript, and all authors commented and approved the final version. **Competing interests:** B.S. was a scientific cofounder of Allcyte GmbH, which has been acquired by Exscientia. B.S. is a shareholder of Exscientia and coinventor on U.S. patent application 15/514,045 relevant to the study. B.S. declares research funding from Roche and speaker fees from Novartis, GSK, and AbbVie. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The LEA MATLAB code and example analysis are available at Dryad (<https://doi.org/10.5061/dryad.2jm63xssk>). The single-cell imaging dataset and deep learning dataset and relevant metadata are available at the FAIR principles (78) compliant repository of the ETH Zürich (<https://doi.org/10.3929/ethz-b-000343106>). T cell RNA-seq measurements used in the study are uploaded to the GEO repository with accession number GSE155093 (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155093).

Submitted 7 December 2021
Accepted 14 September 2022
Published 2 November 2022
10.1126/sciadv.abn5631