Comparing phenotypic manifolds with Kompot: Detecting differential abundance and gene expression at single-cell resolution

Dominik J. Otto[1,2,3], Erica Arriaga-Gomez[4], Elana Thieme[1,2,3,4,5], Ruijin Yang[1,2,3], Stanley C. Lee[4]+, Manu Setty[1,2,3]+

[1] Basic Sciences Division, Fred Hutchinson Cancer Center, Seattle WA

[2] Computational Biology Program, Public Health Sciences Division, Seattle WA

[3] Translational Data Science IRC, Fred Hutchinson Cancer Center, Seattle WA

[4] Translational Science and Therapeutics Division, Fred Hutchinson Cancer Center, Seattle WA

[5] Molecular and Cellular Biology Program, University of Washington, Seattle WA

+ Corresponding authors:

Manu Setty (msetty@fredhutch.org)

Stanley Lee (sclee3@fredhutch.org)

Abstract

Single-cell studies are frequently designed to compare across conditions such as health and disease. However, existing computational approaches typically rely on grouping cells into discrete populations before making comparisons, which can limit resolution for detecting state-dependent changes. Here, we introduce Kompot, a statistical framework for comparative analysis of multi-condition single-cell data. Kompot quantifies both differential abundance, capturing how cells redistribute across the phenotypic space, and differential expression, identifying condition-specific transcriptional changes that may be localized, heterogeneous, or oppositely regulated across states. By modeling cell density and gene expression as continuous functions over a shared cell-state representation, Kompot enables single-cell–resolution inference with principled uncertainty estimates, without requiring predefined clusters or cell types. Applying Kompot to aging murine bone marrow, we identified a continuum of shifts in hematopoietic stem cell and mature cell states, transcriptional remodeling of monocytes independent of compositional changes, and divergent regulation of oxidative stress response genes across cell types. We demonstrate the utility of Kompot in disease settings by identifying cell-state and gene expression changes associated with improved efficacy of combinatorial immunotherapy in melanoma. Additionally, Kompot enables multi-sample comparative analysis by accounting for sample-to-sample heterogeneity. By capturing both global and cell-state–specific effects of perturbation, the Kompot framework is broadly applicable to dissecting condition-specific effects in complex single-cell landscapes.

Introduction

Multi-condition single-cell datasets offer unprecedented opportunities to dissect how mutations, treatments, or perturbations affect differentiation trajectories and disease phenotypes[1-4]. Systematic analyses of such datasets require characterization of two key aspects of condition-induced variability: differential abundance (DA), which tracks changes in cell population frequencies (**Fig. 1A, B**), and differential expression (DE), which identifies molecular responses within those populations (**Fig. 1A, C**).

Differential abundance refers to differences in the composition of cell states between two conditions (**Fig. 1A, B**). Even in the absence of overt phenotypic shifts such as loss or gain of broad cell types, changes in the frequency of specific cell states can provide early indicators of disease[5, 6] or response to treatment[7, 8]. These changes are often nuanced and cell-type specific, necessitating statistical approaches that rigorously assess significance while accounting for underlying biological variability.

Perturbations can also reshape the phenotypic landscape itself through differential expression, where gene expression changes that would be unlikely or impossible in unperturbed conditions can emerge[9] (**Fig. 1A, C**). These transcriptional responses reveal condition-specific molecular programs and regulatory networks that often manifest as subtle yet systematic changes across multiple states[9]. However, these signals can frequently be masked by the inherent variability within biological systems or coarse discretizations.

Several computational approaches have been developed for comparative analysis of multi-condition single-cell data with differential abundance receiving greater attention[10-15]. A number of these methods rely on clustering cells into discrete groups to stabilize variability estimates. While facilitating interpretation, such discretizations introduce a critical trade-off between resolution and statistical power. Broad clusters often mask continuous changes or gradients (**Supp. Fig. 1A**), whereas finer resolution clusters suffer from reduced power to reliably estimate mean expression. These limitations make it difficult to integrate differential expression signals across scales, blurring the distinction between strong but localized changes and weaker trends that span broader regions of the phenotypic space. As a result, conclusions can become sensitive to the chosen clustering resolution and other user-defined parameters. Furthermore, relying on discrete groups defined by the data itself can exacerbate statistical circularity, or "double dipping"[16, 17], and many existing methods lack principled significance measures to distinguish genuine biological signals from technical noise.

Here, we introduce Kompot, a novel computational framework that provides a continuous and statistically rigorous approach for differential abundance and differential expression testing using multi-condition single-cell data. Kompot leverages Gaussian Process (GP) modeling to estimate continuous functions across the cell-state space. This approach enables the quantification of differential abundance and differential expression at single-cell resolution while incorporating uncertainty estimates to provide statistically calibrated comparisons. Kompot differential expression testing uniquely enables the detection of structured gene expression changes across the phenotypic landscape, when the changes are localized to specific cell-states, heterogeneous across cell-states, or even when changes occur in opposing directions in different cell types. By addressing both aspects of multi-condition analysis within a unified statistical framework, Kompot provides a comprehensive toolkit for extracting biological insights from increasingly complex single-cell datasets across a range of biological systems.

We applied Kompot to characterize cell-state and gene expression changes in murine hematopoietic aging, identifying that hematopoietic stem cells not only increase in abundance with age but acquire a state that is less primed towards differentiation. We further demonstrate the utility of Kompot to characterize global gene expression resulting from a single transcription factor knockout during murine embryogenesis and by determining the cell-state and gene expression changes that lead to better responses to combination immunotherapy in melanoma. We extended Kompot to determine sample variability and highlight how the method can determine robust abundance changes in neuroblastoma

disease trajectories while accounting for sample-to-sample heterogeneity. Further, we undertook extensive robustness analyses and demonstrate that Kompot results are robust to a wide range of parameters and scale near-linearly with number of cells and genes.

Results

Overview of the Kompot modeling framework

A single-cell RNA-seq dataset provides a snapshot of cellular states in a biological system, with the molecules measured representing a sample of the transcriptome of the cell. To characterize the system and compare it to similar systems under different conditions (such as mutations or other perturbations), we need to understand what this data tells us about the true composition of transcribed genes in a cell and the true distribution of cell states in the system. With Kompot we leverage high-dimensional, continuous cell-state representations and Bayesian Inference to learn both, the composition of the system and transcriptomes of each cell, enabling a holistic comparison of biological systems at single-cell resolution with uncertainty quantification. This comparison is divided into differential cell-state abundance that reflect changes in cell-state composition, and differential-expression that reflect changes in the transcriptome of cells in equivalent states.

The input to Kompot is a latent representation of co-embedded multi-condition single-cell data. This latent representation can incorporate batch effect correction[18, 19] to ensure that cell states considered equivalent share similar locations in that state space (**Supp. Fig. 1B**). From this foundation, Kompot employs diffusion maps[20] to derive a cell-state representation with biologically meaningful structure (**Fig. 1A**). Diffusion maps are particularly effective for single-cell analysis as they capture the intrinsic geometry of the data while providing biologically meaningful measures of cell-to-cell distance[21, 22]. This distance metric supports both accurate local abundance estimation and principled information sharing when mapping from cell-state space to gene expression, enabling the statistical comparisons at the core of Kompot's differential testing framework.

Kompot differential abundance testing

To quantify differences in cell-state composition between conditions, Kompot first constructs condition-specific density functions on the diffusion map representation. For each condition, Kompot computes a separate continuous density function (**Fig. 1B, Supp. Fig. 2A-B**) using our Mellon framework[23]. Mellon leverages nearest-neighbor distances within a Gaussian Process model to estimate continuous density functions that capture the distribution of cells across the state space[23]. The continuous nature of these density estimates enables evaluation at any position within the cell-state space, including states predominantly observed in the alternate condition (**Supp. Fig. 2B**). By computing the ratio of density estimates at corresponding positions in the cell-state space, Kompot derives a measure of differential abundance that quantifies the fold-change in cell density between conditions at single-cell resolution (**Supp. Fig. 2C**).

While the ratio of density estimates provides a measure of fold-change, rigorous comparison between conditions requires statistical assessment of significance. Kompot addresses this challenge by computing uncertainty estimates for each density function. These uncertainty estimates account for two critical factors: limited data availability in regions of the cell-state space and inherent uncertainty through the sparse function representation. By incorporating these uncertainty estimates into the statistical framework, Kompot calculates posterior tail probabilities (PTPs), an alternative for p-values, for differential abundance at each cell state. This enables the generation of volcano plots for differential abundance testing where each point represents a single cell-state (**Fig. 1B**), allowing for threshold-based significance assessments. Thus, Kompot can identify subtle cell-state transitions and rare subpopulations

that would likely be undetected by clustering-based approaches, providing insights into the continuous spectrum of cellular responses to perturbations (**Fig. 1D**).

Kompot differential expression testing

For differential expression testing between conditions, Kompot constructs condition-specific mappings from cell-states to gene expression space (**Fig. 1A, C**). These mappings are computed using Gaussian Processes (GPs)[24], by predicting log-transformed gene expression as a continuous function of the diffusion map coordinates. The non-linear relationship between diffusion maps and gene expression is effectively captured by GPs[24], essentially serving as a rigorous approach to gene expression imputation [23]. The continuous nature of these mapping functions allows Kompot to estimate gene expression in one condition for cell states observed in another, producing counterfactual expression profiles that reflect how those same cell states would behave under different conditions (**Fig. 1C**). By computing the ratio between the actual expression in a cell and its counterfactual expression, Kompot quantifies gene expression-fold changes for every gene at single-cell resolution across the entire phenotypic landscape (**Fig. 1C**).

Given the large number of genes evaluated, we require a robust significance measure that allows us to identify the most relevant genes, accounting for both uncertainty in the gene expression function and statistical interdependence of imputed values. Kompot therefore employs the Mahalanobis distance, which measures the distance of a point from a multivariate normal distribution, considering both variance and covariance[25]. It effectively measures how many standard deviations a gene expression shift is away from expected variation, while accounting for the correlation structure between neighboring cell states, making it an ideal significance measure for differential expression. Kompot therefore uses Mahalanobis distances as the score in a quasi-volcano plot (**Fig. 1C**) to identify significantly differentially expressed genes either in the full population of cells or in specific subsets (**Fig. 1D**).

Kompot detects cell-composition changes across conditions without relying on predefined populations

Aging is a compelling biological context for evaluating differential abundance methods, as it induces progressive and often subtle shifts in cellular composition rather than dramatic phenotypic transformations. Hematopoiesis has emerged as a particularly informative system for studying aging[26], characterized by declining stem cell function[27], reduced lymphocyte production[28], and increased susceptibility to myeloid malignancies[29]. To investigate these age-related changes, we generated a novel scRNA-seq dataset of murine bone marrow at 10, 63, and 103 weeks of age, representing young, middle-aged, and older timepoints (**Supp. Fig. 3A**). After quality control verification (**Supp. Fig. 3B, C**), we successfully identified most hematopoietic cell types across all ages (**Fig. 2A, B, Supp. Fig. 4**), providing us with a rich dataset to characterize cell-state changes and gene expression changes in aging hematopoiesis using Kompot.

We applied Kompot differential abundance testing on our aging hematopoiesis dataset to determine age-dependent composition changes (**Fig. 2C-E**). We first derived a unified cell-state representation for all cells using diffusion maps. We then calculated age-specific density functions using the subsets of cells from each timepoint (**Supp. Fig. 5A-B**). We evaluated these individual density functions at all cell-states in the dataset to derive condition-specific densities and associated uncertainties (**Supp. Fig. 5C-D**). This approach enabled pairwise comparisons that quantify age-related changes in cell-type composition (**Fig. 2C-D**). A key advantage of our framework is the ability to set thresholds based on both effect size and statistical significance at single-cell resolution, allowing for identification of cell states with significant differential abundance rather than relying solely on fold-change cutoffs (**Fig. 2C-E**).

Our analysis revealed significant expansions and contractions in distinct cell populations over time. Comparing young (10 weeks old) to old (103 weeks old) mice, we observed an expansion of the hematopoietic stem cell (HSC) compartment and a marked shift from naïve to memory cells in both B-cell and T-cell compartments with age (**Fig. 2F-G)**. The depletion of naïve B-cells and enrichment of memory B-cells in older mice is a well-characterized phenomenon of hematopoietic aging[28]. This shift reflects fundamental changes in differentiation dynamics: naive B-cells rely on a continuous influx from hematopoietic progenitors, whereas memory B-cells accumulate over time due to antigen exposure. The observed decline in naïve B-cells reflects the reduced lymphoid differentiation potential of aged HSCs[30], while the expansion of memory B-cells, including age-associated B-cells, is driven by persistent immune activation[31]. Similarly, the T-cell compartment displayed a significant reduction in naïve T-cells with a concurrent increase in memory T-cells with age, a hallmark of immunosenescence that contributes to diminished adaptive immune responses in aging[32].

An increased accumulation of hematopoietic stem cells (HSCs) in older mice is the most significant change we observed in the progenitor cells (**Fig. 2F-G**). Progressive enrichment of HSCs with age is a well-documented phenomenon[33]. HSCs reside at the top of the hematopoietic hierarchy, maintaining blood production through self-renewal and differentiation into mature lineages. With aging, HSC function declines, leading to accumulation in the bone marrow while their ability to differentiate into different lineages becomes impaired[33]. While the accumulation of aged HSCs has been previously observed, it has traditionally been quantified using bulk measurements or clustering-based approaches. We took advantage of the single-cell resolution of Kompot to better characterize the age-induced change in HSCs. By comparing the differential abundance fold-change along pseudotime ordering[21], we observed that the expansion of HSC population is in fact a subtle shift in cell-state, rather than a global increase in a predefined population (**Fig. 2H**). Specifically, HSCs in older mice progressively shift towards a less differentiation-primed state, creating a continuum of altered stem cell states rather than a uniform expansion (**Fig. 2H**). This subtle but consequential repositioning along the differentiation trajectory provides a possible explanation for the functional decline in aged HSCs, revealing how changes in cell abundance are accompanied by fundamental alterations in cellular function and developmental potential.

Comparing differential abundance results across ages, we observed that changes in mature cell populations predominantly occurred between young and middle-aged (63 weeks old) mice (**Fig. 2D, Supp. Fig. 6A**), while alterations in the HSC compartment were more pronounced between middle-aged and old mice (**Fig. 2E, Supp. Fig. 6B**). This temporal separation suggests a sequential progression of aging, where compositional shifts in differentiated cells precede fundamental changes in stem cell states. The early alterations in mature immune cells may represent initial adaptive responses to aging, while the later HSC changes potentially indicate more persistent reprogramming of the hematopoietic system.

Our results demonstrate the ability of Kompot to identify subtle but biologically meaningful abundance changes between conditions. By combining a continuous representation of cell states with robust statistical measures, Kompot enables precise detection of population-level changes while maintaining single-cell resolution.

Kompot captures context-specific gene-expression shifts

We applied Kompot differential expression testing on our mouse aging hematopoiesis dataset to characterize age-related transcriptional changes (**Fig. 3**). We first inferred age-specific mapping functions from diffusion map coordinates to log-normalized gene expression values, using the subset of cells from each timepoint. We then used these mappings to impute gene expression and associated uncertainty at all cell-states across the dataset. This enabled us to compute gene-level fold changes by comparing observed expression values to their counterfactual predictions under alternative age conditions. To assess significance, we calculated Mahalanobis distances for each gene across specified groups of cells,

quantifying deviations from expected variation while accounting for the covariate structure among neighboring cell states. We excluded T and NK cells for this analysis since they do not differentiate in the bone marrow (**Fig. 3H**).

We first applied our approach to characterize gene expression changes in HSCs, a population known to undergo functional and transcriptional changes with age[33], by comparing HSCs from middle-aged to old-aged mice (**Fig. 3A-B**). Aging in HSCs has been extensively studied, and the Hematopoietic Aging Atlas has synthesized results from multiple datasets to identify genes consistently up- or downregulated with age[34]. For each gene, this atlas assigns a consistency score that reflects the reproducibility of age-associated expression changes across studies[34]. Kompot's differential expression results strongly prioritized genes with the highest consistency scores in the Hematopoietic Aging Atlas, with genes showing large fold changes and Mahalanobis distances highly enriched for reproducible aging signatures (**Fig. 3C**, **Supp. Fig. 7**). Notably, this included up-regulation of the cell adhesion molecule Selp, which has been previously identified as a marker of stem cell aging and is functionally linked to increased oxidative stress and decreased hematopoietic reconstitution[34, 35].

Our differential abundance analysis revealed that HSCs in older mice shift towards a less differentiated state along the cell-state continuum (**Fig. 2H**). We next asked whether the genes differentially expressed in aged HSCs reflected this shift. Indeed, we found that gene expression fold changes were structured along the same continuum: genes exhibited larger fold changes in cell states that were more differentially abundant (**Fig. 3D**). These results highlight the transcriptional changes in HSCs that accompany and may underlie the shift toward a less differentiated state with age.

We next investigated age-related gene expression changes in mature cell types. We focused on changes from young to old mice since the abundance changes in mature cell types were most pronounced in this comparison (**Fig. 2D-F**). To distinguish expression changes from abundance-driven effects, we restricted our analysis to cell types that *did not show significant compositional changes*. Monocytes exhibited clear transcriptional changes despite stable abundance (**Fig. 3E-F)**. Gene ontology analysis revealed that downregulated genes were enriched in core monocyte/macrophage functions such as defense response (**Fig. 3G**), consistent with their reduced functional capacity with age[36]. In contrast, upregulated genes were enriched for antigen processing and presentation pathways, driven by increased expression of MHC class II genes (**Fig. 3G**), which have been implicated in age-related immune remodeling[37]. Notably, our results recapitulate findings from a previous study that identified H2-Aa, H2-Ab1, H2-Eb1, Cd74, and Aw112010 as the top dysregulated genes in aging monocytes across both mice and humans, with only Igkc surpassing H2-Eb1 in our ranking[37]. The near-perfect agreement with experimentally validated gene sets underscores the power of our differential expression testing approach to detect biologically meaningful transcriptional shifts. Gene expression differences in other mature cell types were not as pronounced but did show a tendency for upregulation of MHC Class II genes similar to monocytes (**Supp. Fig. 8**).

Our use of Mahalanobis distance as a significance measure enables Kompot to identify genes that exhibit coordinated but opposing changes across subpopulations. For example, consider a gene that is upregulated in one cell type and downregulated in another. While the net fold change may be close to zero, the Mahalanobis distance remains high, reflecting a statistically significant deviation from expected variation if unaffected. When a gene shows consistent but opposite shifts in related subpopulations, this pattern produces a large deviation relative to the expected correlated variation, resulting in a high Mahalanobis distance. Kompot can therefore detect structured differential expression that would be missed by conventional fold-change-based approaches.

To test this capability, we applied Kompot differential expression testing between young and old cells across all cell-types that differentiate in the bone marrow (**Fig. 3H-I**). The top genes with both high positive fold change and high Mahalanobis distance were strongly enriched for MHC class II genes, consistent

with our monocyte-specific results and indicating that MHC class II upregulation is a general feature of hematopoietic aging (**Fig. 3J**). Likewise, genes with high Mahalanobis distance and strong negative fold change were consistently downregulated across multiple cell types (**Fig. 3J**). In contrast, genes ranked highly by Mahalanobis distance but with low absolute fold change exhibited divergent expression patterns across cell types, with upregulation in some and downregulation in others (**Fig. 3J**). To illustrate these distinctions, we visualized expression levels of representative genes from each category across the hematopoietic differentiation landscape (**Fig. 3K**). Genes selected by fold change showed broad, monotonic shifts in expression, whereas Mahalanobis-selected genes revealed distinct, cell-type-specific regulation (**Fig. 3K**).

We next performed gene ontology analysis to assess the biological relevance of genes with high Mahalanobis distance but low absolute fold change. Antioxidant activity emerged as the most significantly enriched term (**Fig. 3L**), a pathway known to be impacted by aging[38]. Notably, our results indicate that while antioxidant activity is broadly impacted across hematopoietic cell types, the direction of expression changes of the pathway genes varies across genes and cell types, suggesting a context-specific adaptation to aging (**Fig. 3M**). Oxidative stress is a known driver of HSC dysfunction, promoting DNA damage, apoptosis, and biased differentiation toward myeloid lineages[38]. The differential expression of antioxidant pathway genes observed in our study likely represents a compensatory mechanism to mitigate reactive oxygen species (ROS) accumulation in aged HSCs[38].

These results demonstrate that Mahalanobis distance captures biologically meaningful heterogeneity in gene expression and provides a powerful approach for aggregating structured expression differences across cell states without requiring prior clustering, making it especially suited for complex, continuous single-cell landscapes.

Global Gene Expression Changes due to Tal1 mutation in Mouse Embryonic Development

As a further test of Kompot's Mahalanobis Distance as a metric to detect expression differences across cell states without clustering, we applied Kompot to a single-cell RNA-seq dataset of mouse embryos at mid-gestation with wild-type and mutant *Tal1*[39] (**Supp. Fig. 9A, B**). *Tal1* is required for the generation of primitive erythroid, megakaryocyte, and myeloid lineages[39], but its effects on other lineages is less characterized. We therefore performed differential expression analysis between *Tal1* wildtype cells to *Tal1*$^{-/-}$ cells specifically within lineages not affected by abundance changes (**Supp. Fig. 9A, B**). As expected, hemoglobin genes showed the greatest changes in the mutant cells along with sex-associated genes such Xist (**Supp. Fig. 9C**). These signals likely reflect technical and experimental factors: hemoglobin transcripts contribute substantially to ambient RNA in single-cell data, and Tal1 mutant chimeras were generated exclusively from female donor cells, resulting in differential detection of imprinting-associated genes (**Supp. Fig. 9C**). After accounting for these confounders, we next checked whether Kompot highlights biologically meaningful differences in the remaining lineages.

We excluded the hemoglobin and imprinting-related genes and selected the top 500 differentially expressed genes ranked by Mahalanobis Distance for downstream analysis (**Supp. Fig. 9D**). Gene ontology analysis revealed significant enrichments for processes such as vasculature development and blood vessel development (**Supp. Fig. 9E**). Genes associated with these processes showed pronounced expression differences between wild-type and mutant cells, particularly within endothelial and hematoendothelial progenitor populations (**Supp. Fig. 9F)**, consistent with the known role of Tal1 in regulating these programs[39]. Further, the magnitude and direction of these changes varied across related cell states (**Supp. Fig. 9F, G**), suggesting a context-dependent role for Tal1 in modulating vasculature-associated gene expression.

Pathway enrichment analysis of the same gene set also highlighted effects of Tal1 loss beyond hematoendothelial populations (**Supp. Fig. 9H**). Genes associated with fluid shear stress pathways were broadly downregulated across multiple lineages (**Supp. Fig. 9I**), suggesting that blood flow during embryogenesis influences not only nutrient delivery but also the transcriptional state of developing tissues. In addition, we observed widespread changes in cyclin family gene expression following Tal1 knockout (**Supp. Fig. 9J**). Although Tal1 has been implicated in cell cycle regulation of hematoendothelial cells, these results suggest more indirect effects on cell cycle–associated programs across the embryo.

Rather than performing differential expression on a lineage-by-lineage or cluster-by-cluster basis, Kompot tests for structured expression differences across all cells simultaneously. This enables detection of both lineage-specific programs, such as vasculature-associated changes in endothelium, and broader patterns, such as altered cell cycle and shear-stress pathways across multiple lineages. Thus, these results illustrate how Kompot differential expression can recover biologically meaningful changes at multiple levels of organization without requiring predefined clusters.

Kompot reveals determinants of immunotherapy in advanced melanoma

We next tested effectiveness of Kompot in disease contexts and applied Kompot to examine how immune checkpoint blockade therapy reshapes CD8 T cell states. CD8 T cell activation and exhaustion span a continuous spectrum[40], making this a natural test case for Kompot's ability to quantify abundance and expression differences without imposing cluster boundaries. Combination treatment with anti-LAG-3 (Relatlimab, Rela) and anti-PD-1 (Nivolumab, Nivo) has been shown to provide greater clinical efficacy than nivolumab alone in advanced melanoma[41]. Cillo et. al.[41] profiled CD8 T cells from patients in a clinical trial of these therapies using single-cell RNA-seq (**Fig. 4A**) and identified that combination therapy promotes an altered differentiation trajectory with increased cytotoxicity and exhaustion contributing to improved efficacy[41]. Using Kompot, we reanalyzed baseline and on-treatment samples within each therapy arm (**Fig. 4B**) to determine whether our cluster-free, continuous framework could recover complementary insights to those reported previously.

Kompot differential abundance analysis revealed significant but heterogeneous shifts across treatment arms relative to baseline (**Fig. 4C, Supp. Fig. 10**). In Nivo monotherapy, we observed a decrease in terminally exhausted cells, whereas Rela treatment led to an increase in TCR-activated exhausted states (**Fig. 4D, Supp. Fig. 10C**). The most prominent change under combination therapy was an expansion of activated cycling CD8 T cells (**Fig. 4D, Supp. Fig. 10C**). Given their central role in tumor cell killing[42], the expansion of activated cycling CD8 T cells represents a striking cellular correlate of the enhanced efficacy observed with combination therapy in advanced melanoma. We therefore next asked whether these therapy-specific abundance shifts detected by Kompot are mirrored in transcriptional programs across the CD8 T cell landscape.

T cell activation and exhaustion trajectories are continuous, producing related but heterogeneous states that complicate comparisons of expression programs using conventional approaches[40]. Kompot addresses this by quantifying covarying differences without clustering or discretization. Using Kompot differential expression, we identified the top 250 genes altered between baseline and treatment within each therapy arm (**Supp. Fig. 11A**). The union of these genes showed broad changes across therapies (**Fig. 4E**), with fold changes generally consistent across states within each arm (**Fig. 4F**). Nivo monotherapy more closely resembled combination therapy than Rela alone (**Fig. 4F**). Interestingly, activated cycling and effector states diverged the most, showing expression changes distinct from those induced by combination treatment (**Fig. 4F**). These results suggest that treatment-specific differences in efficacy may, in part, reflect cell-state–specific transcriptional responses within these key effector populations.

To further interpret these transcriptional differences, we performed gene ontology analysis of the top differentially expressed genes (**Supp. Fig. 11B**). This revealed significant enrichment for translation and T cell activation programs (**Supp. Fig. 11C**). T cell activation genes were upregulated in activated cycling cells under combination therapy but downregulated in the same states under Nivo monotherapy (**Fig. 4G-I, Supp. Fig. 11D**). Ribosomal genes displayed similarly opposing patterns (**Supp. Fig. 11E**). These results indicate that immunotherapy regimens elicit distinct and even opposing gene expression changes within the same effector populations, providing a mechanistic basis for the differential efficacy observed across treatments.

Together, these results show that Kompot provides an effective solution to the challenge of analyzing therapy-induced changes in continuous CD8 T cell landscapes. By jointly testing for abundance and expression differences without predefined clusters, Kompot recovered coherent, state-specific signals across overlapping populations. This underscores Kompot's utility for resolving condition-specific programs in single-cell datasets in which subtle, state-specific changes are distributed across continuous manifolds.

Resolving inter-patient variability with multi-sample differential testing

Multi-sample, multi-condition single-cell datasets present both an opportunity and a challenge: they capture shared therapeutic signals while also reflecting patient-to-patient variability. Most approaches tradeoff between robustness and resolution, either aggregating away heterogeneity or losing power across samples. Kompot addresses this by extending differential abundance and expression testing to the multi-sample setting, estimating condition-specific variance across samples at single-cell resolution and condition-specific covariance across samples at single-gene resolution (**Fig. 5A**). This enables statistically principled detection of both shared patterns and patient-specific differences in large single-cell cohorts.

We applied Kompot to a recently published single-cell atlas of high-risk neuroblastoma[43], comprising over 200,000 neoplastic cells from 22 patients profiled at diagnosis and after treatment (**Fig. 5B-E, Supp. Fig. 12A**). In this study, neoplastic cells were harmonized and assigned to six transcriptional states (**Fig. 5B, C**). Broad therapy-associated shifts in state proportions between diagnosis and treatment were reported, including increases in adrenergic (ADRN)-calcium, ADRN-dopaminergic, Interm-OXPHOS, and MES states and decreases in ADRN-baseline and ADRN-proliferating states[43]. Using Kompot, we recovered these reported shifts (**Fig. 5F, G**) and further resolved heterogeneity within ADRN-baseline and ADRN-proliferating states: while most patients showed decreases after treatment, a subset—particularly those with ALK mutations—displayed increases instead (**Fig. 5H, Supp. Fig. 12B**).

To test whether this patient-level heterogeneity reflects underlying biology, we categorized patients as *Event Free* if they remained free of progression, relapse, or death and as *Event* otherwise (**Fig. 5E**). We then ordered patients by the percentage of states showing significant differential abundance in ADRN-baseline and ADRN-proliferating states (**Fig. 5H)**, the adrenergic programs that dominate at diagnosis and whose persistence has been linked to poor outcomes[43]. We observed that patients with significant reductions in these states were almost uniformly Event Free, with patient 14 as the sole exception. Conversely, patients who either retained these states without significant changes in any cell state, or exhibited increases in these states were all categorized as Event (**Fig. 5H)**. Moreover, all Event Free patients displayed either reductions in these aggressive states or, in the case of patients 17, 20, and 22, compensatory increases in alternative states (**Fig. 5H**), suggesting treatment-induced shifts away from high-risk populations.

Kompot's multi-sample analysis therefore recapitulates broad, cohort-level trends while also resolving patient-specific deviations that align with clinical outcomes. By capturing both shared and heterogeneous

abundance changes within a single unified framework, Kompot provides a general approach for connecting variability in large-scale single-cell studies to biologically and clinically meaningful differences.

Kompot robustness, benchmarking and performance

We next performed robustness analyses of Kompot differential abundance and expression results to ensure confident interpretation of results (**Supp. Figs. 13-16)**. We varied each parameter over a broad range of values and compared the results to baseline using Pearson correlation. We performed all our robustness checks using our murine aging hematopoiesis dataset as representative of a dataset with continuous trajectories and comparison of healthy donors and COVID-19 patients[44] for changes in composition and expression changes in peripheral mononuclear blood cells (PBMCs) as representative of a dataset with discrete clusters.

Number of cells in a dataset is often a technical variable with tissue handling and costs being the predominant factor in the choice of dataset size rather than any representation of underlying biology. We therefore down sampled cells in both datasets and determined the correlation of Kompot differential abundance log fold change between down sampled and the full dataset (**Supp. Fig. 13A, 14A).** Kompot fold changes are robust even when the number of cells is reduced by an order of magnitude. We next varied the number of diffusion components which is a user guided parameter that accounts for the biological complexity of the dataset (**Supp. Fig. 13D, 14D)**. Kompot differential abundance fold changes are robust to large changes in the number of components demonstrating the results are not sensitive to the choice of number of diffusion components. Finally, we tested robustness to two parameters of Mellon density estimation: length scale factor which controls the extent to which to share information between neighboring states and number of landmarks, which is used for sparse Gaussian Processes for efficiency. Differential abundance fold changes are robust to large changes in either of these parameters suggesting that default parameters provide reliable, interpretable results (**Supp. Fig. 13B-C, 14B-C).**

We next performed similar comparisons using differential abundance posterior tail probabilities which is the significance measure (**Supp. Figs. 16, 17).** As expected, the significance measure consistency dropped with number of cells (**Supp. Figs. 16A, 17A)** but was robust to all other parameters including length scale (**Supp. Figs. 16B, 17B)**, number of landmarks (**Supp. Fig. 16C, 17C)** and number of diffusion components (**Supp. Figs. 16D, 17D)**,

We undertook similar robustness analyses for Kompot differential expression results using the same murine aging hematopoiesis and COVID-19 datasets (**Supp. Figs. 17-20).** We computed correlations across genes rather than cells in these experiments. In addition to the parameters for differential abundance testing, differential expression involves an additional parameter sigma, that encodes the expected noise in the manifold to gene expression mapping function. We determined that differential expression results (both gene expression log fold change and Mahalanobis distance) are robust to large changes across all parameters (**Supp. Figs. 17-20).** Notably the Mahalanobis distance metric for assessing the significance of differential expression was also robust to reductions in number of cells (**Supp. Figs. 19A, 20A).** These results collectively highlight the robustness of both differential abundance and differential expression results across continuous and discrete datasets.

We next benchmarked the performance of Kompot relative to other approaches. Since Kompot uniquely offers the capability of structured differential expression testing, we focused our benchmarking on differential abundance analysis. Recent benchmarking studies have highlighted Meld[10] and Milo[11] as amongst the top-performing and most widely used differential abundance analysis approaches[12]. We therefore compared the performance of Kompot differential abundance analysis in comparison to Meld and Milo using the simulated (linear, cluster, branch) and covid-19 dataset[44] described in the benchmarkDA study[45]. Kompot performance in both AUROC and AUPRC measures were comparable to both Meld and Milo with Kompot showing greater robustness to batch effects (**Supplementary Fig.**

21). While Meld was the top performing method in these datasets with no introduced batch effects (**Supplementary Fig. 21**), differentially abundance analysis with Meld does not support a statistical significance measures that can account for sample variance. We note that Kompot combines the individual advantages of both Meld and Milo in providing a single-cell resolution differential abundance framework with confidence measures.

Run-time and memory scalability for important practical considerations for the utility of computational algorithms. We benchmarked both metrics for Kompot differential abundance and expression for datasets with varying sizes (**Fig. 6, Supp. Figs. 22-23**). With a fixed number of landmarks, Kompot exhibits near-linear scaling with the number of cells for both differential and expression analyses (**Fig. 6A-B**). Further, the run-time for differentially expression analyses is nearly independent of the number of genes with number of cells playing a greater role (**Fig. 6B, E**) and memory requirements scale approximately linearly in both (**Supp. Fig. 22**). As expected, higher number of landmarks increases compute and memory requirements (**Fig. 6C, D**), though Kompot results remain robust across a wide range of landmark counts (**Supp. Figs. 13-20 (D)**), making the default of 5000 landmarks suitable for most datasets. Run-times and memory usage are otherwise stable across other parameter choices, including the number of diffusion components and the GP length-scale and sigma parameters (**Supp. Figs. 22-23**).

Discussion

Kompot provides a flexible statistical framework for quantifying condition-specific effects in single-cell data, modeling both cell abundance and gene expression as continuous functions over the cell-state manifold. By estimating condition-specific densities and expression profiles using Gaussian Processes, Kompot performs differential analysis at single-cell resolution while accounting for uncertainty and structured variation along the phenotypic manifold. This approach supports biologically grounded interpretations of perturbation effects, whether they manifest as global shifts in population structure or localized transcriptional responses where subtle, state-dependent changes stand out even against large, heterogeneous variability along the manifold. Further, Kompot also supports estimation of sample variability to assess the significance of differential analysis enabling identification of robust changes between conditions while accounting for sample-to-sample heterogeneity.

We anticipate that the Kompot framework will be broadly applicable to existing and newly-generated multi-condition single-cell datasets to gain new insights about both cell population shifts and gene expression changes resulting from perturbation. We have demonstrated the utility of Kompot through application to diverse biological systems including murine hematopoietic aging, immunotherapy response, changes in mouse embryonic development resulting from gene knockouts and neuroblastoma treatment trajectories. In each case, we were able to both recover well-known behaviors of the respective biological systems highlighting Kompot's ability to infer robust biological findings and determine novel insights through unique capabilities of Kompot. Kompot results are also robust to large changes in parameters, scale near linearly with number of cells and are competitive to state-of-the-art methods in benchmarking studies.

While a number of frameworks have been published for differential abundance analysis of multi-condition single-cell datasets, Kompot is one of the first to present a formal statistical framework to perform differential abundance analysis at single-cell resolution. The posterior tail probability of Kompot enabled us to identify significant changes in both progenitor and mature hematopoietic compartments with aging in mice and identify increased abundance of activated CD8 T cell compartment as the likely predictor of greater response to combinatorial immunotherapy treatment in melanoma. Further, by combining single-cell resolution differential abundance effect size with pseudotemporal ordering, we were able to predict that aged hematopoietic stem cells not only increase in abundance but also progressively shift their state

to become less primed towards differentiation consistent with the observed dysfunction of hematopoietic stem cells with age.

Kompot differential expression framework enables capabilities that have been inaccessible to date. By computing the joint posterior distribution of the gene-expression function across statistically linked cell-states and utilizing this distribution's covariance structure to assess the combined significance of change across multiple cell states, Kompot is uniquely able to identify subtle but consistent gene expression changes across the phenotypic manifold even when the changes are in opposite directions in different regions of the manifold. Further, explicit utilization of these covariances also leads to increased robustness of differential expression analysis to compare gene expression shifts in specific sub populations of cells. Kompot's ability to infer structure gene expression changes enabled us to highlight the importance of antioxidant pathway in aging through diverse gene expression changes across hematopoietic populations, determine systematic changes in embryonic development through changes in cell cycle and laminar flow resulting from *Tal1* mutation and characterize the diverse response of activated CD8 T cells in mono or combination immunotherapy in melanoma. Kompot provides a fundamentally new approach for differential expression testing between conditions: Rather than discretizing the cellular phenotypic manifolds into clusters, differential expression analysis can be performed across the full landscape and changes can then characterized or interpreted at the discrete cell type level taking advantage of the single cell, single gene resolution of Kompot.

It is increasingly common to undertake population level single-cell profiling with tens to hundreds of individuals per condition. The continuous nature of the GP functions underlying Kompot enables it to seamlessly determine condition-specific sample variability at single cell and single gene resolution for both differential abundance and expression analyses. This enables a framework to identify robust changes between conditions while accounting for sample-to-sample heterogeneity. This not only enabled us to determine the changes in neoplastic cell compositions at diagnosis and treatment in neuroblastoma but also identify the determinants of treatment response by taking advantage of single-cell resolution of Kompot framework.

Kompot has certain limitations that should be taken into consideration for applications. First, as with any Gaussian Process–based Bayesian model, a degree of model-selection uncertainty remains, most notably in the choice of covariance kernel and its parameters. Second, Kompot depends on the quality and alignment of the latent cell-state representation, which can influence differential abundance and expression when the two conditions occupy shifted regions of the manifold. Third, Kompot assumes that log-transformed single-cell expression measurements are approximately, and locally, normally distributed around an unknown cell-state-specific mean rather than explicitly modeling the full count distribution for each gene. These limitations and their practical implications are described in detail in **Supplementary Note 5**.

Methods

Kompot framework

Kompot is a toolkit for comparing single-cell phenotypic manifolds across conditions. It provides frameworks for both differential abundance and differential gene expression testing at single-cell resolution without requiring predefined cell-types or clusters. The guiding assumption is that both cell-state abundance and gene expression change continuously along cell-state space[23].

The input to Kompot is a latent representation of co-embedded multi-condition single-cell data (e.g. PCA, Harmony[18], scVI[46], LSI[47]). When needed, upstream batch correction is used to align equivalent biological states across conditions so that cells of similar state occupy similar locations in latent space. This alignment does not remove condition-specific gene-expression differences; rather, it standardizes the coordinate system so Kompot can compare cells we deem equivalent by proximity in state space[19].

Diffusion map representations are used as the default cell-state representation because Euclidean distances in diffusion components capture biologically meaningful differences between cells[20, 22, 48], but Kompot can operate on any embedding with a well-defined distance metric.

Kompot models condition-specific densities and expression profiles with Gaussian processes, enabling differential analysis that accounts for uncertainty and structured variation along the phenotypic manifold. This allows detection of consistent, state-dependent abundance or expression shifts that may be smaller than the feature's global variability, including cases where local effects differ in direction (e.g., reduced dynamic range). By explicitly estimating posterior uncertainty via variational inference, Kompot provides principled significance measures for its findings at minimal additional computational cost (see scaling analysis).

Differential Abundance Testing

Kompot differential abundance testing enables the detection of shifts in the abundance of cell-states across the phenotypic manifold. Using the co-embedded diffusion map representation, Kompot first constructs condition-specific density functions for each condition using Mellon[23]. These density functions describe how cells from each condition are distributed across the phenotypic manifold. Therefore, we can derive measures of differential abundance by comparing how these densities change between two conditions. Finally, since density functions are continuous, Kompot can perform these comparisons across the entire phenotypic landscape, enabling cluster- or cell-type–free differential abundance testing at single-cell resolution.

Our density quantification relies on Gaussian Process (GP) estimation implemented in Mellon[23]. While Mellon was previously introduced as a framework for continuous density modeling, here we extend it with an explicit uncertainty estimation at each cell state. This Bayesian inference procedure yields both a "best estimate" function that maps each point in the cell-state space, x_i to a log density value $\mu(x_i)$, and a quantification of uncertainty $\sigma(x_i)$ separately for each condition. Theoretical descriptions of the source of uncertainty and implementation are in **Supp. Notes 1**.

By comparing two conditions a and b, the posterior distribution of the change in density any cell state x_i is given by

$$\delta(x_i) \sim \mathcal{N}\left(\mu_b(x_i) - \mu_a(x_i), \sigma_a^2(x_i) + \sigma_b^2(x_i)\right).$$

Therefore, the density log-fold change between conditions is defined as $\mu_b(x_i) - \mu_a(x_i)$, with variance $\sigma_a^2(x_i) + \sigma_b^2(x_i)$.

Normalization for comparability across conditions

Mellon density estimates are in the units of log number of cells per volume and as a result, direct comparisons of densities between conditions with different number of cells will not be accurate. Therefore, Kompot transforms these values into log fraction of cells per volume by subtracting the log total number of cells using

$$\mu'_a(x_i) = \mu_a(x_i) - \log N_a$$

where N_a represents the total number of cells measured for condition a. This normalization ensures that density values reflect the relative occupancy of the phenotypic manifold, independent of number of cells. For this to be valid, Kompot assumes that the intrinsic dimensionality of the manifold is estimated correctly. Mellon provides this estimate, and it is treated as constant across the dataset. Thus, the normalized local log-fold change representing abundance difference is

$$\Delta(x_i) = \mu_b(x_i) - \mu_a(x_i) - \log {}^{N_a}/_{N_b}\,.$$

Significance testing

To assess whether observed changes are statistically meaningful, Kompot computes a posterior tail probability (PTP). Intuitively, the PTP measures the probability that the true density difference has the opposite sign to the estimated change, $\Delta(x_i)$ i.e., if $\Delta(x_i)$ is positive, the PTP gives the probability that the true change is negative, and vice versa. Specifically, if the density change $\Delta(x_i)$ is positive, then we compute $\mathbb{P}(\delta(x_i) < 0)$, and $\mathbb{P}(\delta(x_i) > 0)$ when $\Delta(x_i)$ is negative. Since the posterior distribution is a normal distribution, the local Kompot PTP is given by:

$$PTP(x_i) = \Phi\left(\frac{-|\Delta(x_i)|}{\sqrt{\sigma_a^2(x_i) + \sigma_b^2(x_i)}}\right)$$

were Φ is the cumulative distribution function of the standard normal distribution. Thus, Kompot provides both an effect size (density log-fold change) and a principled measure of statistical significance (PTP) for every cell state.

Differential Expression Testing

Kompot differential expression testing enables the identification of gene expression changes between two conditions across the same manifold. The goal is to detect genes with significant differences between conditions, rather than between cell types or clusters within a single condition. Specifically, Kompot is designed to capture subtle, structured transcriptional shifts that vary across cell states, even when those shifts do not follow a consistent direction. Similar to differential abundance testing, Kompot first constructs condition-specific expression functions that map diffusion maps to observed expression values for each gene in each condition using Mellon[23]. This enables testing for gene expression changes across the entire phenotypic landscape without predefined clusters, by explicitly sharing information across neighboring cell states to detect structured patterns.

Similar to the density function, we leverage Mellon's GP implementation for gene expression functions (**Supp. Note 1**). For each condition, Kompot compute a gene-expression function that predicts an expression value for each gene at any location in the cell-state space, specific to that condition. These

functions provide a smooth and continuous representation of gene expression across the manifold and allow for quantifying differential expression between conditions at any point in state space, effectively integrating all data points without requiring predefined cluster boundaries. The resulting gene-expression functions and uncertainty estimates can be used analogously to the cell-state density functions to compute local gene-expression differences, with local fold change and posterior tail probabilities, providing both effect size and uncertainty for each gene.

Beyond asking where in the manifold a gene is differentially expressed, we also want to know whether it shows consistent differences within a cell population or across the entire manifold. Unlike cell-state densities, GP-based expression functions are an imputation that already aggregate information across cells since they map state space to observed expression. Specifically, for gene g, the gene expression values from function $f_g(x_i)$ at different cell states x_i are not independent measurements that can be aggregated directly. Kompot resolves this by explicitly modeling covariance across states using Mellon's uncertainty framework. Mellon provides not only variances for each predicted expression value, but also covariances that reflect the smoothness of gene expression across the manifold i.e., the uncertainty covers not only the variance for a function value $f(x_i)$ but also the covariance to other function values $f(x_j)$ based on the similarity of cell states x_i and x_j.

Structured DE testing with Mahalanobis distance

Formally, under our modelling assumption, the true gene expression values $f_g(x)$ at cell states $x = \{x_i\}_{i=1}^{m}$ are distributed according to a multivariate normal posterior

$$f_g(x) \sim \mathcal{N}\left(\mu^{(g)}(x), \Sigma^{(g)}(\mathrm{x})\right).$$

Therefore, difference in functions between the conditions a and b also follows a normal distribution:

$$\delta_g(x) \sim \mathcal{N}\left(\mu_b^{(g)}(x) - \mu_a^{(g)}(x), \Sigma_{\mathrm{b}}^{(g)}(\mathrm{x}) + \Sigma_{\mathrm{a}}^{(g)}(\mathrm{x})\right).$$

To summarize these correlated local differences into a single test statistic, Kompot uses the Mahalanobis distance (**Supp. Note 2**):

$$D(a,b) \quad = \sqrt{(\mu_a - \mu_b)^T (\Sigma_a + \Sigma_b)^{-1} (\mu_a - \mu_b)}$$

This statistic accounts for the covariance between nearby states, ensuring that consistent changes are emphasized while redundant information is down-weighted. As a result, Kompot can detect structured patterns of differential expression across the manifold, even when local changes differ in magnitude or direction.

Thresholds to identify significantly different genes

To enable robust identification of significantly different genes, Kompot generates an empirical null distribution of Mahalanobis distances by randomly shuffling gene expression values (default: 2000 genes). From the resulting distribution of null Mahalanobis distances, empirical p-values are computed and used to estimate false discovery rates (FDR) via the local FDR methods[49]. By default, Kompot applies with a cutoff of 0.05 on the local FDR, yielding a direct Mahalanobis distance threshold for declaring genes significantly differentially expressed (**Supp. Note 4**).

Multi-Sample Framework

Kompot also supports multi-sample analyses by explicitly incorporating variability across replicates (**Supp. Note 3**). The density functions and gene-expression manifolds are always inferred jointly across all samples, which maximizes statistical power and prevents small or unbalanced samples from distorting the overall estimate. To capture heterogeneity between replicates, Kompot then incorporates an empirical variability term that reflect differences within each condition. This combined uncertainty accounts for both biological variability (e.g., donor or tissue differences) and technical variability (e.g., sample preparation or sequencing batch effects). By integrating these sources of variability, Kompot provides a robust and conservative basis for identifying significant changes in both cell-state densities and gene-expression patterns.

Scalability

Kompot inherits the scalability of the Mellon framework through its sparse Gaussian process implementation, enabling efficient inference across datasets spanning several orders of magnitude in size. We systematically benchmarked runtime scaling across four representative datasets: *aging* (~8,000 cells, 16,000 genes), *BCR-XL* (CyTOF dataset of ~172,000 cells, 24 proteins)[50], *COVID-19* (~422,000 cells, 34,000 genes)[44], and a *post-COVID* variant comprising the same cells but distinct condition comparisons.

To ensure consistent memory reporting, all memory values correspond to the maximum resident set size (MaxRSS) recorded by Slurm. MaxRSS robustly captures total host RAM usage across all components involved in execution, including pure Python allocations and memory consumed by NumPy, JAX, and Dask kernels. Importantly, MaxRSS does not include GPU device memory; VRAM usage during accelerated runs was therefore not tracked in these benchmarks, and the reported values reflect only CPU-side memory consumption.

Scaling with number of cells

Runtime is primarily governed by the number of cells in the compared conditions. Scaling with number of cells were determined using default parameters including 5000 landmarks. When the number of cells exceeds the number of landmarks (default: 5,000), the sparse Gaussian process formulation achieves approximately linear scaling with number of cells (**Supp. Figs. 13-20 (A)**). While DA runtime depends almost exclusively on the number of cells and landmarks, DE additionally scales with the number of genes analyzed.

Scaling with number of landmarks

Runtime increases quadratically with the number of landmarks k, consistent with the expected $\mathcal{O}(k^2 n)$ complexity of sparse Gaussian processes for a number of cells n (**Figs. 6B, D**). The quadratic trend becomes apparent above ~1,000 landmarks, once constant overhead is exceeded. When the landmark count surpasses the number of cells, Kompot reverts to a full Gaussian process, resulting in constant runtime independent of further increases in k. DE was performed with gene batching, resulting in the curves diverge slightly since gene batching reduces memory requirements at the cost of increased per-batch computation time.

Scaling with number of genes and parallelization

DE runtime depends only weakly on the number of genes (**Fig. 6E**). The total complexity is $\mathcal{O}(nk + nkg)$, where n is the number of cells, k the number of landmarks, and g the number of genes. The shared $\mathcal{O}(nk)$ operations (covariance construction and decomposition) are performed once, while the per-gene $\mathcal{O}(nkg)$ term scales linearly with g. Each gene is modeled by an independent Gaussian process with closed-form inference, making computation trivially parallelizable. The extended DE model that explicitly accounts for sample-to-sample variance increases runtime due to $\mathcal{O}(k^2 g)$ complexity, as each gene requires construction of a gene-specific covariance matrix capturing sample effects.

Landmark generation

As in Mellon[23], Landmarks are computed using k-means clustering, which scales as $\mathcal{O}(tkdn)$ with t iterations, k landmarks, d dimensions, and n cells. For very large datasets, this step can become a bottleneck. Kompot therefore allows users to provide custom landmarks, enabling faster or approximate alternatives. Because inference relies on covariance statistics, the exact landmark placement has limited influence as long as landmarks adequately cover the state space. Redundant landmarks contribute negligibly to test statistics, as their strong covariance prevents inflation of significance.

Peak memory usage and runtime analysis

Peak memory usage was measured using SLURM's sacct utility (MaxRSS), which captures the maximum resident set size across all associated processes (JAX, Dask, NumPy, Python). Runtime was measured as the total execution time of Kompot on the AnnData object. Benchmarks were performed on the *covid-19* dataset. Each configuration was repeated 10 times to characterize the variance differences in memory pressure and hardware availability across HPC nodes. Kompot was run with single CPU, 16 CPUs and single GPU to assess the scalability.

Multi-Sample Framework

Kompot also supports multi-sample analyses by explicitly incorporating variability across replicates (**Supp. Note 3**). The density functions and gene-expression manifolds are always inferred jointly across all samples, which maximizes statistical power and prevents small or unbalanced samples from distorting the overall estimate. To capture heterogeneity between replicates, Kompot then incorporates an empirical variability term that reflect differences within each condition. This combined uncertainty accounts for both biological variability (e.g., donor or tissue differences) and technical variability (e.g., sample preparation or sequencing batch effects). By integrating these sources of variability, Kompot provides a robust and conservative basis for identifying significant changes in both cell-state densities and gene-expression patterns.

Aging hematopoiesis single-cell RNA-seq data generation

Marrow collection

Bone marrow mononuclear cells were collected from the femur, tibia, and hip bones from male C57BL/6 mice aged young (10 weeks, n=3), mid (63 weeks, n=2), and old (103 weeks, n=3). Whole marrow samples were washed twice in PBS supplemented with 10% FBS and were gently passed through a 40-micron strainer. The cell suspension was split 1:10 with one part set aside for FACS sorting of mature hematopoietic cells, and the rest for hematopoietic stem and progenitor cell (HSPC) sorting.

HSPC processing

Cells intended for HSPC sorting were first subjected to MACS lineage depletion with BioLegend antibodies targeting the following mature lineage marker antigens: Ter119 (CAT# 79748), B220 (CAT# 103204), CD3e (CAT# 79751), CD4 (CAT# 100404), CD8a (CAT# 100704), CD11b (CAT# 79749), Gr1 (CAT# 79750), FcεR1α (CAT# 134304), CD19 (CAT# 115504), NK1.1 (CAT# 108704), CD11c (CAT# 117304). Cells were incubated in lineage stain for 10 minutes, washed in MACS buffer (0.5% bovine serum albumin and 2 mM EDTA in PBS), then incubated for another 15 minutes with anti-biotin MicroBeads (5 μL for every 10e7 cells, Miltenyi Biotec CAT# 130-090-485), washed in MACS buffer, and passed through a 40 micron filter immediately prior to processing with the AutoMACS Pro Separator. After MACS lineage depletion, samples were incubated for 10 minutes with TruStain FcX PLUS (1%, BioLegend CAT# 156604), then each sample was stained for 30 minutes with a cocktail of (1) a unique hashing antibody (BioLegend CAT#s 155831, 155833, 155835, 155837, 155839, 155841, 155843, or 155845) plus (2) the flow sorting antibodies mCD45, anti-biotin and cKit plus (3) BioLegend CITE Total Seq antibodies targeting the following: CD135 (CAT# 135319), CD16/32 (CAT# 101345), Sca1 (CAT# 108149), CD127 (CAT# 135055), CD150 (CAT# 115951), CD48 (CAT# 103457), CD41 (CAT# 133941), CD34 (CAT# 128625). Samples were pooled then washed three times in Cell Staining Buffer (BioLegend CAT# 420201) then were sorted for DAPI (-), Ter119 (-), mCD45 (+), lineage (-), cKit (+).

Mature processing

Each of the 8 samples intended for Mature sorting were stained for 30 minutes with a cocktail of (1) a unique hashing antibody (BioLegend CAT#s 155831, 155833, 155835, 155837, 155839, 155841, 155843, or 155845) plus (2) the flow sorting antibodies mCD45, and Ter119 plus (3) BioLegend CITE Total Seq antibodies targeting the following: CD4 (CAT# 100573), CD8a (CAT# 100783), Ly-6C (CAT# 128053), CD11b (CAT# 101273), Ly-6G (CAT# 127659), CD44 (CAT# 10371), CD19 (CAT# 115563), CD25 (CAT# 102067), B220 (CAT# 103271), CD115 (CAT# 135543), CD11c (CAT# 117359), CD62L (CAT# 104465), F4/80 (CAT# 123155), I-A/I-E (CAT# 107657), NK1.1 (CAT# 108763), CD3 (CAT# 100257), CD335 (CAT# 137641), CD27 (CAT# 124247), CCR2 (CAT# 150633), CXCR4 (CAT# 146521), CX3CR1 (CAT# 149045). Samples were pooled then washed three times in Cell Staining Buffer (BioLegend CAT# 420201) then were sorted for DAPI (-), Ter119 (-), mCD45 (+).

Sequencing

Single cell barcoded libraries were generated using chromium Next GEM Single Cell 3' Reagent Kits, Dual Index (v3.1 Chemistry) and were sequenced on an Illumina HiSeq.

Data Analysis

Aging hematopoiesis scRNA-seq

Preprocessing

CellRanger (v6.1.2) was used for demultiplexing and count matrix generation. We next performed quality control to remove low-quality cells. Doublets were identified using Scrublet (v0.2.3)[51]. Suspected apoptotic cells were identified as any cell with 10% mitochondrial reads, and cells with 20% or more molecules derived from hemoglobin genes were identified as red blood cell contamination. Cells with abnormally low total molecule counts were also flagged. All cells failing any QC metric were removed. Genes expressed in fewer than 10 cells were excluded.

After QC, the dataset was processed in Scanpy as follows. Raw gene expression counts were normalized using the Scanpy[52] `normalize_total` function so that all cells have the same total count over all genes. A pseudocount of 0.1 was added and base-2 logarithms were taken. The top 25% most highly variable genes (HVGs) were selected using the Scanpy `highly_variable_genes` function with the

787 `cell_ranger` flavor and `batch_key`=Compartment (HSPC vs. Mature compartments). Cell cycle
788 scores were computed using the Scanpy `score_genes_cell_cycle` function, then the cell cycle effect
789 was regressed out with the Scanpy `regress_out`. PCA was carried out on the HVGs, and batch
790 correction across compartments was performed with Harmony[18]. Neighborhood graphs were computed
791 from the Harmony-corrected PCs, followed by UMAP embedding and Leiden clustering. Cell types were
792 assigned manually based on marker gene expression obtained from public data sources

793

794 *Differential Abundance, Differential Expression, and Gene Set Enrichment Analysis*

795 Diffusion maps were computed with Palantir[21] using 40 diffusion components derived from the Harmony-
796 corrected PCA representation, and this diffusion map embedding served as the latent cell-state
797 representation for all Kompot analyses. Differential abundance was computed with Kompot using default
798 parameters (ls_factor = 10, n_landmarks = 5000) for each pairwise comparison between the three age
799 groups (Young vs. Old, Young vs. Mid, and Mid vs. Old). Differential expression was performed on the
800 same comparisons with Kompot defaults (ls_factor = 10, n_landmarks = 5000, sigma = 1). For DE, we
801 used the Scanpy-normalized expression matrix described above, applying a log10 transformation to
802 normalized counts after adding a 0.1 pseudocount, and included all 16 285 genes in the analysis.

803 Gene set enrichment analysis was carried out using the Kompot `StringDBReport` function, which
804 interfaces with the STRING[53] database through its API. Gene names were mapped to STRING protein
805 identifiers using the default mapping, unmapped genes were skipped, and enrichment was computed
806 using STRING's default annotation categories for any gene set derived from the Kompot DE results.

807

808 **Mouse gastrulation data**

809 Processed and annotated data comprising of wild type and *Tal1-/-* cells (56122 cells) from E8.5 mouse
810 embryos were downloaded from ArrayExpress (E-MTAB-7325)[39]. Cell types that were exclusively present
811 in the wild type embryos (>99%) were excluded from analyses since our focus was on differential gene
812 expression between Tal1 wild type and knock out cells. Pre-computed, batch corrected PCA was used
813 input to compute diffusion maps with 50 diffusion components. Kompot differential expression analysis
814 was performed with default parameters. Top 500 differentially expressed genes ranked by Mahalanobis
815 Distance were used for all interpretive analyses.

816

817 **Melanoma CD8 T cell landscape**

818 Processed and annotated data was downloaded from GEO (GSE225063)[41]. The dataset consisted of
819 75160 cells spanning different CD8 T cell states from melanoma patients at baseline and following
820 treatment with different ICI. All cells from baseline, week4 and week16 timepoints were used for analysis
821 (74644 cells), with the union of week4 and week16 cells used as cells in the response condition. Pre-
822 computed scVI latent representation was used to compute diffusion maps with 10 components.

823 Kompot differential abundance testing was performed separately for Nivo, Rela and Combination
824 treatments to determine changes from baseline to response. Cell states with posterior tail probability <
825 1e-2 and absolute log fold change > 1 were identified to be significantly differentially abundant in each
826 treatment.

827 Kompot differential expression testing was performed separately for each treatment using genes detected
828 in at least 100 cells in the dataset (ls (15247 genes). Top 250 differentially expressed genes based on
829 Mahalanobis Distance were identified for each treatment condition and the union of these genes were
830 used interpretive analyses. STRING gene ontology analysis was performed with default parameters.

Neuroblastoma cell atlas

Processed and annotated data was downloaded from cellxgene[43]. This dataset consisted of 205153 neoplastic cells across 19 high-risk neuroblastoma patients at diagnosis and post-induction chemotherapy. Pre-computed, batch corrected PCA was used as input to compute diffusion maps with 10 components. Kompot differential abundance testing was performed using the multi-sample mode with default settings.

Data Availability

Raw single-cell data for the murine aging hematopoiesis will be deposited to GEO (GSE309924). Processed anndata objects with annotations are available to download from Zenodo[54].

Code Availability

Kompot is available as a Python module at https://github.com/settylab/Kompot. Jupyter notebooks detailing the usage of Kompot differential expression and differential abundance testing are available at https://github.com/settylab/kompot/tree/main/examples and documentation at https://kompot.readthedocs.io/en/latest/. Jupyter notebooks for reproducing the figures are available at https://github.com/settylab/kompot_notebooks, for stability analysis are available at https://github.com/settylab/kompot_de_benchmark, for memory and run time scaling at https://github.com/settylab/kompot_scaling and for differential abundance benchmarking at https://github.com/settylab/kompot_benchmarkDA.

Acknowledgements

This study was supported by National Institute of General Medical Studies grant no. R35 GM147125 to M.S.; National Cancer Institute grant no. R01CA292932, Mark Foundation for Cancer Research Endeavor Award and Edward P. Evans Foundation Discovery Research Grant to S.C.L; Brotman Baty Institute Pilot Award and Translational Data Science IRC New Collaboration Award to M.S. and S.C.L; National Institute of General Medicine Studies grant no. K99GM159044 to D.O.; National Institute of General Medicine Studies grant no. T32GM136534 to E.T.; National Institutes of Health grant no. ORIP S10OD028685 to support high-performance computing at the Fred Hutchinson Cancer Center.

Competing Interests

The authors declare no competing interests.

Figure Legends

Figure 1: Overview of Kompot's framework for differential abundance and expression analysis.

(A) Illustration of high-dimensional single-cell transcriptomic data from two conditions, projected into a shared low-dimensional cell-state space using diffusion maps.

(B) Differential abundance analysis in Kompot. Condition-specific cell-state densities with associated uncertainty estimates are computed and compared to yield density log-fold changes. Significance is quantified using posterior tail probability (PTP), visualized in a quasi-volcano plot.

(C) Differential expression analysis. Left: Schematic of inferred phenotypic manifolds for both conditions with counterfactual cell pairs illustrating expression comparison. Middle: Gene-specific differences are illustrated for an example gene, showing condition-specific expression levels, statistical interdependence, and uncertainty. Right: A quasi-volcano plot summarizes gene-level results with the average log-fold change on the x-axis and Mahalanobis distance (accounting for variance and covariance) as a multivariate significance score on the y-axis.

(D) Visualization of differential abundance (left) and differential expression (right) across the cell-state space, colored by the magnitude of log-fold change.

Figure 2: Kompot differential abundance analysis in aging hematopoiesis.

(A) UMAP projection of aging murine bone marrow scRNA-seq dataset, colored by medium-resolution cell types.

(B) UMAPs highlighting cells from mice grouped by age: Young (10 weeks), Mid (63 weeks), and Old (103 weeks).

(C) Differential abundance results for the Young to Mid (left), Young to Old (middle) and Mid to Old (right) comparisons. Quasi-cell volcano plot with density log-fold change on the x-axis and Posterior Tail Probabilities (PTP) on the y-axis. Significantly different cell-states (absolute log fold change > 1, PTP < 10^{-3}) are highlighted by their cell type annotations from (A).

(D) UMAPs displaying the estimated local density log-fold changes across the full joint cell-state space derived from all samples (Young, Mid, and Old) for comparisons in (C).

(E) UMAPs highlighting the subset of cell states with statistically significant abundance changes, as defined by the thresholds shown in the volcano plots in (C).

(F) Differential abundance volcano plots showing density log-fold changes across cells, stratified by major cell types and colored by subtype in the Young to Old comparison. y-axis indicates posterior tail probability. Violin plots in the background indicate abundance of cells at the respective density log-fold change in this group.

(G) Direction and magnitude of abundance changes per cell type in the Young to Old comparison. Bars show the fraction of significantly changing cell-states within each annotated cell type, color-coded by the direction of change. Cell types are ordered along a progenitor-to-mature axis.

(H) Kompot results for early hematopoietic progenitors. Left: UMAP highlighting HSC (Hematopoietic Stem Cells), and LMPP (Multipotent Progenitors) subsets. Right: Scatterplot of cells colored by density log-fold change along pseudotime (x-axis), showing dynamic abundance shifts during early differentiation.

Figure 3: Kompot differential expression analysis in aging hematopoiesis.

(A) UMAP from Fig. 2A, with HSCs (Hematopoietic Stem Cells) highlighted.

(B) Differential expression results for Mid to Old HSCs. Quasi-cell volcano plot with expression log-fold change on the x-axis and Mahalanobis distance on the y-axis. Significantly different genes (absolute log fold change > 0.15, Mahalanobis distance > 3) are highlighted.

(C) Same as (B), colored by weighted average of reported "Aging Consistency" using standard Gaussian Kernel after scaling x- and y-axis to standard deviation 1.

(D) Same as Fig. 2H, colored by Mid to Old HSC gene expression log fold change for a gene downregulated with age (Cd52, left) and an upregulated gene (Ifitm1, right).

(E) UMAP from Fig. 2A, with monocytes highlighted.

(F) Same as (B), for Young to Old monocytes.

(G) Gene ontology analysis from STRING[53] for downregulated genes (left) and upregulated genes (right) from Fig. 3F.

(H) UMAP from Fig. 2A, highlighting all cell types except T and NK cells

(I) Differential expression results for all cells in (H) for Young to Old comparisons. Top 10 genes with positive fold change and high Mahalanobis distance are shown in red, bottom 10 genes with negative fold change and high Mahalanobis distance are shown in blue, and genes with log absolute fold change but high Mahalanobis distance are shown in green (absolute log fold change < 0.07, Mahalanobis distance > 6). 20 genes are labeled.

(J) Heatmap showing Young to Old gene expression log fold changes across different cell types for genes highlighted in (I).

(K) Top: Imputed expression of gene H2-Q7 using young cells (left), using old cells (right) and fold change from young to old (middle). Middle, Bottom: Same as the top row for Mef2c and Dynll1. These genes serve as examples of genes with high positive fold change, high Mahalanobis but low absolute fold change, and high negative fold change respectively.

(L) Gene ontology analysis using genes with high Mahalanobis distance and low absolute fold change from (I).

(M) Heatmap showing Young to Old gene expression log fold changes across different cell types for genes involved in antioxidant activity.

Figure 4: Kompot reveals determinants of immunotherapy in advanced melanoma

(A) UMAP highlighting the CD8 T cell state (left) landscape from immunotherapy recipients with advanced melanoma[41]. Right: Same UMAP as (left) with cells colored by treatment type.

(B) Same as (A) with cells colored by treatment condition for Nivolumab (anti-PD1, left), Relatlimab (anti-LAG-3, center) and combination of Nivolumab and Relatlimab (right).

(C) Kompot differential abundance results highlighting significantly altered cell states (p-value < 1e-3, absolute log-fold change > 1) for the three treatment conditions.

(D) Boxplot comparing the abundance fold changes for different CD8 T cell states across treatment conditions. Activated Cycling cells which have greater abundance in the combination treatment are highlighted.

(E) Kompot differential expression results for the three treatment conditions. The union of top 250 genes ranked by Mahalanobis Distance, across the three treatment conditions are highlighted in the quasi-volcano plots.

(F) Heatmap showing the pairwise correlation between average fold change from baseline to response in different cell states and treatment conditions. Genes highlighted in (E) were used for computing correlations. Correlations of Activated cycling cells in the different treatment conditions are highlighted.

(G) Heatmap showing the log fold change from baseline to response conditions for T cell activation genes. These are a subset of genes in (E) selected based on gene ontology analysis. Columns of activated cycling cells in the three treatment conditions are highlight. Columns are in the same order as (F).

(H) Same UMAP as (A), with activated cycling T cells highlighted.

(I) Cells in (H) colored by gene expression fold change from baseline to response in Nivo (left) or the combination treatment (right). Same genes show opposing expression changes with different treatments.

Figure 5: Inter-patient variability with multi-sample differential abundance testing

(A) Illustration of multi-sample differential abundance testing in Kompot. Variability is computed using density estimates for each sample within the condition for significance assessment.

(B) UMAP of neoplastic cells from high-risk neuroblastoma patients colored by cell-states[43].

(C-D) Same as (B), colored by treatment condition and patient id

(E) UMAP of the same cells in (B), prior to batch correction. Colored by patient id (left), cell type (middle) and patient event category (right).

(F) Kompot differential abundance volcano plot with significantly altered cell-states from diagnosis to post-induction chemotherapy (p-value < 5e-2, absolute log-fold change > 1) colored by cell types from (B).

(G) Direction and magnitude of abundance changes per cell type. Bars show the fraction of significantly changing cell-states within each annotated cell type, color-coded by the direction of change.

(H) Same as (G), but separately for each patient. Patients are ordered based on whether event free or suffered an adverse event. Within the two categories, patients are sorted by fraction of differentially abundant ADRN-proliferating and ADRN-Baseline cells. A greater reduction associates with event free survival.

(Relatlimab, Rela) and anti-PD-1 (Nivolumab, Nivo)

Figure 6: Kompot runtime scaling analysis.

(A) Runtime scaling of Kompot differential abundance analysis as a function of number of cells. Benchmarks were performed on four real datasets: "aging" (Young vs. Old), "bcr-xl" (Ref vs. BCR-XL), "covid" (Healthy vs. COVID), and "covid-post" (Healthy vs. Post). Each dataset was subset to varying

number of cells and run with different number of CPUs, using the default 5000 landmarks. Dashed line indicates linear scaling.

(B) Same as (A) for differential expression analysis.

(C) Runtime scaling of differential abundance analysis as a function of the number of landmark cell states. Number of cells for each dataset are indicated in the legend.

(D) Same as (C) for differential expression analysis.

(E) Runtime scaling of differential expression analysis as a function of the number of genes tested, using the default 5000 landmarks. Number of cells for each dataset are indicated in the legend.

994 **Supplementary Figure Legends**
995
996

997 Supplementary Figure 1: Kompot differential expression

998 (A) Point-wise fold changes used in Kompot on the left, and diminished fold change with large confidence
999 intervals of aggregated pseudobulk in comparison on the right.

1000 (B) Batch correction aligns comparable cell-states in cell-state space.

1001

1002

1003 Supplementary Figure 2: Kompot differential abundance

1004 (A) Same as 1A: Diffusion map representation of co-embedded cell-states from two conditions.

1005 (B) Left: Subset of cells from each condition. Center: Continuous density function for each condition.
1006 Right: Condition-specific density evaluated for all cell states across conditions using the two separate
1007 density functions.

1008 (C) Abundance log fold change computed as the ratio of the two condition specific densities at single-cell
1009 resolution.

1010

1011

1012 Supplementary Figure 3: CITE-seq dataset of murine aging hematopoiesis.

1013 (A) Schematic illustrating the generation of our murine aging hematopoiesis CITE-seq dataset

1014 (B) Plots comparing molecule counts (x-axis) and fraction of mitochondrial molecules (y-axis) for each
1015 sample. Cells with greater than 10% mitochondrial molecules were excluded from downstream analysis.

1016 (C) Plots comparing molecule counts and gene counts for each sample. Cells are colored using
1017 annotations in Fig. 2A

1018

1019

1020 Supplementary Figure 4: CITE-seq dataset of murine aging hematopoiesis.

1021 (A) UMAP of murine aging hematopoiesis dataset colored by medium resolution cell types.

1022 (B) Same as (A), colored by high resolution cell types.

1023

1024

1025 Supplementary Figure 5: Age-specific densities and uncertainties

1026 (A) UMAPs highlighting cells from different age (Same as Fig. 2B).

1027 (B) Age-specific density computed using the subset of cells in (A).

1028 (C) Age-specific density evaluated all cells using the continuous density function.

1029 (D) Uncertainties associated with density estimates in (C).

1030

Supplementary Figure 6: Kompot differential abundance analysis in aging hematopoiesis.

Same as Fig. 2F, for Young to Mid and Mid to Old comparisons.

Supplementary Figure 7: Consistency of Kompot differential expression results with hematopoietic aging atlas

(A) Enrichment of aging-associated genes among Kompot's top differentially expressed genes. Bars show the fraction of genes within the top k ranked by Mahalanobis significance that are part of a published hematopoietic aging signature, defined as genes with a reported consistency score ≥ 4. The dashed horizontal line indicates the background frequency of such genes across all genes tested.

(B) Comparison of average log-fold changes (logFC) per gene between Kompot (Mid to Old transition; y-axis) and the published aging signature (x-axis). Genes with consistency scores ≥ 4 are highlighted in green; those with scores ≥ 8 are additionally labeled. Pearson and Spearman correlation coefficients between the two datasets are shown.

Supplementary Figure 8: Kompot differential expression analysis in aging hematopoiesis.

Same as Fig. 3F, for neutrophils, basophils, pDCs and cDCs.

Supplementary Figure 9: Kompot differential expression analysis in *Tal1-/-* mouse embryos

(A) UMAP colored by Tal1 status (left) and cell types (right) for cells from E8.5 mouse embryos[39].

(B) Subset of cells from (A) with at least 1% of cells from the cell type from Tal *1-/-* embryos.

(C) Kompot differential expression results from wild type to knockout highlight the top 20 genes.

(D) Same as (C), highlighting top 500 genes that change from wild type to knockout based on Mahalanobis distance

(E) Gene ontology analysis results from genes in (D)

(F) Heatmap showing the gene expression fold changes in different cell types for the genes involved in vasculature development. Genes shown are a subset of highlighted genes in (D) selected based on gene ontology analysis.

(G) Top: UMAP from (B) showing the hematoendothelial progenitors and endothelial cells colored by cell type (left) and Tal1 status (right). Bottom: UMAPs colored by gene expression change from wild type to mutant for three representative vasculature genes highlighting the heterogenous nature of gene expression changes in these cell types in response to Tal1 knockout.

(H) Gene ontology analysis for the genes in (D) using the KEGG pathway annotations.

(I-J). Similar to heatmaps in (H), highlighting genes involved in fluid shear stress and cell cycle.

Supplementary Fig 10: Kompot differential abundance results comparing baseline to response in different immunotherapy treatments in advanced melanoma

(A) Same as the UMAP in Fig. 4A, colored by Kompot differential abundance fold change from baseline to response in different immunotherapy treatments. This accompanies Fig. 4C.

(B) Same as Fig. 4C

(C) Direction and magnitude of abundance changes per cell type with different treatments. Bars show the fraction of significantly changing cell-states within each annotated cell type, color-coded by the direction of change.

(D) Differential abundance volcano plots showing density log-fold changes across cells, stratified and colored by cell state. y-axis indicates posterior tail probability. Violin plots in the background indicate abundance of cells at the respective density log-fold change in this group.

Supplementary Fig 11: Kompot differential expression results comparing baseline to response in different immunotherapy treatments in advanced melanoma

(A) Differential expression pseudo-volcano plots highlighting the top 250 genes that change from baseline to response in each treatment condition. Genes are ranked using Mahalanobis Distance. This accompanies Fig. 4E.

(B) Gene ontology results from genes highlighted in Fig. 4E.

(C) STRING connectivity graph for genes in Fig. 4E.

(D) Same as Fig. 4H-I, with gene expression changes across all cells.

(E) Same as Fig. 4G, with heatmap showing expression changes of ribosomal genes.

Supplementary Fig. 12: Multi-sample differential abundance testing in high-risk neuroblastoma

(A) Same as Fig. 4B, with cells of each patient highlighted and colored by cell state.

(B) Same as (A), with cells of each patient colored by differential abundance fold change from diagnosis to post-induction.

Supplementary Fig. 13: Robustness of Kompot differential abundance mean log fold change for the hematopoietic aging dataset

Benchmarking was performed by computing the Pearson correlation of differential abundance log fold change across all cell-states between pairs of parameter choices.

Top: Heatmaps shows Pearson correlations pairs of parameter values, computed for the Young vs. Old comparison in the aging dataset. Each column shows a different parameter being varied: (A) number of cells (dataset subset to varying number cells), (B) length-scale factor (default: 10), (C) number of landmarks (default: 5000), and (D) number of diffusion map components (default: 50). Default values are indicated by black lines on the matrix

Middle: Margin plots mapping each column of the correlation matrix to its corresponding parameter value for the respective datasets(note log scaling for some parameters).

Bottom: Correlation of abundance LFC to a reference setting as a function of the varied parameter. The reference is the maximum cell count in (A) and the default parameter value in (B-D), indicated by a red dashed line. This corresponds to extracting a single row from the heatmap in (A).

Supplementary Fig. 14: Robustness of Kompot differential abundance mean log fold change for the COVID-19 dataset

(A-D) Same as Supplementary Fig. 13 for the COVID-19 dataset.

Supplementary Fig. 15: Robustness of Kompot differential abundance posterior tail probabilities for the hematopoietic aging dataset

(A-D) Same as Supplementary Fig. 13 showing robustness of posterior tail probabilities instead of log fold change.

Supplementary Fig. 16: Robustness of Kompot differential abundance posterior tail probabilities for the COVID-19 dataset

(A-D) Same as Supplementary Fig. 15 for the COVID-19 dataset.

Supplementary Fig. 17: Robustness of Kompot differential expression mean log fold change for the hematopoietic aging dataset

(A-D) Same as Supplementary Fig. 13 showing robustness of differential expression mean log fold change across all genes instead of abundance log-fold change across all cells.

Supplementary Fig. 18: Robustness of Kompot differential expression mean log fold change for the COVID-19 dataset

(A-E) Same as Supplementary Fig. 17 for the COVID-19 dataset.

Supplementary Fig. 19: Robustness of Kompot differential expression Mahalanobis distance for the hematopoietic aging dataset

(A-E) Same as Supplementary Fig. 17 showing robustness of Mahalanobis distance instead of mean log fold change.

Supplementary Fig. 20: Robustness of Kompot differential expression Mahalanobis distance for the COVID-19 dataset

(A-E) Same as Supplementary Fig. 19 for the COVID-19 dataset.

Supplementary Fig. 21: Benchmarking of differential abundance analysis.

Benchmarking was performed the framework defined in benchmarkDA[12]. Three simulated datasets (linear, branch and cluster) and one real-world datasets with artificially generated labels (covid-19) were used. Kompot performed was compared against Meld[10] and Milo[11]. AUROC and AUPRC were used for performance comparison.

(A-B) AUROC (A) and AUPRC (B) for Kompot, Meld and Milo for the linear dataset with different batch effect levels. Statistics were computed for different resolutions and cell types.

(C-D) AUROC (C) and AUPRC (D) for Kompot, Meld and Milo for the linear dataset with resolutions. Statistics were computed for different batch effects and cell types.

(E-H) Same as (A-D) for the branch dataset.

(I-L) Same as (A-D) for the cluster dataset.

(M-P) Same as (A-D) for the covid-19 dataset.

Supplementary Fig. 22: Kompot peak memory usage scaling analysis.

All benchmarks were performed on the "covid" dataset (Healthy vs. COVID comparison). Lines show median run times for 10 experiments and shaded regions indicate the 5th-95th percentile intervals. SV: Sample variance estimation (the most computationally expensive option). Gene counts are abbreviated (e.g., "200g" = 200 genes), 16CPU: indicates parallel execution with 16 CPUs. GPU: GPU acceleration (NVIDIA L40S). The left column shows peak memory usage times for differential abundance analysis and the right column shows differential expression analysis.

(A-C) Peak memory usage for differential abundance analysis as a function of (A) number of cells using 5000 landmarks and 50 diffusion components, (B) number of landmarks using 422,220 cells and 50 diffusion components, (C) number of diffusion components using 422,220 cells and 5000 landmarks.

(D-G) Peak memory usage for differential expression analysis as a function of (D) number of cells using 5000 landmarks and 50 diffusion components, (E) number of landmarks using 422,220 cells and 50 diffusion components, (F) number of diffusion components using 422,220 cells and 5000 landmarks, (G) number of genes tested using 422,220 cells and 5000 landmarks.

Supplementary Fig. 23: Kompot runtime scaling analysis.

All benchmarks were performed on the "covid" dataset (Healthy vs. COVID comparison). Lines show median run times for 10 experiments and shaded regions indicate the 5th-95th percentile intervals. SV: Sample variance estimation (the most computationally expensive option). Gene counts are abbreviated (e.g., "200g" = 200 genes), 16CPU: indicates parallel execution with 16 CPUs. GPU: GPU acceleration (NVIDIA L40S). The left column shows run times for differential abundance analysis and the right column shows differential expression analysis.

(A-C) Runtime analysis for differential abundance analysis as a function of (A) number of cells using 5000 landmarks and 50 diffusion components, (B) number of landmarks using 422,220 cells and 50 diffusion components, (C) number of diffusion components using 422,220 cells and 5000 landmarks.

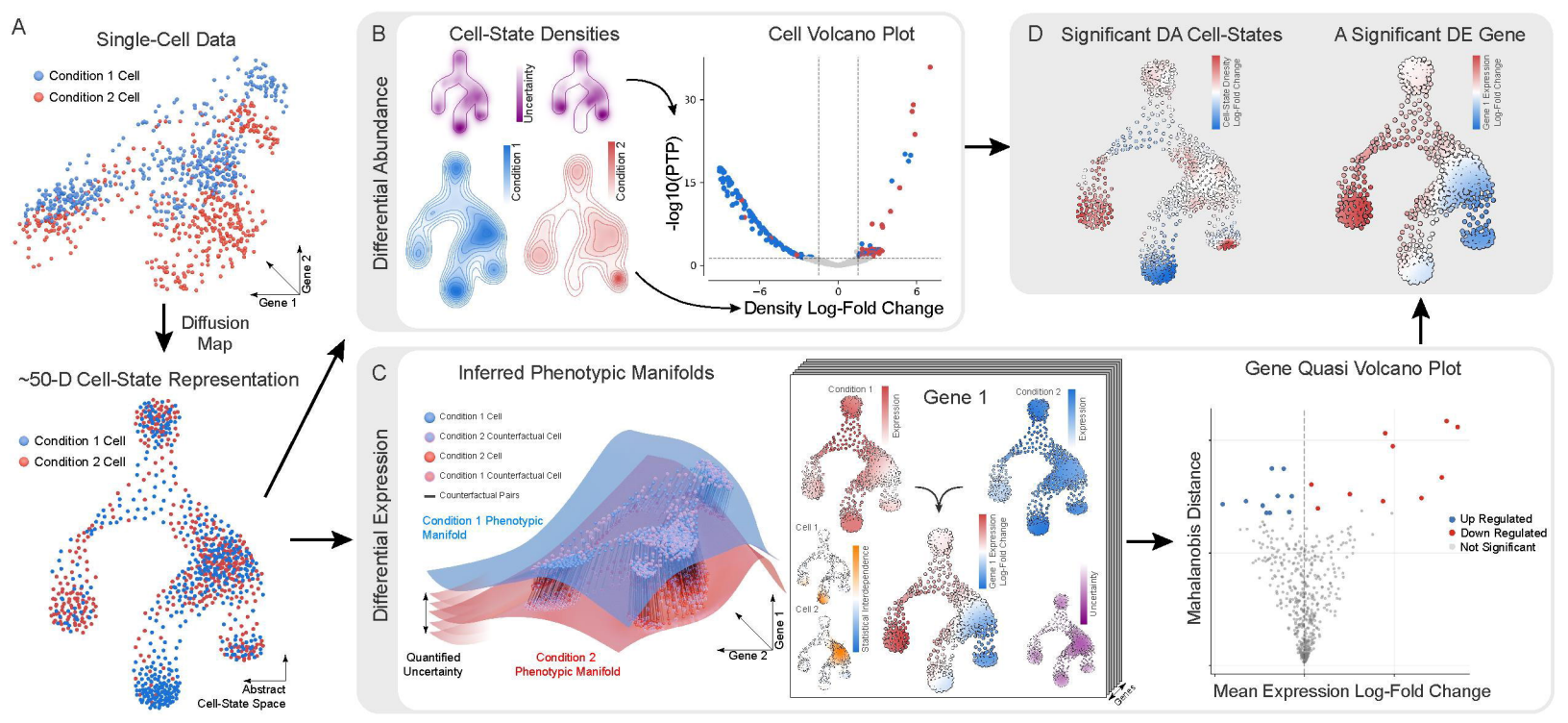
(D-G) Runtime analysis for differential expression analysis as a function of (D) number of cells using 5000 landmarks and 50 diffusion components, (E) number of landmarks using 422,220 cells and 50 diffusion

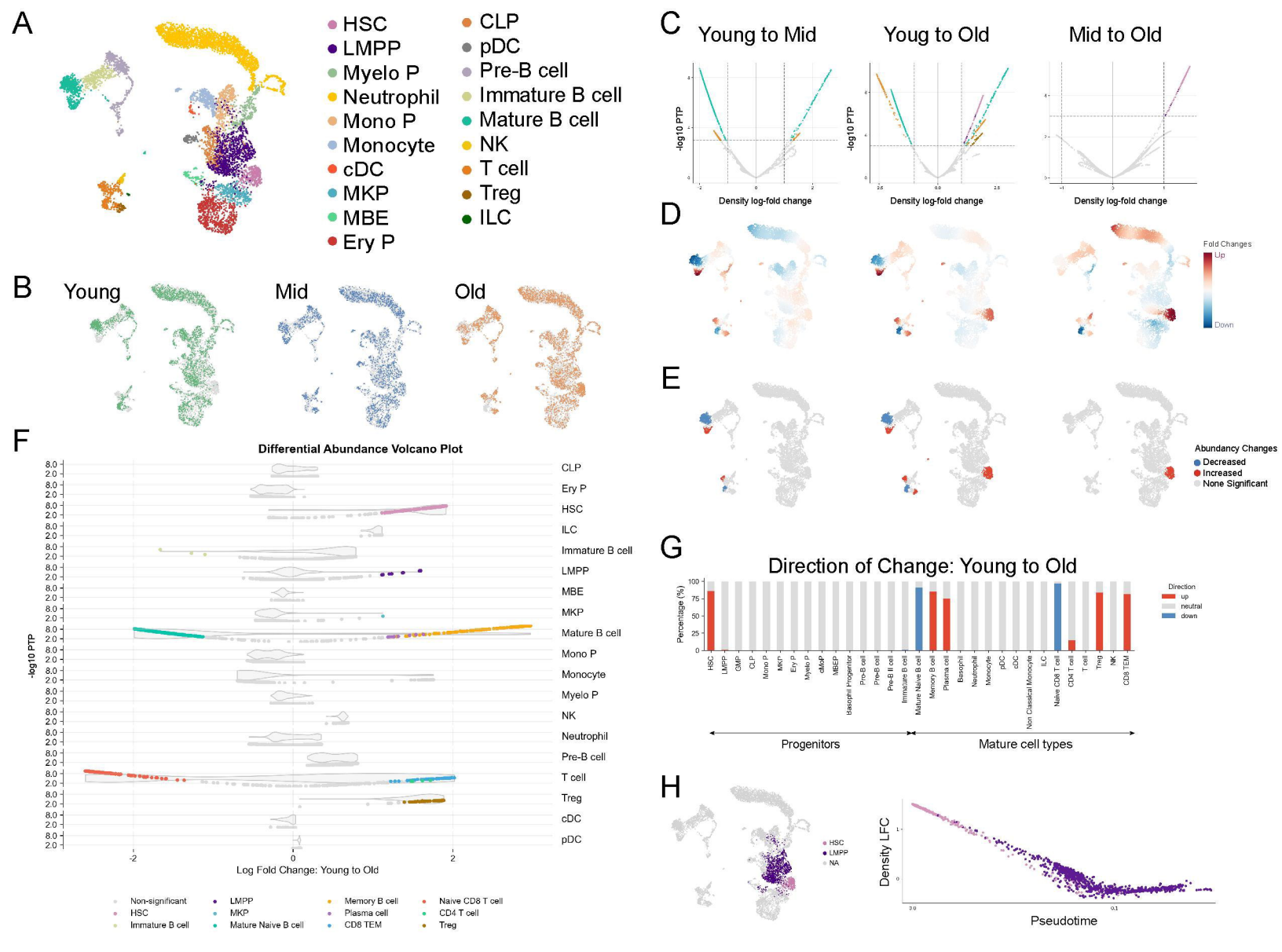
components, (F) number of diffusion components using 422,220 cells and 5000 landmarks, (G) number of genes tested using 422,220 cells and 5000 landmarks.

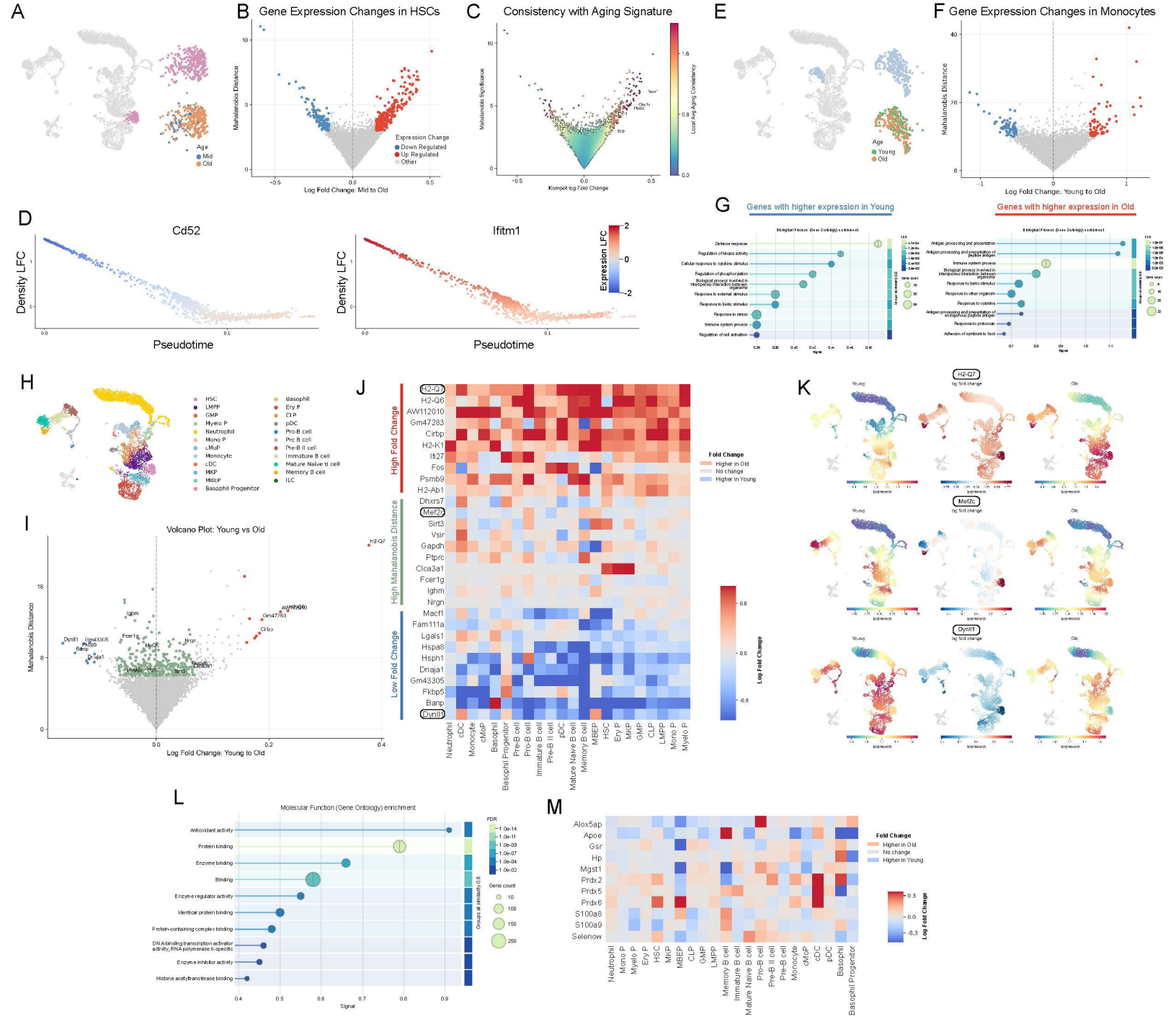
References

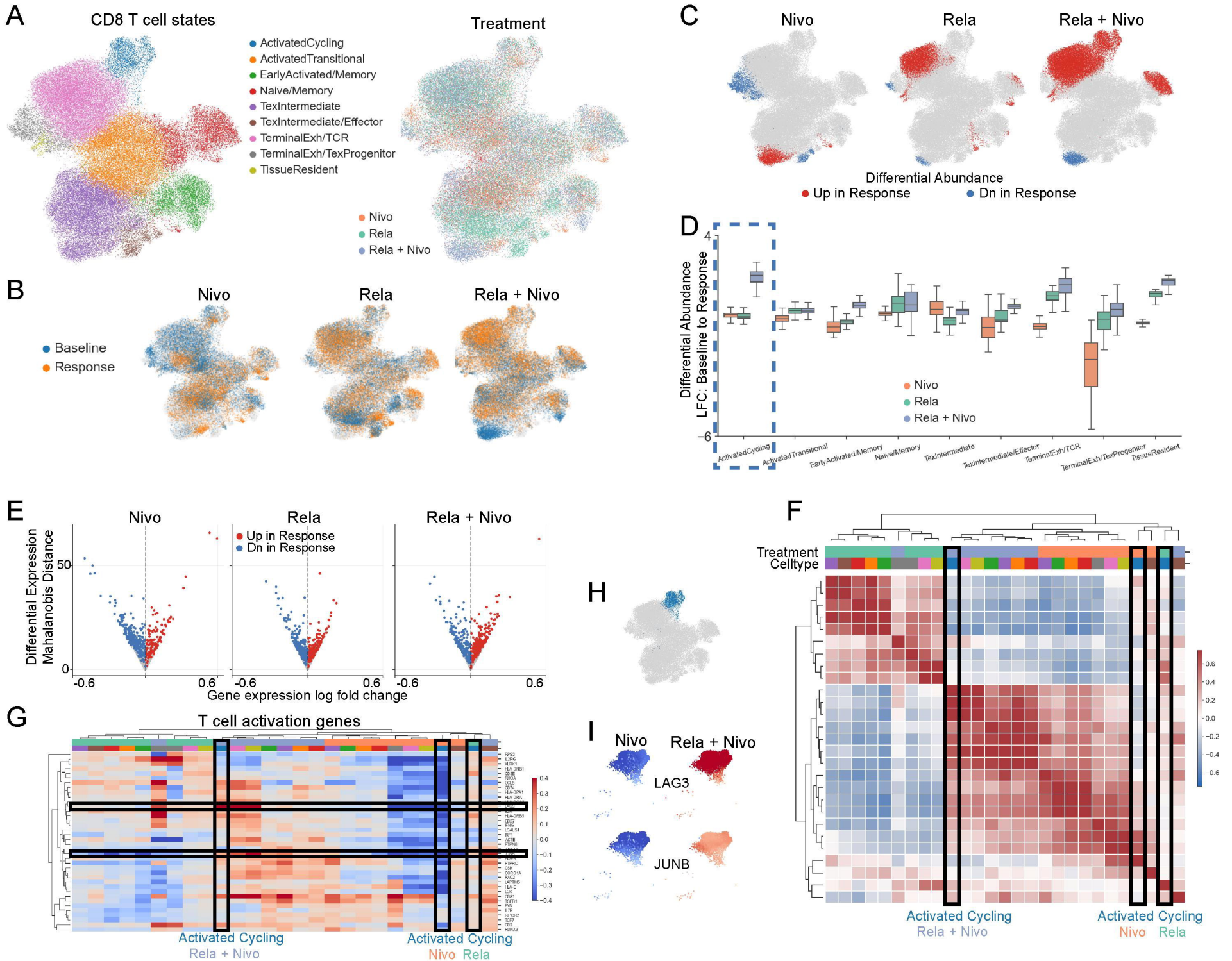
1. Tabula Muris C. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature. 2020;583(7817):590-5, PMC8240505.
2. Konturek-Ciesla A, Dhapola P, Zhang Q, Sawen P, Wan H, Karlsson G, et al. Temporal multimodal single-cell profiling of native hematopoiesis illuminates altered differentiation trajectories with age. Cell Rep. 2023;42(4):112304.
3. Herault L, Poplineau M, Mazuel A, Platet N, Remy E, Duprez E. Single-cell RNA-seq reveals a concomitant delay in differentiation and cell cycle of aged hematopoietic stem cells. BMC Biol. 2021;19(1):19, PMC7851934.
4. Kucinski I, Campos J, Barile M, Severi F, Bohin N, Moreira PN, et al. A time- and single-cell-resolved model of murine bone marrow hematopoiesis. Cell Stem Cell. 2023.
5. Gazestani V, Kamath T, Nadaf NM, Dougalis A, Burris SJ, Rooney B, et al. Early Alzheimer's disease pathology in human cortex involves transient cell states. Cell. 2023;186(20):4438-53 e23, PMC11107481.
6. Mayr CH, Simon LM, Leuschner G, Ansari M, Schniering J, Geyer PE, et al. Integrative analysis of cell state changes in lung fibrosis with peripheral protein biomarkers. EMBO Mol Med. 2021;13(4):e12871, PMC8033531.
7. Rehman SK, Haynes J, Collignon E, Brown KR, Wang Y, Nixon AML, et al. Colorectal Cancer Cells Enter a Diapause-like DTP State to Survive Chemotherapy. Cell. 2021;184(1):226-42 e21, PMC8437243.
8. van der Leun AM, Thommen DS, Schumacher TN. CD8(+) T cell states in human cancer: insights from single-cell analysis. Nat Rev Cancer. 2020;20(4):218-32, PMC7115982.
9. Replogle JM, Saunders RA, Pogson AN, Hussmann JA, Lenail A, Guna A, et al. Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq. Cell. 2022;185(14):2559-75 e28, PMC9380471.
10. Burkhardt DB, Stanley JS, 3rd, Tong A, Perdigoto AL, Gigante SA, Herold KC, et al. Quantifying the effect of experimental perturbations at single-cell resolution. Nat Biotechnol. 2021;39(5):619-29, PMC8122059.
11. Dann E, Henderson NC, Teichmann SA, Morgan MD, Marioni JC. Differential abundance testing on single-cell data using k-nearest neighbor graphs. Nat Biotechnol. 2022;40(2):245-53.
12. Yi H, Plotkin A, Stanley N. Benchmarking differential abundance methods for finding condition-specific prototypical cells in multi-sample single-cell datasets. Genome Biology. 2024;25(1):9.
13. Missarova A, Dann E, Rosen L, Satija R, Marioni J. Leveraging neighborhood representations of single-cell data to achieve sensitive DE testing with miloDE. Genome Biol. 2024;25(1):189, PMC11256449.
14. Boyeau P, Hong J, Gayoso A, Kim M, McFaline-Figueroa JL, Jordan MI, et al. Deep generative modeling of sample-level heterogeneity in single-cell genomics. bioRxiv. 2024:2022.10.04.510898.
15. Ahlmann-Eltze C, Huber W. Analysis of multi-condition single-cell data with latent embedding multivariate regression. Nat Genet. 2025;57(3):659-67, PMC11906359.
16. Song D, Li K, Ge X, Li JJ. ClusterDE: a post-clustering differential expression (DE) method robust to false-positive inflation caused by double dipping. Res Sq. 2023, PMC10418557.
17. Gao LL, Bien J, Witten D. Selective Inference for Hierarchical Clustering. J Am Stat Assoc. 2024;119(545):332-42, PMC11036349.
18. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16(12):1289-96, PMC6884693.
19. Luecken MD, Buttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. Nat Methods. 2022;19(1):41-50, PMC8748196.
20. Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, et al. Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. Proc Natl Acad Sci U S A. 2005;102(21):7426-31, PMC1140422.
21. Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe'er D. Characterization of cell fate probabilities in single-cell data with Palantir. Nat Biotechnol. 2019;37(4):451-60.
22. Haghverdi L, Buttner M, Wolf FA, Buettner F, Theis FJ. Diffusion pseudotime robustly reconstructs lineage branching. Nat Methods. 2016;13(10):845-8.
23. Otto D, Jordan C, Dury B, Dien C, Setty M. Quantifying Cell-State Densities in Single-Cell Phenotypic Landscapes using Mellon. Nature Methods. 2024, PMC10369887.
24. Rasmussen CE, Williams CKI. Gaussian Processes for Machine Learning2006.
25. PC M. Reprint of: Mahalanobis, P.C. (1936) "On the Generalised Distance in Statistics.". Sankhya A. 2018;80(1):1-7.
26. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science. 2019;366(6465), PMC8050831.
27. Chambers SM, Shaw CA, Gatza C, Fisk CJ, Donehower LA, Goodell MA. Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. PLoS Biol. 2007;5(8):e201, PMC1925137.
28. Li H, Cote P, Kuoch M, Ezike J, Frenis K, Afanassiev A, et al. The dynamics of hematopoiesis over the human lifespan. Nat Methods. 2025;22(2):422-34, PMC11908799.

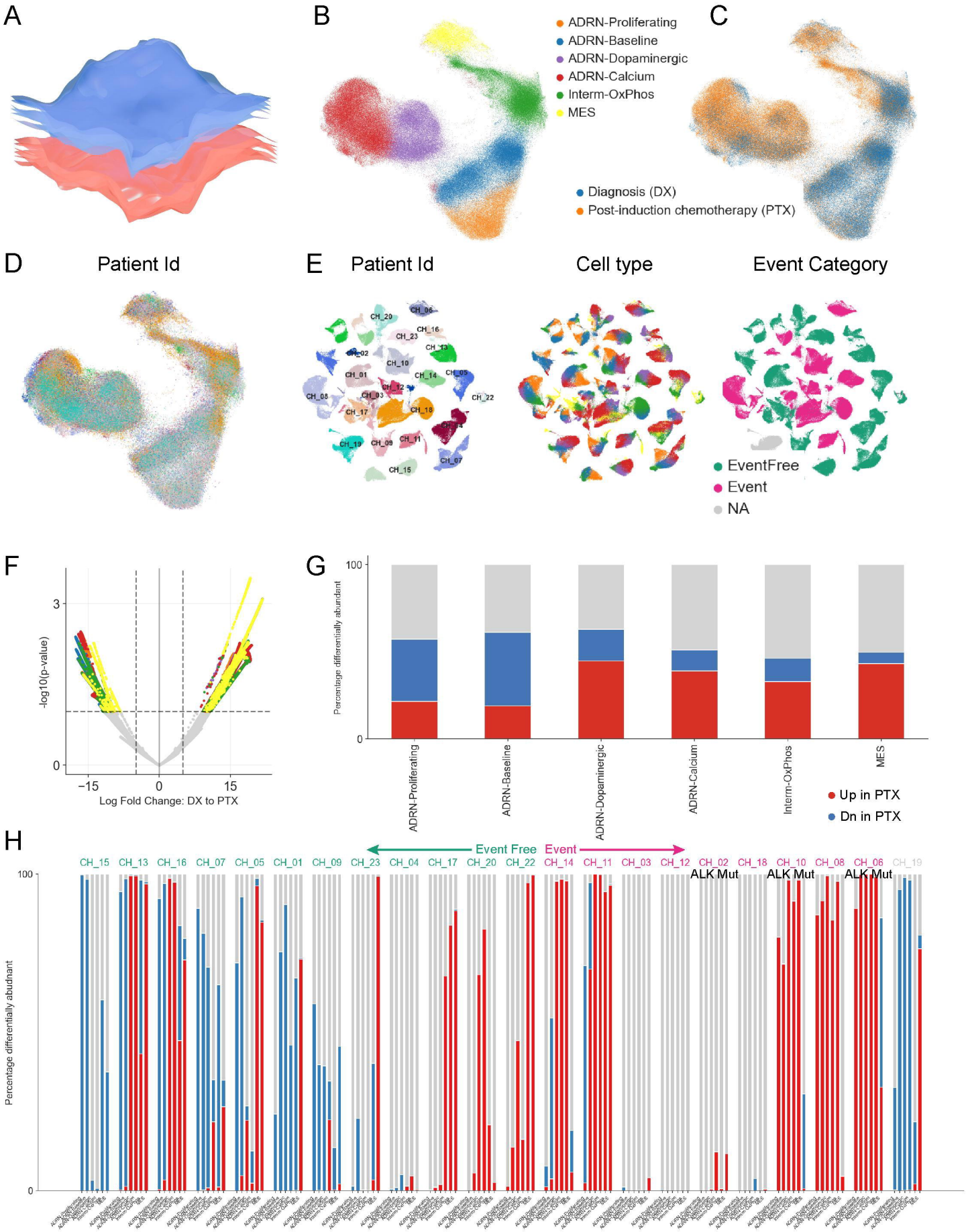
29. Appelbaum FR, Gundacker H, Head DR, Slovak ML, Willman CL, Godwin JE, et al. Age and acute myeloid leukemia. Blood. 2006;107(9):3481-5, PMC1895766.
30. Florian MC, Klose M, Sacma M, Jablanovic J, Knudson L, Nattamai KJ, et al. Aging alters the epigenetic asymmetry of HSC division. PLoS Biol. 2018;16(9):e2003389, PMC6168157.
31. Bulati M, Caruso C, Colonna-Romano G. From lymphopoiesis to plasma cells differentiation, the age-related modifications of B cell compartment are influenced by "inflamm-ageing". Ageing Res Rev. 2017;36:125-36.
32. Goronzy JJ, Weyand CM. Immune aging and autoimmunity. Cell Mol Life Sci. 2012;69(10):1615-23, PMC4277694.
33. de Haan G, Lazare SS. Aging of hematopoietic stem cells. Blood. 2018;131(5):479-87.
34. Flohr Svendsen A, Yang D, Kim K, Lazare S, Skinder N, Zwart E, et al. A comprehensive transcriptome signature of murine hematopoietic stem cell aging. Blood. 2021;138(6):439-51.
35. Yang D, Skinder N, Kao YR, Chen J, Thiruthuvanathan V, Flohr Svendsen A, et al. Aberrant engagement of P-selectin drives hematopoietic stem cell aging in mice. Nat Aging. 2025;5(6):1010-24.
36. De Maeyer RPH, Chambers ES. The impact of ageing on monocytes and macrophages. Immunol Lett. 2021;230:1-10.
37. Barman PK, Shin JE, Lewis SA, Kang S, Wu D, Wang Y, et al. Production of MHCII-expressing classical monocytes increases during aging in mice and humans. Aging Cell. 2022;21(10):e13701, PMC9577948.
38. Shao L, Li H, Pazhanisamy SK, Meng A, Wang Y, Zhou D. Reactive oxygen species and hematopoietic stem cell senescence. Int J Hematol. 2011;94(1):24-32, PMC3390185.
39. Pijuan-Sala B, Griffiths JA, Guibentif C, Hiscock TW, Jawaid W, Calero-Nieto FJ, et al. A single-cell molecular map of mouse gastrulation and early organogenesis. Nature. 2019;566(7745):490-5, PMC6522369.
40. Daniel B, Yost KE, Hsiung S, Sandor K, Xia Y, Qi Y, et al. Divergent clonal differentiation trajectories of T cell exhaustion. Nat Immunol. 2022;23(11):1614-27, PMC11225711.
41. Cillo AR, Cardello C, Shan F, Karapetyan L, Kunning S, Sander C, et al. Blockade of LAG-3 and PD-1 leads to co-expression of cytotoxic and exhaustion gene modules in CD8(+) T cells to promote antitumor immunity. Cell. 2024;187(16):4373-88 e15, PMC11346583.
42. Wang W, Green M, Choi JE, Gijon M, Kennedy PD, Johnson JK, et al. CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. Nature. 2019;569(7755):270-4, PMC6533917.
43. Yu W, Biyik-Sit R, Uzun Y, Chen CH, Thadi A, Sussman JH, et al. Longitudinal single-cell multiomic atlas of high-risk neuroblastoma reveals chemotherapy-induced tumor microenvironment rewiring. Nat Genet. 2025;57(5):1142-54, PMC12081299.
44. Stephenson E, Reynolds G, Botting RA, Calero-Nieto FJ, Morgan MD, Tuong ZK, et al. Single-cell multi-omics analysis of the immune response in COVID-19. Nat Med. 2021;27(5):904-16, PMC8121667.
45. Yi H, Plotkin A, Stanley N. Benchmarking differential abundance methods for finding condition-specific prototypical cells in multi-sample single-cell datasets. bioRxiv. 2023.
46. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. Nat Methods. 2018;15(12):1053-8, PMC6289068.
47. Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HY, et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat Genet. 2021;53(3):403-11, PMC8012210.
48. Setty M, Tadmor MD, Reich-Zeliger S, Angel O, Salame TM, Kathail P, et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat Biotechnol. 2016;34(6):637-45, PMC4900897.
49. Efron B, Tibshirani R, Storey JD, Tusher V. Empirical Bayes Analysis of a Microarray Experiment. J Am Stat Assoc. 2001;96(456):1151-60.
50. Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, Bruggner RV, et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. Nat Biotechnol. 2012;30(9):858-67, PMC3627543.
51. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. 2019;8(4):281-91 e9, PMC6625319.
52. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15, PMC5802054.
53. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D13, PMC6323986.
54. Otto D, Arriaga-Gomez E, Thieme E, Yang R, Lee S, Setty M. Single-cell CITE-seq of murine bone marrow across aging (Young, Mid, Old): Zenodo; 2025 [Available from: https://doi.org/10.5281/zenodo.15587767.

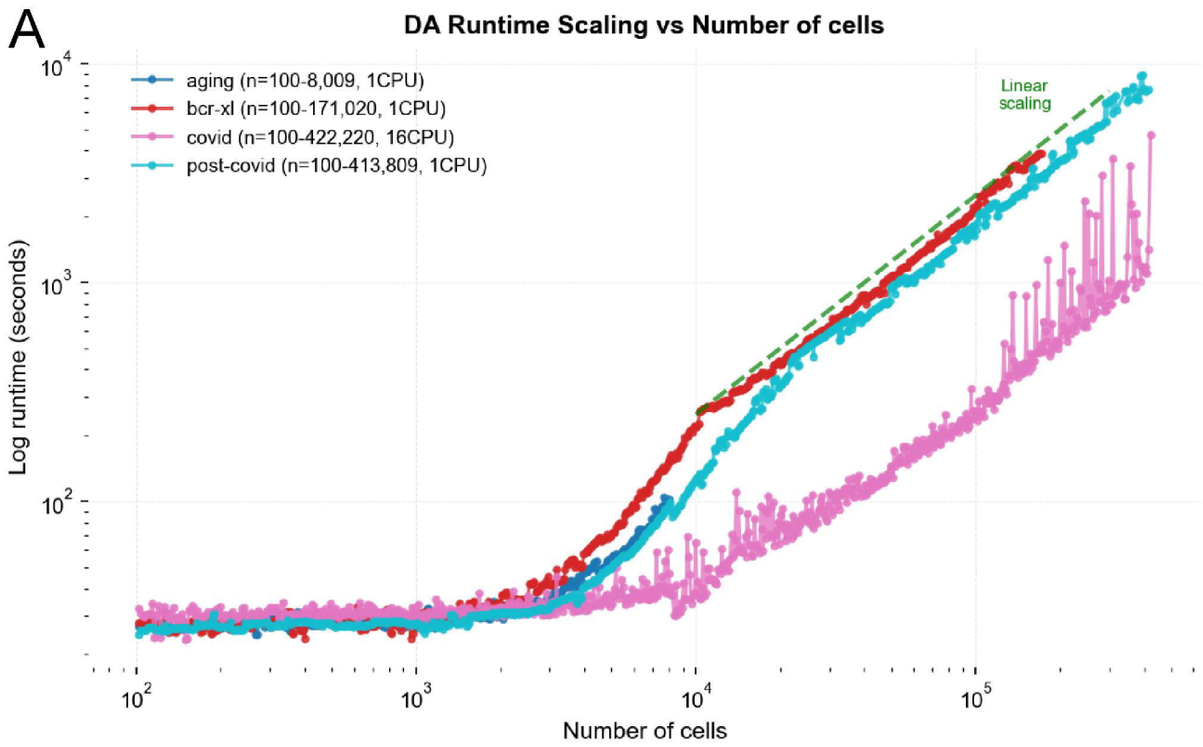

A
Single-Cell Data
Condition 1 Cell
Condition 2 Cell
Gene 1
Gene 2
Diffusion Map
~50-D Cell-State Representation
Condition 1 Cell
Condition 2 Cell
Abstract Cell-State Space
B
Differential Abundance
Cell-State Densities
Uncertainty
Condition 1
Condition 2
Cell Volcano Plot
-log10(PTP)
30
15
0
−6
0
6
Density Log-Fold Change
C
Differential Expression
Inferred Phenotypic Manifolds
Condition 1 Cell
Condition 2 Counterfactual Cell
Condition 2 Cell
Condition 1 Counterfactual Cell
Counterfactual Pairs
Condition 1 Phenotypic Manifold
Quantified Uncertainty
Condition 2 Phenotypic Manifold
Gene 2
Gene 1
Gene 1
Condition 1
Expression
Condition 2
Expression
Cell 1
Cell 2
Statistical Interdependence
Gene 1 Expression Log-Fold Change
Uncertainty
Genes
Gene Quasi Volcano Plot
Mahalanobis Distance
Up Regulated
Down Regulated
Not Significant
Mean Expression Log-Fold Change
D
Significant DA Cell-States
Cell-State Density Log-Fold Change
A Significant DE Gene
Gene 1 Expression Log-Fold Change

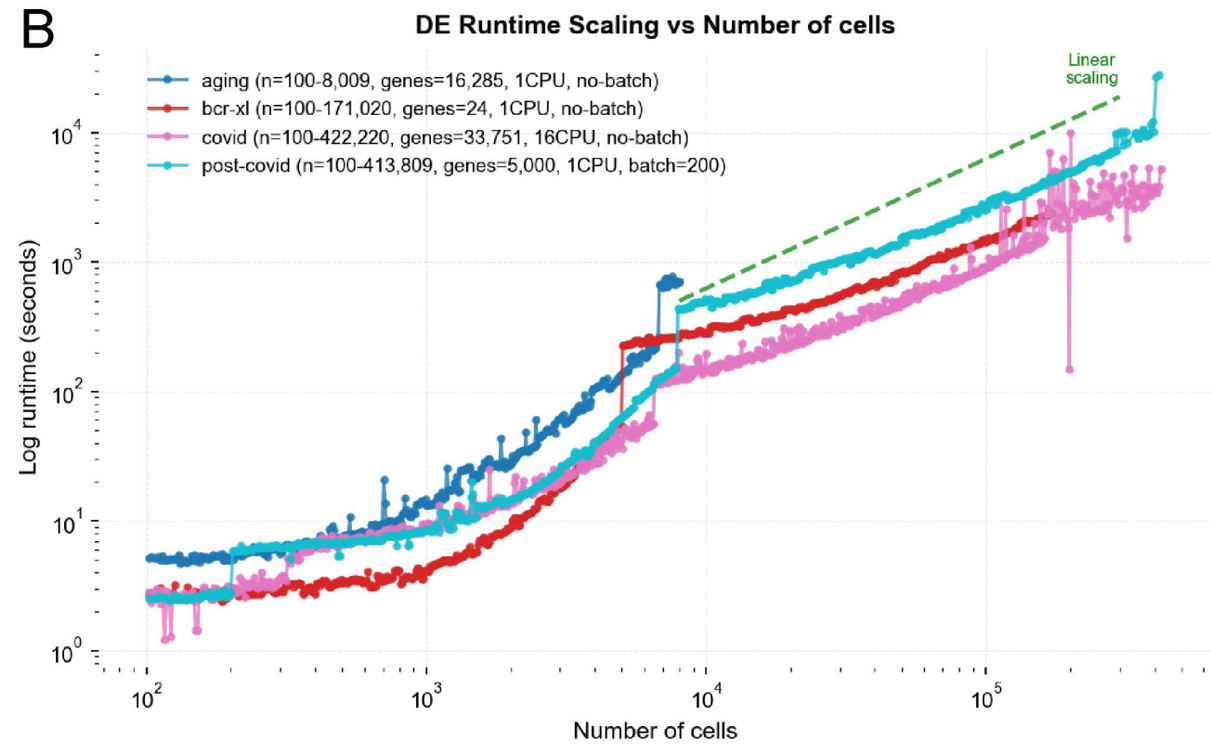

A
HSC
LMPP
Myelo P
Neutrophil
Mono P
Monocyte
cDC
MKP
MBE
Ery P
CLP
pDC
Pre-B cell
Immature B cell
Mature B cell
NK
T cell
Treg
ILC
B
Young
Mid
Old
C
Young to Mid
Youg to Old
Mid to Old
-log10 PTP
Density log-fold change
D
Fold Changes
Up
Down
E
Abundancy Changes
Decreased
Increased
None Significant
F
Differential Abundance Volcano Plot
-log10 PTP
Log Fold Change: Young to Old
Non-significant
HSC
Immature B cell
LMPP
MKP
Mature Naive B cell
Memory B cell
Plasma cell
CD8 TEM
Naive CD8 T cell
CD4 T cell
Treg
G
Direction of Change: Young to Old
Percentage (%)
Direction
up
neutral
down
Progenitors
Mature cell types
H
HSC
LMPP
NA
Density LFC
Pseudotime

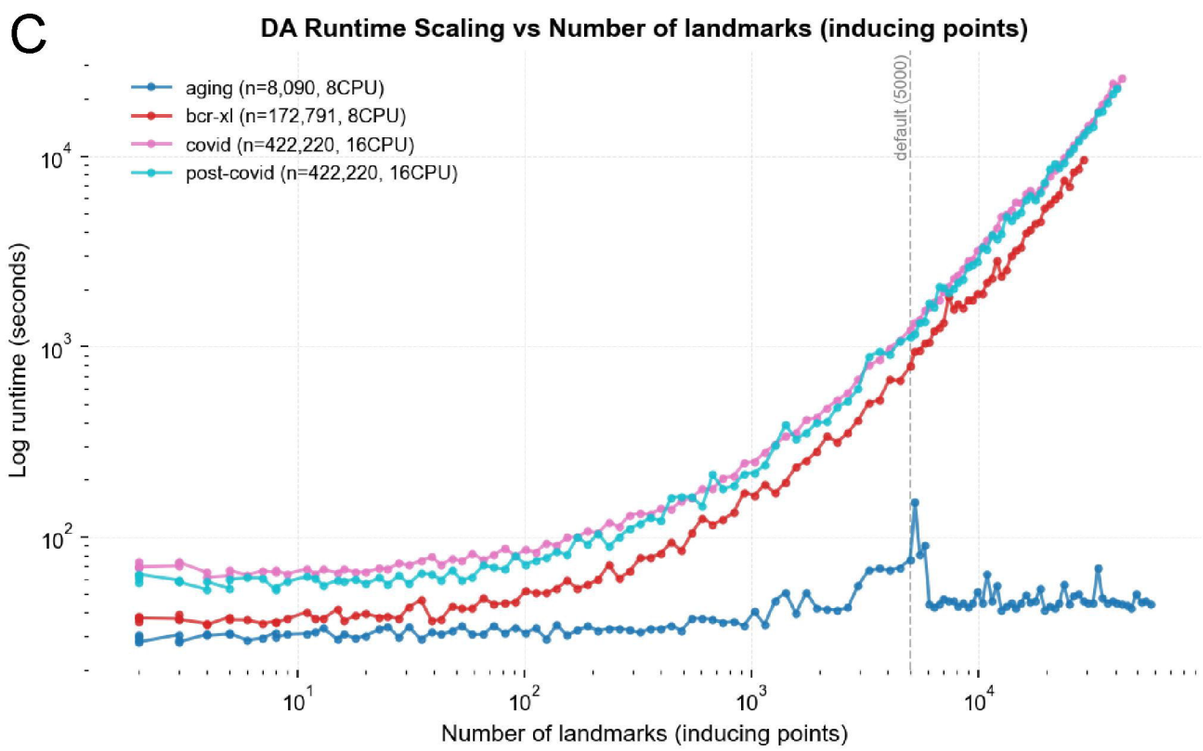

A
Age
Mid
Old
B
Gene Expression Changes in HSCs
Mahalanobis Distance
Log Fold Change: Mid to Old
Expression Change
Down Regulated
Up Regulated
Other
C
Consistency with Aging Signature
Mahalanobis Significance
Kompot log Fold Change
Local Avg Aging Consistency
E
Age
Young
Old
F
Gene Expression Changes in Monocytes
Mahalanobis Distance
Log Fold Change: Young to Old
D
Cd52
Density LFC
Pseudotime
Ifitm1
Density LFC
Pseudotime
Expression LFC
G
Genes with higher expression in Young
Genes with higher expression in Old
Biological Process (Gene Ontology) enrichment
Defense response
Regulation of kinase activity
Cellular response to cytokine stimulus
Regulation of phosphorylation
Biological process involved in interspecies interaction between organisms
Response to external stimulus
Response to biotic stimulus
Response to stress
Immune system process
Regulation of cell activation
Antigen processing and presentation
Antigen processing and presentation of peptide antigen
Immune system process
Biological process involved in interspecies interaction between organisms
Response to biotic stimulus
Response to other organism
Response to cytokine
Antigen processing and presentation of endogenous peptide antigen
Response to protozoan
Adhesion of symbiont to host
Signal
Gene count
H
HSC
LMPP
GMP
Myelo P
Neutrophil
Mono P
cMoP
Monocyte
cDC
MKP
MBEP
Basophil Progenitor
Basophil
Ery P
CLP
pDC
Pro-B cell
Pre B cell
Pre-B II cell
Immature B cell
Mature Naive B cell
Memory B cell
ILC
I
Volcano Plot: Young vs Old
Mahalanobis Distance
Log Fold Change: Young to Old
J
High Fold Change
High Mahalanobis Distance
Low Fold Change
H2-Q7
H2-Q6
AW112010
Gm47283
Cirbp
H2-K1
I627
Fos
Psmb9
H2-Ab1
Dhxrs7
Mef2c
Sirt3
Vsir
Gapdh
Ptprc
Clca3a1
Fcer1g
Ighm
Nrgn
Macf1
Fam111a
Lgals1
Hspa8
Hsph1
Dnaja1
Gm43305
Fkbp5
Banp
Dynll1
Fold Change
Higher in Old
No change
Higher in Young
Log Fold Change
Neutrophil
cDC
Monocyte
cMoP
Basophil
Basophil Progenitor
Pre-B cell
Pro-B cell
Immature B cell
Pre-B II cell
pDC
Mature Naive B cell
Memory B cell
MBEP
HSC
Ery P
MKP
GMP
CLP
LMPP
Mono P
Myelo P
K
H2-Q7
Mef2c
Dynll1
Young
Old
log fold change
expression
L
Molecular Function (Gene Ontology) enrichment
Antioxidant activity
Protein binding
Enzyme binding
Binding
Enzyme regulator activity
Identical protein binding
Protein-containing complex binding
DNA-binding transcription activator activity, RNA polymerase II-specific
Enzyme inhibitor activity
Histone acetyltransferase binding
FDR
Gene count
Groups at similarity 0.8
Signal
M
Alox5ap
Apoe
Gsr
Hp
Mgst1
Prdx2
Prdx5
Prdx6
S100a8
S100a9
Selenow
Fold Change
Higher in Old
No change
Higher in Young
Log Fold Change
Neutrophil
Mono P
Myelo P
Ery P
HSC
MKP
MBEP
CLP
GMP
LMPP
Memory B cell
Immature B cell
Mature Naive B cell
Pro-B cell
Pre-B II cell
Pre-B cell
Monocyte
cMoP
cDC
pDC
Basophil
Basophil Progenitor

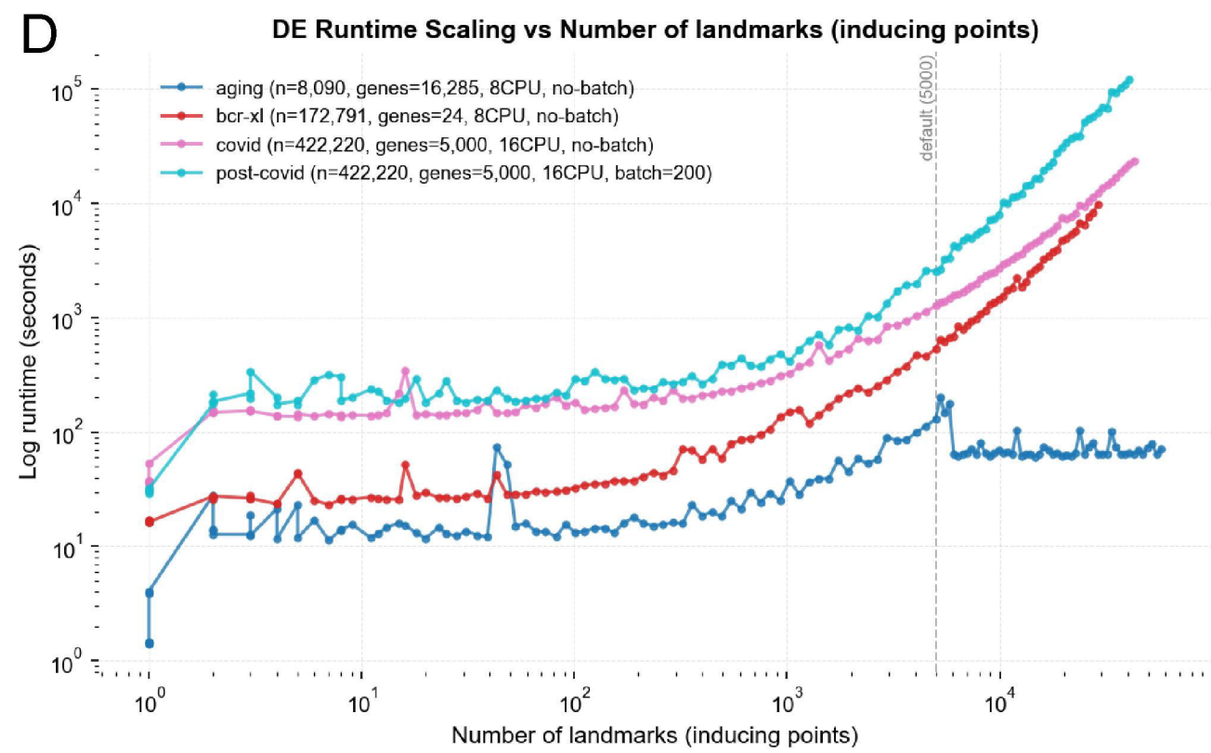

A
CD8 T cell states
ActivatedCycling
ActivatedTransitional
EarlyActivated/Memory
Naive/Memory
TexIntermediate
TexIntermediate/Effector
TerminalExh/TCR
TerminalExh/TexProgenitor
TissueResident
Treatment
Nivo
Rela
Rela + Nivo
B
Nivo
Rela
Rela + Nivo
Baseline
Response
C
Nivo
Rela
Rela + Nivo
Differential Abundance
Up in Response
Dn in Response
D
Differential Abundance
LFC: Baseline to Response
4
−6
Nivo
Rela
Rela + Nivo
ActivatedCycling
ActivatedTransitional
EarlyActivated/Memory
Naive/Memory
TexIntermediate
TexIntermediate/Effector
TerminalExh/TCR
TerminalExh/TexProgenitor
TissueResident
E
Nivo
Rela
Rela + Nivo
Up in Response
Dn in Response
Differential Expression
Mahalanobis Distance
50
0
-0.6
0.6
Gene expression log fold change
F
Treatment
Celltype
Activated Cycling
Rela + Nivo
Activated Cycling
Nivo
Rela
G
T cell activation genes
Activated Cycling
Rela + Nivo
Activated Cycling
Nivo
Rela
H
I
Nivo
Rela + Nivo
LAG3
JUNB

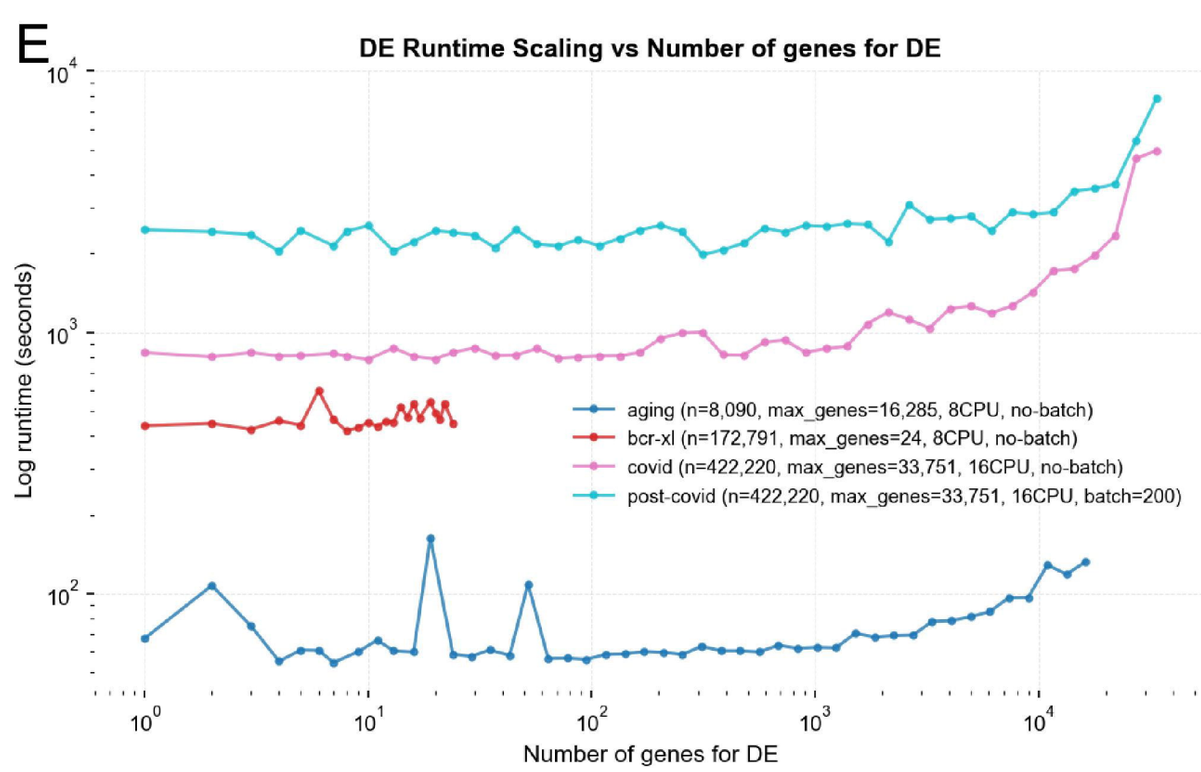

A
B
ADRN-Proliferating
ADRN-Baseline
ADRN-Dopaminergic
ADRN-Calcium
Interm-OxPhos
MES
C
Diagnosis (DX)
Post-induction chemotherapy (PTX)
D
Patient Id
E
Patient Id
Cell type
Event Category
EventFree
Event
NA
F
-log10(p-value)
Log Fold Change: DX to PTX
G
Percentage differentially abundant
ADRN-Proliferating
ADRN-Baseline
ADRN-Dopaminergic
ADRN-Calcium
Interm-OxPhos
MES
Up in PTX
Dn in PTX
H
Event Free
Event
ALK Mut
Percentage differentially abudnant


A
DA Runtime Scaling vs Number of cells
aging (n=100-8,009, 1CPU)
bcr-xl (n=100-171,020, 1CPU)
covid (n=100-422,220, 16CPU)
post-covid (n=100-413,809, 1CPU)
Linear scaling
Log runtime (seconds)
Number of cells


B
DE Runtime Scaling vs Number of cells
aging (n=100-8,009, genes=16,285, 1CPU, no-batch)
bcr-xl (n=100-171,020, genes=24, 1CPU, no-batch)
covid (n=100-422,220, genes=33,751, 16CPU, no-batch)
post-covid (n=100-413,809, genes=5,000, 1CPU, batch=200)
Linear scaling
Log runtime (seconds)
Number of cells


C
DA Runtime Scaling vs Number of landmarks (inducing points)
aging (n=8,090, 8CPU)
bcr-xl (n=172,791, 8CPU)
covid (n=422,220, 16CPU)
post-covid (n=422,220, 16CPU)
default (5000)
Log runtime (seconds)
Number of landmarks (inducing points)


D
DE Runtime Scaling vs Number of landmarks (inducing points)
aging (n=8,090, genes=16,285, 8CPU, no-batch)
bcr-xl (n=172,791, genes=24, 8CPU, no-batch)
covid (n=422,220, genes=5,000, 16CPU, no-batch)
post-covid (n=422,220, genes=5,000, 16CPU, batch=200)
default (5000)
Log runtime (seconds)
Number of landmarks (inducing points)


E
DE Runtime Scaling vs Number of genes for DE
aging (n=8,090, max_genes=16,285, 8CPU, no-batch)
bcr-xl (n=172,791, max_genes=24, 8CPU, no-batch)
covid (n=422,220, max_genes=33,751, 16CPU, no-batch)
post-covid (n=422,220, max_genes=33,751, 16CPU, batch=200)
Log runtime (seconds)
Number of genes for DE