



DestVI identifies continuums of cell types in spatial transcriptomics data

Romain Lopez^{1,9}, Baoguo Li^{2,9}, Hadas Keren-Shaul^{3,9}, Pierre Boyeau¹, Merav Kedmi³, David Pilzer³, Adam Jelinski², Ido Yofe², Eyal David², Alon Wagner¹, Can Ergen¹, Yoseph Addadi^{1D,3}, Ofra Golani^{1D,3}, Franca Ronchese^{1D,4}, Michael I. Jordan^{1D,2,5}, Ido Amit^{1D,2,10}✉ and Nir Yosef^{1D,6,7,8,10}✉

Most spatial transcriptomics technologies are limited by their resolution, with spot sizes larger than that of a single cell. Although joint analysis with single-cell RNA sequencing can alleviate this problem, current methods are limited to assessing discrete cell types, revealing the proportion of cell types inside each spot. To identify continuous variation of the transcriptome within cells of the same type, we developed Deconvolution of Spatial Transcriptomics profiles using Variational Inference (DestVI). Using simulations, we demonstrate that DestVI outperforms existing methods for estimating gene expression for every cell type inside every spot. Applied to a study of infected lymph nodes and of a mouse tumor model, DestVI provides high-resolution, accurate spatial characterization of the cellular organization of these tissues and identifies cell-type-specific changes in gene expression between different tissue regions or between conditions. DestVI is available as part of the open-source software package scvi-tools (<https://scvi-tools.org>).

Spatial transcriptomics (ST) opens up new opportunities to define the organization of cellular niches and cross-talk that modulate cellular function¹. It has been applied to study the organization of complex tissues, such as the mouse brain² and the human heart³. Research on human pathologies, such as tumors, is also an important avenue for ST^{4,5} because the tumor microenvironment (TME) consists of a rich milieu of cell types and states organized in different anatomical niches. A variety of experimental assays exist for performing ST analyses of tissue sections, although all assays are data rich and require automated and quantitative computational analyses. For example, fluorescence in situ hybridization (FISH)-based methods (MERFISH⁶, osmFISH^{2,6} and seqFISH⁷) have near single-transcript resolution but rely on pre-selected marker genes. On the other hand, sequencing-based methods (Slide-Seq^{8,9} and 10x Visium¹⁰) provide measurements of the whole transcriptome, although, in current versions, a single spot may include multiple cells, thereby providing a low-resolution view of their cellular content.

To overcome this limitation, such datasets are often matched with a single-cell RNA sequencing (scRNA-seq) dataset from the same tissue. The convention for analyzing such pairs of datasets^{9–15} is to first infer a dictionary of cell types from the scRNA-seq data and then estimate the proportion of each cell type within each spot using a linear model. This approach has had promising results in brain tissue sections for which the diversity of cellular composition is well captured by a discrete view of cell types¹⁶. In more challenging application scenarios where there is no clear way to stratify cells into discrete types or subtypes (for example, after different inflammatory signals)¹⁷, current algorithms leave the user with the choice of setting the granularity in which the data are to be analyzed. However, there is an inherent tradeoff: deeper clustering of the

scRNA-seq data provides more granular transcriptomic resolution but makes the deconvolution problem more difficult and the results potentially less accurate.

Here we introduce DestVI, a Bayesian model for multi-resolution deconvolution of cell types in ST data. Unlike other methods, DestVI learns both discrete cell-type-specific profiles and continuous sub-cell-type latent variations using a conditional deep generative model¹⁸. This way, it recovers cell type proportions (CTPs) as well as a cell-type-specific snapshot of the transcriptional state at every spot. DestVI features a post hoc analysis pipeline that helps guide different forms of downstream analysis by highlighting the main axes of spatial variation. This pipeline also enables the user to extract molecular signatures that characterize a given tissue section or different areas inside the same tissue using cell-type-specific differential expression (DE).

We show that DestVI outperforms discrete deconvolution approaches applied at different levels of cell state granularity on simulations and real data from several popular ST protocols. We then apply DestVI to two very different biological models using 10x Visium measurements: the murine lymph node and a mouse tumor model. In the lymph node, DestVI recapitulates the spatial organization of cell types at steady state and accurately identifies the spatial and transcriptional organization of monocytes that are activated upon immunization¹⁹. In the mouse tumor model, DestVI delineates the spatial coordinates of main immune cells within the TME and identifies a discrete sub-population of immunosuppressive macrophages, expressing a hypoxia activation signature within the tumor core²⁰. DestVI is implemented in the scvi-tools package²¹ and is readily available as open source, along with the accompanying tutorials.

¹Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley CA, USA. ²Department of Immunology, Weizmann Institute of Science, Rehovot, Israel. ³Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel. ⁴Malaghan Institute of Medical Research, Wellington, New Zealand. ⁵Department of Statistics, University of California, Berkeley, Berkeley CA, USA. ⁶Center for Computational Biology, University of California, Berkeley, Berkeley CA, USA. ⁷Chan Zuckerberg Biohub, San Francisco CA, USA. ⁸Ragon Institute of MGH, MIT and Harvard, Cambridge MA, USA. ⁹These authors contributed equally: Romain Lopez, Baoguo Li, Hadas Keren-Shaul. ¹⁰These authors jointly supervised this work: Ido Amit, Nir Yosef. ✉e-mail: ido.amit@weizmann.ac.il; niryosef@berkeley.edu

Results

Multi-resolution deconvolution for spatial transcriptomics. DestVI uses two different latent variable models (LVMs)²² for inferring CTPs as well as cell-type-specific continuous sub-states. DestVI takes as input a pair of datasets: a query ST data and a reference scRNA-seq data from the same tissue, annotated with cell type labels (Fig. 1). The output consists of the CTPs for every spot and a continuous estimate of cell state for every cell type in every spot (Fig. 1a). Such information may then be used for downstream analysis and formulation of biological hypotheses (Methods).

To model the reference scRNA-seq data, the first LVM (single-cell latent variable model (scLVM); Fig. 1b) of DestVI posits that, for each gene g and cell n , the number of observed transcripts, x_{ng} , follows a negative binomial distribution²³. The distribution is parameterized as (r_{ng}, p_g) , with mean $\frac{p_g r_{ng}}{(1-p_g)}$, where p_g is a gene-specific parameter determining the mean-variance relationship at every spot. The rate parameter r_{ng} of the negative binomial depends on the overall number of detected molecules in the cell (l_n), its assigned cell type (c_n) and a low-dimensional latent vector γ_n that captures the variability within cells of that type. A neural network f maps γ_n and c_n to a high-dimensional latent vector ρ_n that represents the relative abundance of each gene, such that $r_{ng} = l_n \rho_{ng}$ (Methods). scLVM is closely related to scVI²⁴, although here we capture transcriptional variation that is cell type specific. scLVM is fit with amortized variational inference²⁵, after which we obtain, for every cell, a distribution $q_\phi(\gamma_n | c_n, x_n)$ that quantifies the cell state (Methods).

To model the ST data, the second LVM (spatial transcriptomic latent variable model (stLVM); Fig. 1c) of DestVI posits that, for each gene g and each spot s , the number of observed transcripts, x_{sg} , also follows a negative binomial distribution (as in refs. ^{13,15}). The rate parameter r_{sg} depends on factors that capture technical variation (α_g —a multiplicative factor to correct for gene-specific bias between spatial and scRNA-seq measurements) and biological variation, decomposed over cell types. The latter factors come at two levels: β_s^c is a scalar proportional to the relative part of cells of type c inside the spot, and γ_s^c is a low-dimensional vector that estimates the average state of these cells. To facilitate consistency with the scRNA-seq measurements, stLVM uses the same decoder neural network trained by scLVM and an empirical prior for γ_s^c , based on the cell-type-specific values of γ inferred for the scRNA-seq data²⁶. To fit stLVM, DestVI relies on an amortized maximum a posteriori (MAP) inference scheme, in which the parameters for the CTP are kept as free, but the γ_s^c values are tied with a neural network (Methods). The CTPs π_s^c are obtained by normalizing β_s^c (to sum to 1).

The resulting model enables several types of tasks for analysis of novel datasets, some of which we implemented in an automated pipeline for analysis downstream of DestVI. In the first part of this pipeline, we report CTPs for each cell type in every spot and then

highlight cell types that tend to occur at specific niches using Geary's C autocorrelation. In the second part, we identify variability within cell types, thus going beyond the functionality that is available in current pipelines. We first select, for each type, spots that have sufficiently high proportion of cells of that type and then, with this constrained view, report the main axes of variation in each cell type, using the spatial covariates (Methods). This analysis helps highlight and visualize the most dominant transcriptional programs in every cell type as well as exploring their dependence on the cells' locations. We also propose a cell-type-specific DE procedure, helpful for comparative analysis of tissues with different conditions or areas within the same tissue (Methods).

Validation of DestVI using simulated data. Most benchmarking analyses of deconvolution focus on the ability of a given algorithm to recapitulate the proportion of cell types in every spot and, therefore, use as benchmark synthetic 'spot' measurements by sampling cells from different clusters of an scRNA-seq dataset with a given ground truth proportion. However, to fully assess the performance of DestVI to infer continuous cell states in addition to CTPs, we built a more nuanced simulation framework that also accounts for variability within cell types (Fig. 2a and Methods). In this simulation scheme, each spot is defined by a CTP as well as the state of cells in every type. We simulate the continuum of cell states as a low-dimensional linear LVM for every cell type with a negative binomial likelihood, in which the model parameters are learned on an scRNA-seq dataset. We then generate a spatial dataset by sampling, for each spot, its CTP and the low-dimensional cell state representations (Methods and Supplementary Fig. 1).

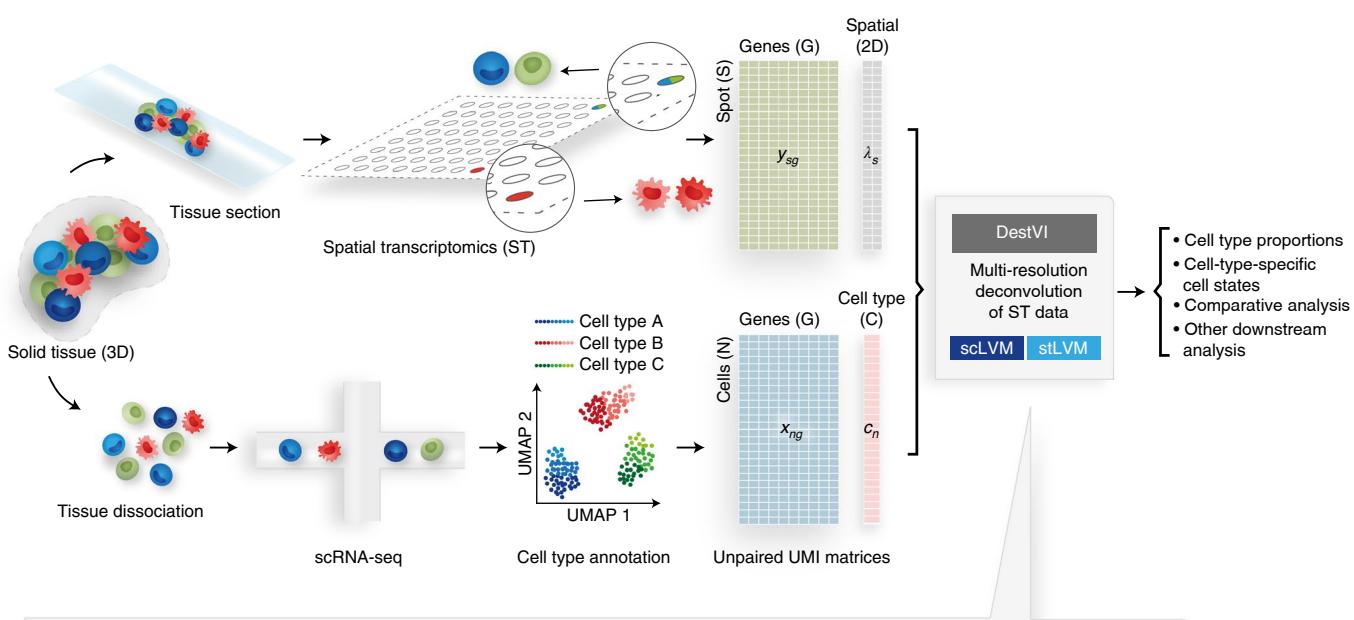
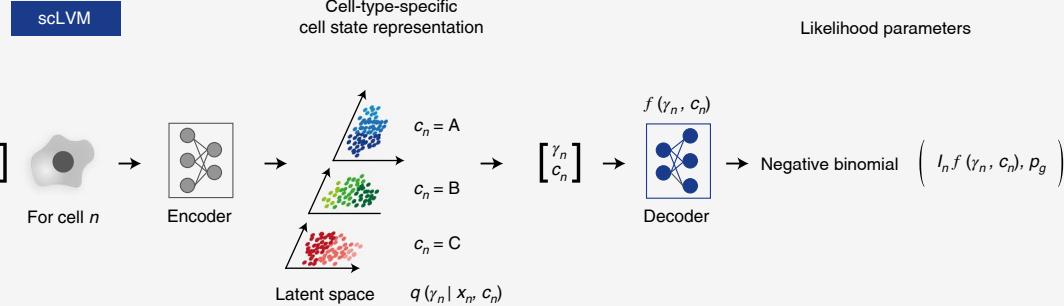
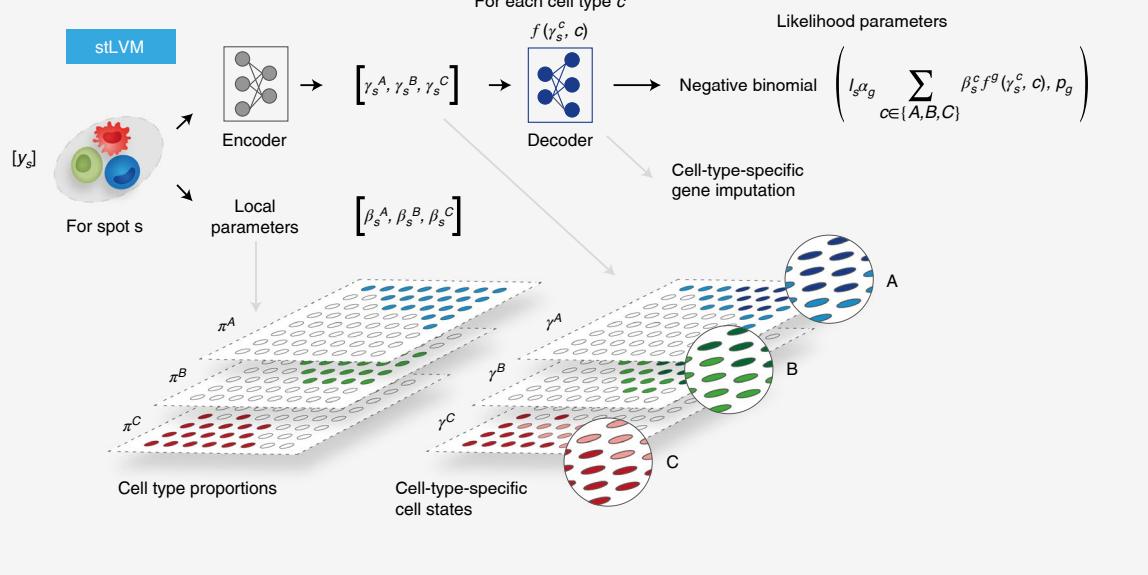
We benchmarked DestVI against several methods, evaluating the accuracy of the inferred CTP and cell-type-specific gene expression at each simulated spot. First, we applied discrete deconvolution approaches—RCTD¹¹, SPOTLight¹², Stereoscope¹³ and Seurat²⁷—with different levels of sub-clustering (Fig. 2b,c). The lowest resolution of clustering corresponds to the original cell types, and every subsequent resolution further partitions each type into distinct states (Supplementary Fig. 2). We also benchmarked against a second set of methods that embeds the ST data with the scRNA-seq data—scVI²⁴, Harmony²⁸ and Scanorama²⁹—for which inferences about CTP or cell-type-specific gene expression are done using k-nearest neighbors imputation (Methods).

In terms of CTP estimation (Fig. 2d), we found that the embedding methods have a generally lower performance than methods specifically designed for deconvolution. This has been already reported (for example, ref. ¹⁵) and is expected, as these embedding methods (for example, scVI) do not explicitly consider that spatial spots may include a mixture of cell types. Among the deconvolution methods, the effect of the clustering resolution on accuracy depends on the algorithm (a slight increase for Stereoscope and RCTD but a slight

Fig. 1 | Schematic representation of the ST analysis pipeline with DestVI. **a**, A ST analysis workflow relies on two data modalities, producing unpaired transcriptomic measurements, each in the form of count matrices. The ST data measures the gene expression y_s in a given spot s and its location λ_s . However, each spot may contain multiple cells. The scRNA-seq data measure the gene expression x_n in a cell n , but the spatial information is lost because of tissue dissociation. After annotation, we may associate each cell with a cell type c_n . These matrices are the input to DestVI, composed of two LVMs: the scLVM and the stLVM. DestVI outputs a joint representation of the single-cell data and the spatial data by estimating the proportion of every cell type in every spot and projecting the expression of each spot onto cell-type-specific latent spaces. These inferred values may be used for performing downstream analysis, such as cell-type-specific DE and comparative analyses of conditions. **b**, Schematic of the scLVM. RNA counts and cell type information from the scRNA-seq data are jointly transformed by an encoder neural network into the parameters of the approximate posterior of γ_n , a low-dimensional representation of cell-type-specific cell state. Next, a decoder neural network maps samples from the approximate posterior of γ_n along with the cell type information c_n to the parameters of a count distribution for every gene. The superscript notation f^g denotes the g -th entry ρ_{ng} of the vector ρ_n . **c**, Schematic of the stLVM. RNA counts from the ST data are transformed by an encoder neural network into the parameters of the cell-type-specific embeddings γ_s^c . Free parameters β_s^c encode the abundance of cell type c in spot s and may be normalized into CTP π_s^c (Methods). The decoder from the scLVM model maps cell-type-specific embeddings γ_s^c to estimates of cell-type-specific gene expression. These values are summed across all cell types, weighted by the abundance parameters β_s^c , to obtain the parameter r_{sg} approximating the gene expression of the spot. After training, the decoder may be used to perform cell-type-specific imputation of gene expression across all spots.

decrease for SPOTLight). However, DestVI compares favorably to all methods, regardless of their level of clustering. For imputation of cell-type-specific gene expression (Fig. 2e and Supplementary Fig. 3), the performance of deconvolution methods depends more

heavily on the level of clustering. Indeed, these approaches outperform embedding-based methods at higher resolution (more than four clusters per cell type) but perform less well at low sub-clustering resolution. Adding more cell states for deconvolution is, however,

a**b****c**

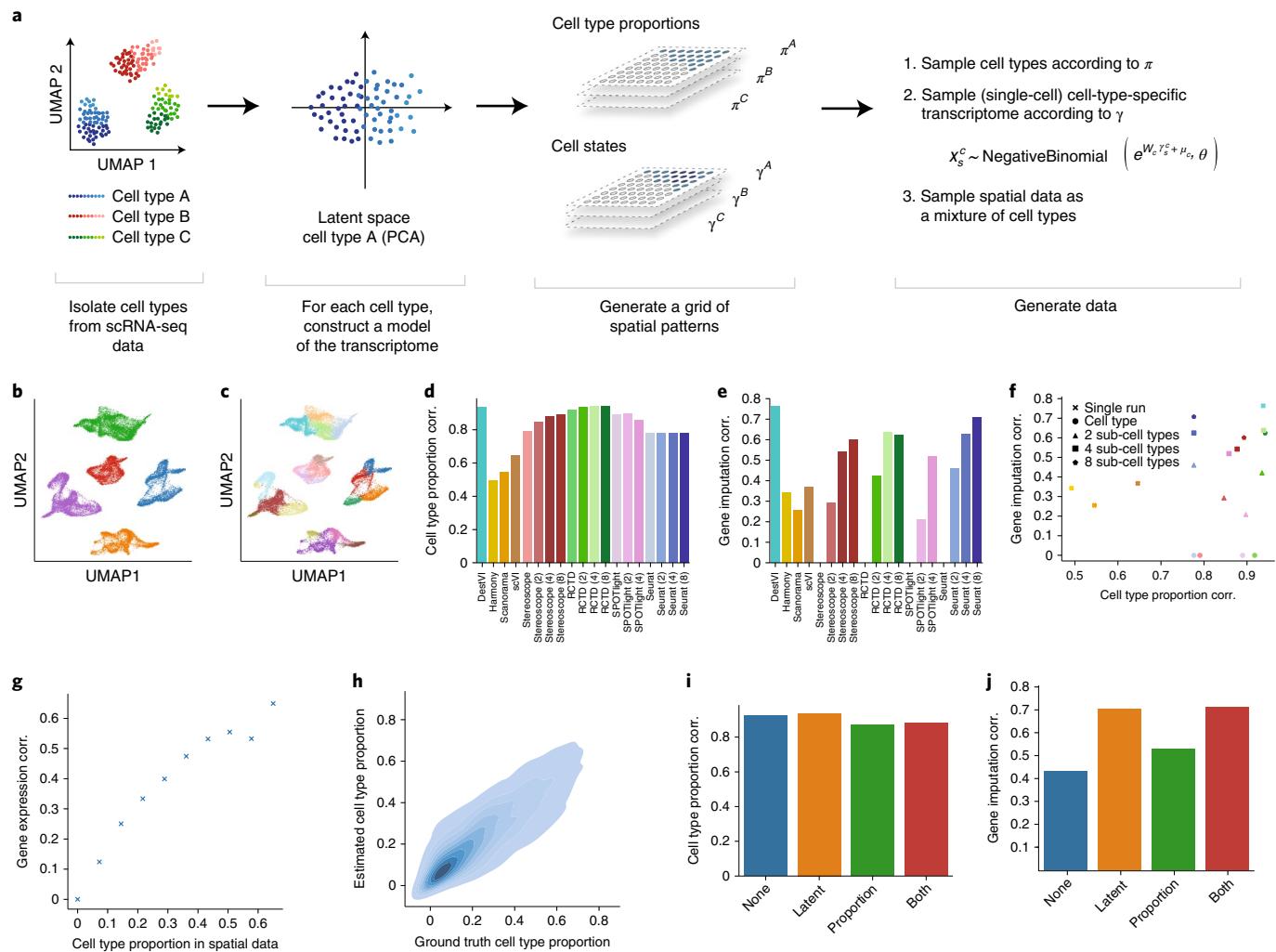


Fig. 2 | Evaluating the performance on DestVI on simulations. **a**, Schematic view of the semi-simulation framework. For each cell type of an scRNA-seq dataset, we learned a continuous model of gene expression. We sampled spatially relevant random vectors on a grid to encode the proportion of every cell type in every spot π_s^c as well as the cell-type-specific embeddings γ_s^c . Then, we feed those parameters into the learned continuous model to generate ST data (Methods). **b, c**, Visualization of the single-cell data and the cell state labels used for comparison to competing methods (UMAP embeddings of the single-cell data; 32,000 cells). **b**, Cells are colored by cell type. **c**, Cells are colored by the sub-cell types, obtained via hierarchical clustering (five clusters). **d–f**, Comparison of DestVI to competing algorithms, possibly applied to different clustering resolutions. Performance is not reported for cases that did not terminate by 3 hours (SPOTLight with eight sub-clusters; Methods). **d**, Spearman correlation of estimated CTP compared to ground truth for all methods. **e**, Spearman correlation of estimated cell-type-specific gene expression compared to ground truth, for combinations of spot and cell type for which the proportion is >0.4 for the parent cluster (not applicable to algorithms run at the coarsest level, as they do not provide cell type proportions at any sub-cell-type level). **f**, Scatter plot of both metrics that shows the tradeoff reached by all methods. Colors in this panel are in concordance with the ones from **e** and **f**. **g, h**, Follow-up stress tests for DestVI. **g**, Accuracy of imputation, measured by Spearman correlation as a function of the cell type proportion in a given spot. **h**, Head-to-head comparison of estimated cell type proportion against ground truth across all spots and cell types (8,000 combinations of spot and cell type). **i, j**, Ablation studies for the amortization scheme used by DestVI. ‘None’ stands for vanilla MAP inference. ‘Latent’ and ‘Proportion’ refer to only the inference of the latent variables and only the cell type abundance being amortized with a neural network, respectively. ‘Both’ refers to fully amortized MAP inference. **i**, Spearman correlation of estimated CTP compared to ground truth. **j**, Spearman correlation of estimated cell-type-specific gene expression compared to ground truth. UMAP, uniform manifold approximation and projection.

prone to complications, such as computational burden (for example, SPOTLight did not terminate after 3 hours with eight sub-clusters per cell type) and performance plateau (for example, RCTD’s correlation did not improve using eight sub-clusters compared to four sub-clusters). Because it is difficult, in practice, to estimate the number of sub-clusters, it limits the ability of discrete deconvolution to interpret spatial data. Conversely, the sub-clustering-free approach of DestVI outperforms all methods in this task. Because none of our benchmark methods exploits the hierarchical nature of the input sub-populations (that is, clusters and sub-clusters), we implemented

an additional baseline that exploits this information and show that DestVI compares favorably to it (Supplementary Note 1).

In summary, we found that DestVI is able to provide more accurate estimations, both at the discrete level of CTP and at the continuous level of cell-type-specific gene expression (Fig. 2f). Of note, we observe (as expected) that the ability of DestVI to predict cell-type-specific gene expression decreases for cell populations that occupy a small fraction of a spot (Fig. 2g). We do observe, however, much less effect on the accuracy of cell type proportion estimates (Fig. 2h). DestVI can, therefore, provide an internal

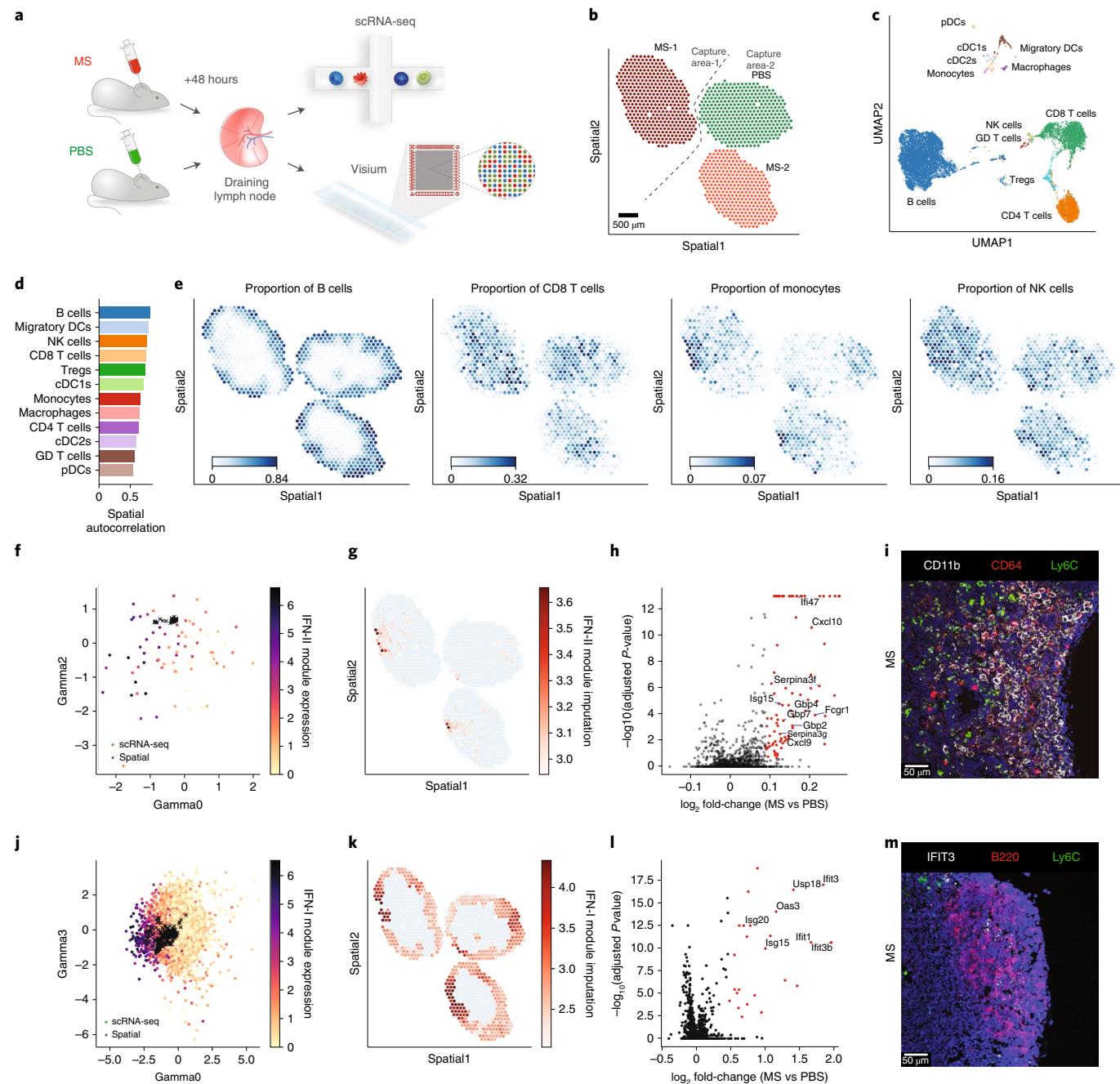


Fig. 3 | Application of DestVI to the murine lymph nodes. **a**, Schematics of the experimental pipeline. We processed murine lymph nodes with ST (10x Visium) and scRNA-seq (10x Chromium) after 48-hour stimulation by MS compared to PBS control (two sections from each condition). **b**, ST data (1,092 spots; only three sections passed the quality check) (Supplementary Methods). Sample MS-1 and samples PBS/MS-2 were processed on different capture areas of the same Visium gene expression slide. **c**, UMAP projection of the scRNA-seq data (14,989 cells). **d**, Spatial autocorrelation of the CTP. **e**, Spatial distribution of CTP for B cells, CD8 T cells, monocytes and NK cells, as inferred by DestVI. **f**, Embedding of the monocytes (circles; 128 single cells) alongside the monocyte-abundant spots (crosses; 79 spots). Single cells are colored by expression of IFN-II genes identified by Hotspot (Fcgr1, Cxcl9 and Cxcl10; Supplementary Figs. 12–14). **g**, Imputation of monocyte-specific expression of the IFN-II marker genes for the monocyte-abundant spots of the spatial data (log-scale). **h**, Monocyte-specific DE analysis between MS and PBS lymph nodes (2,000 genes, 79 spots, total 10,980 samples from the generative model). Red dots designate genes with statistical significance, according to our DE procedure (two-sided Kolmogorov-Smirnov test, adjusted for multiple testing using the Benjamini-Hochberg procedure; Methods). **i**, Immunofluorescence imaging from an MS lymph node, staining for CD11b, CD64 and Ly6C in the IFA. Scale bar, 50 μm. **j**, Embedding of the B cells (circles, 8,359 single cells) alongside the B-cell-abundant spots (crosses, 579 spots). Single cells are colored by expression of the IFN-I genes identified by Hotspot (Ifit3, Ifit3b, Stat1, Ifit1, Usp18 and Isg15; Supplementary Figs. 17 and 18). **k**, Imputation of B-cell-specific expression of the IFN-I gene module on the spatial data (log-scale), reported on B-cell-abundant spots. **l**, B-cell-specific DE analysis between MS and PBS lymph nodes (2,000 genes, 579 spots, 6,160 samples). Red dots designate genes with statistical significance, according to our DE procedure (two-sided Kolmogorov-Smirnov test, adjusted for multiple testing using the Benjamini-Hochberg procedure; Methods). **m**, Immunofluorescence imaging from an MS lymph node, staining for IFIT3, B220 and Ly6C in B cell follicle near the inflammatory IFA. Scale bar, 50 μm. UMAP, uniform manifold approximation and projection. cDC2, type 2 conventional dendritic cell; GD, gamma delta; pDC, plasmacytoid dendritic cell.

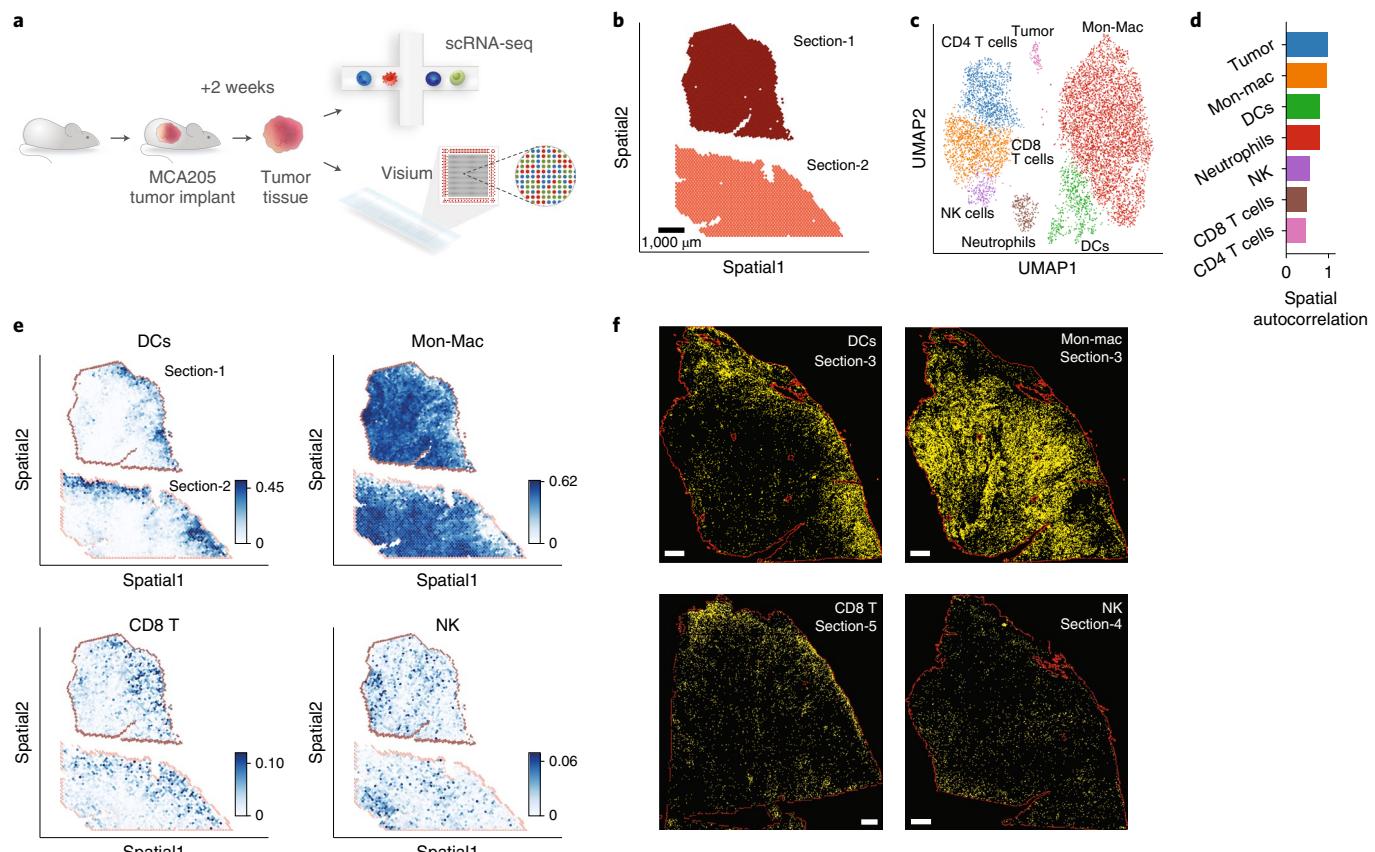


Fig. 4 | Application of DestVI to a MCA205 tumor sample. **a**, Schematics of the experimental pipeline. We performed ST (10x Visium) and scRNA-seq (single-cell MARS sequencing protocol) on MCA205 tumor that contains heterogeneous immune cell populations 14 days after intracutaneous transplantation into the wild-type mouse (two sections). **b**, Visualization of the ST data for two MCA205 tumor sections, after quality control (4,027 spots). Scale bar, 1,000 μm . The two sections were processed on the different capture areas of the same Visium gene expression slide. **c**, UMAP projection of the scRNA-seq data (8,051 cells), embedded by scVI and manually annotated. **d**, Spatial autocorrelation of the CTP for every cell type, computed using Hotspot. **e**, Spatial distribution of CTP for DCs, monocytes and macrophages (Mon-Mac), CD8 T cells and NK cells. **f**, Immunofluorescence imaging from neighboring tumor sections, using antibodies for MHCII⁺ cells showing for DCs (Section-3, +20 μm from Section-2), F4/80⁺MHCII⁻ cells showing for Mon-Mac (Section-3, +20 μm from Section-2), TCRb⁺ cells showing for CD8 T cells (Section-5, +60 μm from Section-2) and NK1.1⁺ cells showing for NK cells (Section-4, +30 μm from Section-2). All scale bars denote 500 μm . Red solid lines indicate the section boundary. Right side is the MCA205 tumor marginal boundary. The cells positive for staining marker are segmented and annotated using QuPath and showing yellow color here with changed brightness and contrast (Supplementary Methods). UMAP, uniform manifold approximation and projection.

control for which spots should be removed when conducting a cell-type-specific analysis of gene expression or cell state (automated as part of our automated pipeline; Methods). We also used the simulation to benchmark variants of DestVI with several amortization schemes (that is, which parameters to treat as free or as a function of the input data) (Methods and Fig. 2*i,j*).

Validation on a single-cell spatial transcriptomics dataset. The lack of ground truth in Visium or Slide-seq experiments is an obstacle for assessing DestVI's performance on real-world datasets. However, sci-Space³⁰ provides a way to evaluate DestVI because it includes transcriptional information at a single-cell resolution with coarse spatial information (about 200 μm). For evaluation with sci-Space data, we, therefore, treated each spatial position as a single 200- μm 'spot' (summing the transcriptomes of all the included cells) and use these as an input to DestVI and the other benchmark methods. We then treated the known (yet hidden from the algorithms) cellular content of each spot as ground truth, as in our simulation analysis (Methods). For the dataset of mouse embryo cortex, the accuracy is lower than what was observed in the simulation study for all methods that we inspected. However, consistent

with our simulations, DestVI is still the top-performing method in this more realistic scenario, both in predicting CTP and in inferring cell-type-specific gene expression (Supplementary Fig. 4).

DestVI identifies a spatial immune response in lymph nodes. For a first application of DestVI, we aimed to study spatial patterns of antigen-specific immunity and profiled murine auricular lymph nodes after 48-hour stimulation by *Mycobacterium smegmatis* (MS), a Gram-positive bacteria that induces a robust CD4⁺ T cell response characterized by IFN γ ¹⁹. For spatial analysis, we used the Visium platform to profile four lymph node sections (two sections stimulated with MS and two sections from control (PBS) injections, processed on two capture areas of the same Visium slide). We profiled other lymph nodes with 10x Chromium to obtain a matching scRNA-seq dataset of MS and PBS treatments (Fig. 3*a* and Supplementary Methods).

After quality control, we discarded one PBS-treated section from further analysis as its number of unique molecular identifiers (UMIs) per spot was significantly lower than the other sections (Supplementary Fig. 5, Fig. 3*b* and Supplementary Methods). We began with a clustering analysis of the raw data using scanpy

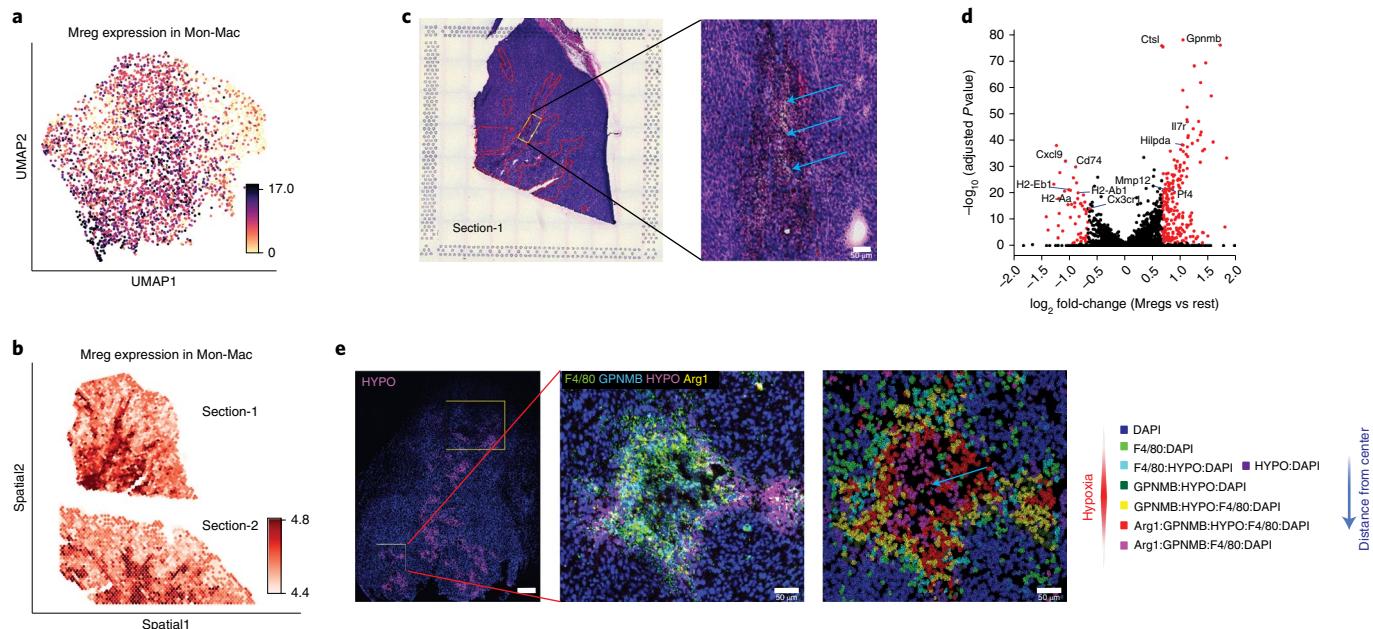


Fig. 5 | DestVI identifies a hypoxic population of macrophages in the tumor core. **a**, Visualization of the hypoxia gene expression module on the Mon-Mac cells from the scRNA-seq data (4,400 cells), on the embedding from scVI (identified using Hotspot; Supplementary Figs. 28 and 29). **b**, Imputation of gene expression for this module on the spatial dataset (log-scale), reported only on spots with high abundance of Mon-Mac (3,906 spots across the two sections). Imputation for other modules is shown in Supplementary Fig. 30. **c**, H&E-stained histology of Section-1 (left), with overlapping Mreg-identified regions from DestVI showing red polygons (as identified in Supplementary Fig. 32). Blue arrows show the location of cells from the necrotic core. H&E-stained histology showing a magnification of the necrotic core of the yellow frame in Section-1 (right). Scale bar, 55 μ m. **d**, Mon-Mac cell-specific DE analysis between the Mreg-enriched areas and the rest of the tumor section (2,886 genes; 379 spots for the Mreg-enriched area and 361 randomly sampled spots from the rest of the tumor; total of 2,220 samples from the generative model). Red dots designate genes with statistical significance, according to our DE procedure (two-sided Kolmogorov-Smirnov test, adjusted for multiple testing using the Benjamini-Hochberg procedure; Methods). **e**, Representative image of the multiplexed immunofluorescence staining. Left, Hypoxic areas as identified by the Hypoxyprobe (HYPO) in a whole MCA205 tumor section. Two yellow frames show the hypoxic areas with necrotic cores. Scale bar, 500 μ m. Middle, Magnification of a necrotic core with F4/80, Arg1, GPNMB, HYPO and DAPI staining. Scale bar, 50 μ m. Right, Annotation of different macrophages surrounding the necrotic core. Different colors shown in the legend bar show different staining combinations. Red spindle shows the extent of hypoxia. Blue arrow shows the location of cells from the necrotic core. Scale bar, 50 μ m.

(Supplementary Fig. 6a). We noted the high reproducibility of cluster compositions for the MS conditions as well as marked differences between the MS stimulated and PBS tissue sections (Supplementary Fig. 6b). The matching scRNA-seq data yielded 14,989 cells after quality filtering (Supplementary Methods and Supplementary Table 1). Data were processed by scVI²⁴ and annotated based on a public dataset³¹ using scANVI³². This first round of annotation was followed by manual labeling of rare cell types (Supplementary Methods, Fig. 3c and Supplementary Fig. 7). We observed a cellular MS-specific response, including changes in natural killer (NK) cell and monocyte abundance (Supplementary Fig. 8), as previously reported¹⁹.

With DestVI, we explored, for each cell type how infection-induced differences in transcriptional states are associated with changes of spatial organization. The reports from our automated pipeline (Supplementary Report 1) helped guide our exploration of the data, and we summarize the main findings in the remainder of this section. We began by exploring the spatial distribution of cells from each type (Fig. 3d,e and Supplementary Fig. 9). As expected, the first striking pattern is the organization of the lymph node sections into the B cell follicles (external area) and a T cell compartment (internal area) (Fig. 3e and Supplementary Fig. 10)^{33,34}. We also found that monocytes tend to form spatially coherent niches, with a stronger extent of co-localization in the stimulated lymph nodes compared to the control (comparing the Geary's C values of their inferred proportions; Fig. 3e and Supplementary

Fig. 11). This finding is consistent with our previous work, which showed that, after immunization with MS, NK cells are recruited to the lymph nodes and produce IFN γ . This signaling axis further promotes the upregulation of IFN γ signaling in monocytes¹⁹. We refer to this multi-cellular immune response niche, located between the B cell follicles as shown in Fig. 3e, as the interfollicular area (IFA)³⁵.

We next explored the inferred spatial organization of cell states within every cell type, a unique feature of our pipeline compared to other methods for spatial analysis (Fig. 3f-m). Starting with the monocyte population, analysis of the scRNA-seq data reveals interferon signaling as an important source of variation. This is indicated by Hotspot analysis, which searches for sets of co-expressed genes that distinguish different states (Supplementary Figs. 12-15). Through joint analysis with DestVI, we further contextualized this variation with the location of the monocytes and their unique spatial niche in the lymph node (Supplementary Report 1). To this end, we used our automated pipeline to study the main axes of spatial variation in the monocyte population. We found that these axes indeed reflect heterogeneity in the expression of a type II interferon response gene signature (Fcgr1, Cxcl9 and Cxcl10; Fig. 3f). Moving to the tissue coordinates, we found a clear localization of inferred monocyte-specific expression of IFN-II genes in the IFA (Fig. 3g). This visualization suggests that the amount of expression of IFN-II genes is markedly higher in the MS lymph node versus the PBS lymph node, in those constrained (IFA) niches. DE analysis across lymph nodes with DestVI supported these

differences and highlighted a rich set of additional IFN-II response genes, beyond those used in our signature analysis (for example, Gpb2, Serpina3g and Ifi47) (Methods, Fig. 3h and Supplementary Table 2). To verify this discovery using alternative spatial measurements, we performed immunofluorescence staining of lymph node from naive and MS-treated mice, with CD64 (Fcgr1, a marker for IFN γ that we identified with DestVI) and CD11b and Ly6C (markers for monocytes). In line with the results identified by DestVI, we found a higher number of inflammatory monocytes in the MS-treated lymph nodes that are localized to the IFA (Fig. 3i and Supplementary Fig. 16).

Our automated pipeline also points to spatial heterogeneity within the B cell compartment, with a strong enrichment for type I interferon signature (for example, Ifit1, Ifit3 and Isg15; Supplementary Report 1 and Fig. 3j). This heterogeneity is also observed in analysis of the scRNA-seq data alone (Supplementary Figs. 17–19). To explore this, we used the B-cell-specific gene expression estimates from DestVI to inspect the spatial organization of this signature (Fig. 3k). We found that the signature genes are expressed across all the lymph nodes but with higher levels close to the IFA. We also found that these genes are expressed at lower levels in the PBS lymph node than in the MS lymph node (Supplementary Fig. 19). This observation is also supported by DE analysis, comparing B cells in the MS lymph node versus the PBS lymph node (Fig. 3l and Supplementary Table 3) but also comparing B cells within MS samples in each interferon-enriched zone to B cells in the rest (Supplementary Table 4 and Supplementary Fig. 20). In both cases, we indeed noticed a similar strong signature of type I interferon signaling. These observations were further validated by immunofluorescence staining for the B cells and IFN-I markers IFIT3, B220 and Ly6C (Supplementary Methods). We identified IFIT3⁺B220⁺ cells on the MS sample in the B cell follicle near the inflammatory IFA (Fig. 3m) but not in control samples (Supplementary Fig. 21).

In summary, the unique features of our analysis enable robust spatial characterization of cell types and states within naive and pathogen-challenged lymph nodes. DestVI identifies a clear and specialized immune niche involving IFN signaling of different cell types, including monocytes and B cells activated by MS-infected mice and localized to the peripheral subcapsular sinus and IFA.

DestVI identifies a hypoxic state of macrophages in the tumor core. To apply DestVI to a more complex and less structured tissue, we spatially profiled a syngeneic mice tumor model (MCA205) using Visium. Fourteen days after intracutaneous transplantation of MCA205 tumor cells, we characterized the tumor using scRNA-seq and Visium (Fig. 4a, Supplementary Fig. 22 and Supplementary Methods). A total of 2,125 spots (Section-1) and 1,902 spots (Section-2) were used for analysis with DestVI after quality metrics and filtering (Fig. 4b). We also collected cells from a separate MCA205 tumor for scRNA-seq (8,051 immune cells and tumor cells; Supplementary Methods and Supplementary Table 5). We annotated the scRNA-seq data by labeling clusters of the latent space from scVI³⁴ based on marker genes (Supplementary Methods and Fig. 4c). We then explored the spatial distribution of these cell types and states using DestVI. We included the analysis generated by our automated pipeline in Supplementary Report 2.

We first inspected the inferred CTP of the major immune subsets (CD8 T cells, monocytes, macrophages, dendritic cells (DCs) and NK cells; Fig. 4d,e; remaining cell types in Supplementary Fig. 23). To validate these predictions, we performed immunofluorescence staining on three additional sections of the same tumor sample (labeled as Section-3, Section-4 and Section-5, at distances 20 μ m, 30 μ m and 60 μ m from Section-2, respectively; Supplementary Table 6). Both major types of T cells (CD4⁺ and CD8⁺) were predicted to be highly abundant on the boundary of the MCA205 tumor, a finding that was confirmed by staining for a T cell marker (TCRb) on

Section-5 (Fig. 4f)³⁶. Quantification of the fluorescent signal at different areas of the tissue further showed strong correlation to the cell type density inferred by DestVI ($\rho=0.91$; Supplementary Methods and Supplementary Fig. 24). NK cells were also predicted to occupy very specific niches, which was supported by staining for NK1.1 on Section-4 (Fig. 4f; $\rho=0.94$; Supplementary Fig. 25). Antigen-presenting DCs were also inferred to exhibit a non-uniform spatial organization, with marked localization at the boundary of the MCA205 tumor (Fig. 4d,e), a property that we validated by staining for MHC-II on Section-3 (Fig. 4f; $\rho=0.87$; Supplementary Fig. 26). Finally, monocytes and macrophages (jointly labeled as Mon-Mac, because they could not be clearly separated in the scRNA-seq data) were present broadly in the MCA205 tumor, with no specific pattern (Fig. 4d,e). We verified this by staining for F4/80 on Section-3 (Fig. 4f, $\rho=0.75$; Supplementary Fig. 27). Together, these results suggest that DestVI is able to provide a precise and detailed view of the spatial organization of major immune subsets in the MCA205 tumor.

Because the Mon-Mac population did not have a specific spatial pattern, we hypothesized that their spatial coordinates may reflect different cell states, a hypothesis that was also suggested by our automated pipeline (Supplementary Report 2). To explore this, we applied Hotspot to the Mon-Mac population from the scRNA-seq data (Supplementary Figs. 28 and 29). Two of the detected sets of co-expressed genes are specific to tumor-infiltrating myeloid suppressive cell populations that we previously identified using non-spatial analysis³⁷. The first module corresponds to tumor-associated macrophages (TAMs), expressing C1qa, C1qb, C1qc, Ms4a7 and Apoe (Supplementary Fig. 30a). The second one corresponds to a regulatory myeloid (Mreg) population discovered previously³⁷, with expression of Trem2, Gpnmb, Mmp12 and Il7r as well as markers of hypoxia Hmox1 and Hilpda (Fig. 5a).

To inspect the spatial organization of these sub-populations, we used DestVI to infer the Mon-Mac-specific expression of the corresponding gene sets in our Visium data (Fig. 5b; the spatial distribution of other modules is displayed in Supplementary Fig. 30e–h). We found that both populations were mostly abundant in the inner layers of the MCA205 tumor. This result is consistent with previous observations on both Mreg and TAM in human tumors³⁸ and murine models³⁷, and it can help interpret the association between such myeloid suppressive cells (normally identified as Arg1⁺) and poor anti-tumor response³⁹. Of these two tumor-infiltrating myeloid populations, we found a marked localization of the Mreg cells (Fig. 5b and Supplementary Fig. 31). Cell-type-specific DE analysis in the Mreg-enriched areas versus the rest of the tumor identified markers for this population that are consistent with our single-cell data and our previous study, including Ctsl, Il7r and the hypoxia-inducible gene Hilpda³⁷ (Fig. 5d and Supplementary Table 7). To understand the nature of this putatively hypoxic Mreg-enriched niche, we overlapped our Visium data with images obtained by hematoxylin and eosin (H&E) staining (Fig. 5c). We noticed that the Mreg-enriched niches strongly correspond to perinecrotic areas of the MCA205 tumor^{40–42}. This is consistent with previous work, as cells with similar phenotypes have been observed to congregate in hypoxic tissue niches (cancerous⁴³ and non-cancerous⁴⁴).

We further validated the localization of the Mreg cells with multiplexed immunofluorescence staining in additional MCA205 tumor samples ($n=12$) (Supplementary Methods). We measured the abundance of macrophage and Mreg markers (F4/80, Arg1 and GPNMB) as well as a specific probe for hypoxic regions (Hypoxyprobe; Fig. 5e). The Hypoxyprobe⁺ staining helped identify a clear pattern in which necrotic core areas are consistently surrounded by the hypoxic areas. Additional staining of F4/80, GPNMB and Arg1 showed that F4/80⁺GPNMB⁺HYP⁺ cells and F4/80⁺GPNMB⁺Arg1⁺HYP⁺ cells are enriched in hypoxic regions and in perinecrotic areas.

In summary, DestVI correctly maps cell types of immune cells onto the spot coordinates and identifies a clear and specialized niche involving metabolic changes in response to hypoxia within the macrophage population. Our further experimental validation also suggests that DestVI could be a reliable tool for exploring complex cell-type-specific phenotypic changes in different tumor models under immunotherapy treatments.

Discussion

We introduce DestVI, a multi-resolution approach to deconvolution of ST profiles using an auxiliary scRNA-seq dataset. Via simulations, we show that classical deconvolution approaches, which are based on clustering the scRNA-seq data, might be difficult to apply and can miss important information in the case of continuous variation within cell types. DestVI addresses this issue by learning cell-type-specific latent variables on the scRNA-seq data, using a deep generative model, and mapping those latent variables onto the spatial data. Coupled to the automated pipeline that we developed, DestVI can provide interpretable analyses to allow comparison of within-cell-type gene expression levels across different conditions or between different niches of the same tissue section.

In this study, we employed a fixed set of hyperparameters for DestVI. For example, we used five latent variables for the cell-type-specific cell states γ_s^c and two for the number of hidden layers in the decoder neural networks. We discuss motivations for this design as well as stability with respect to the hyperparameter choice in Supplementary Note 2.

DestVI can be further developed in terms of the modeling perspective: for instance, the output of the DE analysis could be improved by make use of a more principled false discovery rate (FDR) control procedure^{45,46} (Supplementary Note 3). Furthermore, it can be beneficial to explicitly incorporate spot locations into the stLVM model. We explore such variants of DestVI and show improvements on simulated data in Supplementary Note 4.

Several practical considerations are necessary for a successful analysis with DestVI. For scRNA-seq data generation, we refer to the recent review on deconvolution⁴⁷ and provide more specific recommendations in Supplementary Note 5. We further discuss important considerations: the many possible sources of mismatch between the two modalities that may lead to decreased performance. We discuss those, and benchmark DestVI in the setting of missing cell types, in Supplementary Note 6.

Despite the conceptual advantage of modeling continuous transcriptomic variation within cell types, one might wonder what its effect is on the hypotheses described in this manuscript. In other words, could we have gained the same biological insights with classical deconvolution approaches? To answer this question, we focus on the interferon response from the monocyte population in MS lymph nodes versus PBS lymph nodes (Fig. 3g,h) that we previously characterized¹⁹. We found that both RCTD and Stereoscope correctly identify the location of basic cell types; however, their monocyte-specific gene expression imputation fails to characterize this specific interferon response (Supplementary Note 7).

Both applications of DestVI described here focus on the 10x Visium protocol, which has typically low spatial resolution, compared to emerging technologies such as Slide-SeqV2 (ref.⁸), HDST⁴⁸ and Seq-Scope⁴⁹. Even so, analysis of data generated from most of those high-resolution methods is still challenging, because one cannot guarantee perfect spatial delineation of single cells. To illustrate this, we applied our method to a liver tissue profiled with Seq-Scope as well as refined simulations (Supplementary Note 8). DestVI recapitulates the main finding of the original Seq-Scope study and provides more aggregable cell type proportion throughout the tissue compared to the single-cell pipeline originally used. On the simulations, DestVI still performs competitively on a sparser dataset, with lower library size and fewer cell types per spot.

ST is a promising approach for unravelling cell–cell interactions⁵⁰ and other forms of cellular communication and function in a tissue³⁹. We expect that approaches such as DestVI will provide the necessary level of resolution and further enhance understanding of the local signaling environments and how they affect cell functions and spatial cues, such as interaction among specific cellular subsets, chemical gradients and metabolic cross-talk.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41587-022-01272-8>.

Received: 10 May 2021; Accepted: 7 March 2022;

Published online: 21 April 2022

References

- Wagner, A., Regev, A. & Yosef, N. Revealing the vectors of cellular identity with single-cell genomics. *Nat. Biotechnol.* **34**, 1145–1160 (2016).
- Codeluppi, S. et al. Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods* **15**, 932–935 (2018).
- Asp, M. et al. Spatial detection of fetal marker genes expressed at low level in adult human heart tissue. *Sci. Rep.* **7**, 12941 (2017).
- Hunter, M. V., Moncada, R., Weiss, J. M., Yanai, I. & White, R. M. Spatial transcriptomics reveals the architecture of the tumor/microenvironment interface. *Nat. Commun.* **12**, 6278 (2021).
- Ji, A. L. et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* **182**, 1661–1662 (2020).
- Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging: Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
- Eng, C. L. et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019).
- Stickels, R. R. et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* **39**, 313–319 (2020).
- Rodrigues, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
- Stähli, P. L. et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
- Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-021-00830-w> (2021).
- Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I. & Heyn, H. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.* **49**, e50 (2021).
- Andersson, A. et al. Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Commun. Biol.* **3**, 565 (2020).
- Song, Q. & Su, J. DSTG: deconvoluting spatial transcriptomics data through graph-based artificial intelligence. *Brief. Bioinform.* **22**, bbaa414 (2021).
- Kleshcheynikov, V. et al. Comprehensive mapping of tissue cell architecture via integrated single cell and spatial transcriptomics. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-021-01139-4> (2022).
- Ortiz, C. et al. Molecular atlas of the adult mouse brain. *Sci. Adv.* **6**, eabb3446 (2020).
- Lähnemann, D. et al. Eleven grand challenges in single-cell data science. *Genome Biol.* **21**, 31 (2020).
- Lopez, R., Gayoso, A. & Yosef, N. Enhancing scientific discoveries in molecular biology with deep generative models. *Mol. Syst. Biol.* **16**, e9198 (2020).
- Blecher-Gonen, R. et al. Single-cell analysis of diverse pathogen responses defines a molecular roadmap for generating antigen-specific immunity. *Cell Syst.* **8**, 109–121 (2019).
- Kumar, S. et al. Intra-tumoral metabolic zonation and resultant phenotypic diversification are dictated by blood vessel proximity. *Cell Metab.* **30**, 201–211 (2019).
- Gayoso, A. et al. scvi-tools: a library for deep probabilistic analysis of single-cell omics data. A Python library for probabilistic analysis of single-cell omics data. *Nat. Biotechnol.* **40**, 163–166 (2022).
- Blei, D. M. Build, compute, critique, repeat: data analysis with latent variable models. <http://www.cs.columbia.edu/~blei/papers/Blei2014b.pdf> (2014).

23. Grün, D., Kester, L. & van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11**, 637–640 (2014).
24. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
25. Kingma, D. P. & Welling, M. Auto-encoding variational Bayes. *Int. Conf. Learning Representations* (2014).
26. Tomczak, J. M. & Welling, M. VAE with a VampPrior. *Int. Conf. on Artificial Int. and Stat.* (2018).
27. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
28. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
29. Hie, B., Bryson, B. & Berger, B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat. Biotechnol.* **37**, 685–691 (2019).
30. Srivatsan, S. R. et al. Embryo-scale, single-cell spatial transcriptomics. *Science* **373**, 111–117 (2021).
31. Gayoso, A. et al. Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* **18**, 272–282 (2021).
32. Xu, C. et al. Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol. Syst. Biol.* **17**, e9620 (2021).
33. Garraud, O. et al. Revisiting the B-cell compartment in mouse and humans: more than one B-cell subset exists in the marginal zone and beyond. *BMC Immunol.* **13**, 63 (2012).
34. Hampton, H. R. & Chtanova, T. The lymph node neutrophil. *Semin. Immunol.* **28**, 129–136 (2016).
35. Kastenmüller, W., Torabi-Parizi, P., Subramanian, N., Lämmermann, T. & Germain, R. N. A spatially-organized multicellular innate immune response in lymph nodes limits systemic pathogen spread. *Cell* **150**, 1235–1248 (2012).
36. Hatfield, S. M. et al. Immunological mechanisms of the antitumor effects of supplemental oxygenation. *Sci. Transl. Med.* **7**, 277ra30 (2015).
37. Katzenelenbogen, Y. et al. Coupled scRNA-seq and intracellular protein activity reveal an immunosuppressive role of TREM2 in cancer. *Cell* **182**, 872–885 (2020).
38. Massi, D. et al. Arginine metabolism in tumor-associated macrophages in cutaneous malignant melanoma: evidence from human and experimental tumors. *Hum. Pathol.* **38**, 1516–1525 (2007).
39. Yosef, N. & Regev, A. Writ large: genomic dissection of the effect of cellular environment on immune response. *Science* **354**, 64–68 (2016).
40. Steinberg, J. D. et al. Negative contrast Cerenkov luminescence imaging of blood vessels in a tumor mouse model using ^{168}Ga gallium chloride. *EJNMMI Res* **4**, 15 (2014).
41. Wippold, F. J. 2nd, Lämmle, M., Anatelli, F., Lennerz, J. & Perry, A. Neuropathology for the neuroradiologist: palisades and pseudopalisades. *AJNR Am. J. Neuroradiol.* **27**, 2037–2041 (2006).
42. Godet, I. et al. Fate-mapping post-hypoxic tumor cells reveals a ROS-resistant phenotype that promotes metastasis. *Nat. Commun.* **10**, 4862 (2019).
43. Carmona-Fontaine, C. et al. Metabolic origins of spatial organization in the tumor microenvironment. *Proc. Natl Acad. Sci. USA* **114**, 2934–2939 (2017).
44. Duque-Correa, M. A. et al. Macrophage arginase-1 controls bacterial growth and pathology in hypoxic tuberculosis granulomas. *Proc. Natl Acad. Sci. USA* **111**, E4024–E4032 (2014).
45. Lopez, R., Boyeau, P., Yosef, N., Jordan, M. I. & Regier, J. Decision-making with auto-encoding variational Bayes. 34th Conference on Neural Information Processing Systems. <https://papers.nips.cc/paper/2020/file/357a6fd7642bf815a88822c447d9dc4-Paper.pdf> (2020).
46. Boyeau, P. et al. Deep generative models for detecting differential expression in single cells. Preprint at <https://www.biorxiv.org/content/10.1101/794289v1?rss=1> (2019).
47. Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* **22**, 627–644 (2021).
48. Vickovic, S. et al. High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019).
49. Cho, C.-S. et al. Microscopic examination of spatial transcriptome using Seq-Scope. *Cell* **184**, 3559–3572 (2021).
50. Armingol, E., Officer, A., Harismendy, O. & Lewis, N. E. Deciphering cell–cell interactions and communication from gene expression. *Nat. Rev. Genet.* **22**, 71–88 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

Methods

Model of scRNA-seq data. *Assumption and model for the single-cell data (scLVM).* Let n designate a cell in the scRNA-seq dataset. We assume that each cell is annotated with cell type label c_n , but those labels are broad enough such that the introduction of continuous covariates γ_n into the model helps explain additional variance in gene expression (that is, within-cell-type variation). For example, c_n represents a discrete cell type (for example, B cells or CD8 T cells), whereas γ_n is a continuous vector summarizing a sub-cell state of interest (for example, B cell activation and CD8 T cell exhaustion).

In the following, we assume that c_n is observed (for example, obtained via clustering) and that γ_n , however, is unobserved and treated as a latent variable. We posit the following generative model for our data:

$$\begin{aligned} \gamma_n &\sim \text{Normal}(0, I) \\ x_{ng} &\sim \text{NegativeBinomial}\left(l_n f^g(c_n, \gamma_n), p_g\right), \end{aligned}$$

where l_n is the library size of cell n , f is a two-layer neural network and p is a G -dimensional vector. f takes as input the concatenation of the one-hot encoding of c_n as well as the scalar γ and outputs a G -dimensional vector. f has a softplus non-linearity at its output to ensure positivity.

This generative model has substantial overlap with our previous proposed method, scVI²⁴, which is also a conditional deep generative model. On top of the conceptual difference that scVI conditions on the batch identifier, whereas scLVM conditions on the cell type information, there are several technical points in which these two models differ. First, we use the standard parametrization (r, p) for the negative binomial distribution: the number of successful independent and identical Bernoulli trials before r failures are achieved, in which p is the probability of failure of each Bernoulli trial. This is in contrast to scVI that relies on the mean dispersion parameterization for the negative binomial distribution and is necessary to make the definition of the spot gene expression level as the sum of contributions from individual cells correct (as emphasized in ref. ¹³). Furthermore, changes were required to the neural network architecture for the transfer learning to work adequately. Indeed, we found that using a decoder with randomness such as dropout⁵¹ or with running parameters as in batch normalization⁵² did not work, so we replaced those with layer normalization⁵³.

Variational inference. We use auto-encoding variational Bayes²⁵ to optimize the marginal likelihood $\log p_\theta(x_n | l_n, c_n)$. We use a mean-field Gaussian variational distribution $q_\phi(\gamma_n | x_n, c_n)$, parameterized by a two-layer neural network g . This neural network takes as input the concatenation of the gene expression vector x_n as well as the one-hot encoding of the cell type label and outputs the mean and the variance of the variational distribution for γ_n . We optimize the evidence lower bound:

$$\begin{aligned} \log p_\theta(x_n | l_n, c_n) &\geq E_{q_\phi(\gamma_n | x_n, c_n)} \log p_\theta(x_n | \gamma_n, l_n, c_n) \\ &\quad - KL(q_\phi(\gamma_n | x_n, c_n) ; p_\theta(\gamma_n)), \end{aligned}$$

where $p_\theta(\gamma_n)$ denotes the prior likelihood for γ_n . For fitting this model, we sub-sample the observations in mini-batches, and we sample from the variational distribution using the re-parameterization trick. Additionally, we re-weight cells by their inverse cell type proportion (capped with a minimal proportion of 5%). We have found this to be an effective method for learning a better representation of the lowly abundant cell types (for example, monocytes in the lymph node).

Model of spatial transcriptomics data. *Assumption for the spatial data.* In the spatial data, we assume that the gene expression of each spot is the sum of contributions from multiple cells, each potentially being from different cell types. A standard modeling assumption is that a spot s has for expression x_s , the sum of individual cells¹³. Similarly, let us assume that each spot has $C(s)$ cells and that each cell n in spot s is generated from latent variables (c_{ns}, γ_{ns}) . We then have a distribution of gene expression:

$$x_{nsg} \sim \text{NegativeBinomial}\left(l_s \alpha_g f^g(c_{ns}, \gamma_{ns}), p_g\right),$$

with l_s is a spot-specific scaling factor and α_g is a correction term for the difference in experimental assays. From the standard property of the rate-shape parameterization of the negative binomial distribution, the distribution of the total gene expression x_{sg} in spot s and gene g is:

$$x_{sg} \sim \text{NegativeBinomial}\left(l_s \alpha_g \sum_{n=1}^{C(s)} f^g(c_{ns}, \gamma_{ns}), p_g\right).$$

We now assume that all the cells from a given cell type c in a given spot s must all be generated from the same covariate γ_s^c . Instead of determining the cell identity of all individual cells in the spot, we focus on determining the density into broad cell types as well as the archetype of the sub-cell state, which is a simpler problem. More concretely, we assume that there cannot be both significantly different cell states of the same cell types within a radius of 50 μm (that is, a spot).

Generative model. With these points in mind, we parameterize the sum in the previous equation to be over cell types. We obtain the following generative process:

$$x_{sg} \sim \text{NegativeBinomial}\left(l_s \alpha_g \sum_{c=1}^C \beta_{sc} f^g(c, \gamma_s^c), p_g\right),$$

where f and p denote, respectively, the decoder network and the rate parameter of the negative binomial, transferred from scLVM. The gene-specific multiplicative factor α explicitly controls for discrepancies between the technologies. Parameters β_{sc} are positive and designate the abundance of every cell type in every spot. These parameters may be normalized per spot to return an estimate of the cell type proportion. In our implementation, we also add a constant term that serves as an unknown cell type, as in ref. ¹³.

An important technical component of DestVI is the empirical prior that we use for the per-spot-per-cell type γ_s^c latent variable. Indeed, the model is susceptible to factor technology discrepancies into the latent space γ instead of the multiplicative factor α if an informative prior is not used, and we noticed this pathological behavior with an isotropic normal prior. Consequently, we designed an empirical prior based on the single-cell data for each cell type c :

$$\gamma_s^c \sim \frac{1}{K} \sum_{k=1}^K q_\phi(\gamma^c | u_{kc}, c),$$

where $\{u_{kc}\}_{k=1}^K$ designates a set of cells from cell type c in the scRNA-seq dataset, and q_ϕ designates the variational distribution from the scLVM. In another context, this prior over γ^c is referred to as a variational aggregated mixture of posteriors (VampPrior²⁶). However, the objective here is simply to use the points $\{u_{kc}\}_{k=1}^K$ as a more informative prior for deconvolution; the VampPrior method seeks to learn a multi-modal, more complex prior to better fit the data.

MAP inference. We infer point estimates for random variables γ^c and for parameters α and β using a penalized likelihood method. In addition to vanilla MAP inference, we introduced two key ideas that stabilized the performance of DestVI. First, we added to the likelihood a variance penalty for the parameter α , calculated across all the genes. This strategy was applied previously by ZINB-WaVE to regularize estimates of dispersion parameters in their likelihood-based matrix factorization of single-cell data⁵⁴. Second, compared to a standard deconvolution model that has exactly C parameters per spot, scLVM has C parameters and dC latent variables per spot, where d denotes the dimension of the latent space learned by scLVM. To avoid overfitting, we, therefore, proposed to use a neural network to parameterize the latent variables as a function of the input data (as in auto-encoding variational Bayes). Namely, we proposed several variants of the algorithm in which both, part or of none of β and γ may be parameterized by neural networks. Intuitively, the use of a neural network for inference of γ^c may be helpful whenever there are shared transcriptomics profiles across cell types (such as inflammatory signals). In addition, we optionally incorporated a sparsity prior for the cell type abundance per spot β_{sc} for settings where we expect the presence of very few cell types per spot. This penalty is enforced for the high-resolution liver data analysis as well as the Sci-Space data. Because we may not easily perform classical cross-validation in our inference scenario, we manually increased the regularization strength by a constant factor until the abundance was mostly zero for most cell types. With these points in mind, the objective function to train this generative model is simply composed of (1) the negative binomial likelihood, (2) the likelihood of the empirical prior, (3) the variance penalization for α and (4) (optionally) the sparsity prior for the cell type abundance.

Simulations and data generative process. The benchmarking of spatial deconvolution methods often relies on using an scRNA-seq dataset with cell type annotation and aggregating multiple cells into a pseudo-bulk using a ground truth proportion. The algorithms are then evaluated on the prediction of the cell type proportions¹³. This approach is not entirely adequate to our setting, as we would like to assess imputation of cell-type-specific variations of the transcriptome that are lost when combining together random cells of a given cluster. Instead, we used a real dataset of scRNA-seq data to simulate paired spatial and single-cell transcriptomics data. This helps properly benchmark our method against existing deconvolution methods.

Learning parameters from single-cell data. We propose to build a simulation process for generating single-cell data and learn its parameters from real data. Concretely, the goal is an algorithm for sampling gene expression counts x of a single cell. We let the distribution that governs these counts be a function of the cell's type c , as well as a low-dimensional embedding γ . The intuition behind this design choice is to have clear differences of transcriptome from one cell type to another but continuous variations of cellular states within a type. A similar approach was taken in the design of the SimSym simulator for scRNA-seq data⁵⁵.

As a model for the counts, we propose to map the low-dimensional embedding γ to the gene expression space using a negative binomial generalized linear model for every cell type c :

$$x | \gamma, c \sim \text{NegativeBinomial} \left(\exp \{ \mu_c + W_c^\top \gamma \} - 1, \theta \right),$$

where μ_c describes a cell-type-specific offset of gene expression, and W_c is a linear map. This is a simplification of SimSym that relies on a beta-Poisson likelihood model. However, the negative binomial has been shown to perform well in modeling scRNA-seq data^{23,56}. We learned the parameters μ_c , W_c and θ based on the lymph node scRNA-seq dataset of this manuscript. Out of all the cells, we kept five cell types: B cells, CD4 T cells, CD8 T cells, migratory DCs and Tregs. Those cell types were selected because they were the most abundant in the dataset.

Although approximate inference of these parameters is, in principle, possible—for example using maximum likelihood as in ZINB-WaVE⁵⁴—we wanted to enforce sparsity on the gene-loading matrix W_c motivated by the idea that only a subset of the genes contributes to the variation of gene expression within a unique cell type. Also, sparsity seemed necessary because, for some cell types, the number of cells was much lower than the number of genes (for example, around $n = 300$ cells for the regulatory T cells and $G = 2,000$ genes). Consequently, we learned patterns of transcriptomic variation within each cell type using a sparse principal component analysis (PCA) model on log-normalized data⁵⁷. The counts were normalized using scanpy⁵⁸ with a target count of 10,000 UMIs, and the sparse PCA model was fit using sklearn⁵⁹, using four components and a Lagrange multiplier of 5 for the $\|\cdot\|_1$ penalty. The output of this procedure is a cell-type-specific embedding for every cell γ_s , a mean expression profile for every cell type μ_c and a dictionary of within-cell-type transcriptomic variation W_c . The dispersion parameter θ_g for every gene g is estimated from the data using scVI²⁴.

Generating spatial maps of transcriptome. To build the spatial transcriptomics data, we first constructed a regular grid of dimension 40×40 . Each spot s (that is, point on the grid) is associated with spatial coordinates t_s . For the cell type proportions π_s , we built a covariance matrix K based on the spatial coordinates:

$$[K]_{ss'} = \exp \left\{ -\frac{\|t_s - t_{s'}\|_2^2}{\lambda\sigma^2} \right\},$$

where σ denotes the median distance between all pairs of points in the grid. The λ parameter controls the level of spatial smoothness of the stochastic process and is fixed to 0.1. This kernel matrix K may be used to sample c independent draws from a Gaussian process:

$$\phi^c \sim \text{Normal}(0, K),$$

and interpret them as an energy to derive cell type proportions at every spot:

$$\pi_s^c = \frac{\exp \left\{ \frac{\phi^c}{T} \right\}}{\sum_c \exp \left\{ \frac{\phi^c}{T} \right\}},$$

where T is a temperature parameter, set to 1. Large T would correspond to all proportions to be equal to 20%, whereas small T would tend to make the proportions binary. Regarding the embedding variables for every cell type γ_s^c , we treat them as 4C independent draw from the same kernel.

For every spot s , we sample a fixed number of cells K . For every single cell, we decide on its cell type c based on a draw from the categorical distribution parameterized by the proportions π_s . We then use the previously introduced simulation model to generate the transcriptome of that cell (of cell type c and embedding γ_s^c). In our simulation, the scRNA-seq dataset has $K = 20$ cells per spot. Finally, for every spot, we average the mean parameter of the negative binomial distribution across all K cells, and we sample from the same observation model.

Comparison to competing methods. Our major claim is that DestVI is able to infer cell type proportions but also to detect within-cell-type transcriptomic variations in the spatial transcriptomics data. Although cell type proportion estimation methods are reasonably simple to benchmark based on simulations, there is more ambiguity with respect to the second task. We provide a robust evaluation of the performance of algorithms at identifying continuous cell states within each cell type, based on cell-type-specific gene expression imputation.

We, therefore, provided a slight modification of every algorithm so that it may be used to impute gene expression at a given cell type c and a given spot s . For example, DestVI directly performs this task by accessing the inferred variable γ_s^c and decoding as $f(c, \gamma_s^c)$, which serves as an unnormalized gene expression. For algorithms based on embeddings and nearest neighbors imputation (scVI, Harmony and Scanorama), we impute by calculating the average gene expression of the k-nearest neighbors of the embedding of the spot s restricted to the single-cell data of cell type c only. Similarly, we estimate the proportion using the empirical proportion of cell types for the k-nearest neighbors of the embeddings of the spot s restricted to all the single-cell data. Finally, we also evaluate the performance of discrete deconvolution algorithms (Stereoscope, Seurat, RCTD and SPOTLight) for within-cell-type gene expression. For this, we re-cluster the single-cell data for each cell type at several depths, using hierarchical clustering (two, four and eight clusters per cell type). Then, we run the discrete deconvolution methods on the

re-clustered data. We also calculate the mean gene expression for every cluster based on the single-cell data. Finally, we impute gene expression for cell type c at spot s by averaging the gene expression of every cluster in cell type c , weighted by the conditional proportions of every cluster at the spot. Given these modifications, we evaluated the imputation based on correlation metrics (for example, Spearman correlation) across an oracle list of genes for each cell type. The list of genes for each cell type is given by the indices of the non-zero coefficients of the matrix W_c , learned via sparse PCA. We have found that this selection of genes helps make the evaluation more robust. SPOTLight did not terminate after 3 hours for the most granular clustering (eight clusters per cell type).

Pre-processing for the sci-Space dataset. We obtained the data for the 14 sagittal sections derived from two embryonic day (E) 14 (E14.0) mouse embryos (C57BL/6N) from the original publication. We pre-processed the data as per the cortex-specific analysis conducted in Fig. 5 of the sci-Space manuscript. More precisely, we filtered for sections 8, 9, 11, 13 and 14 and kept only the cells whose anatomical annotation was cortex. Within those cells, we filtered for the top four abundant cell types (8,186 neurons, 1,610 radial glias, 581 connective tissue progenitors and 507 endothelial cells) included in the original study. We additionally selected the top 4,000 highly variable genes using scanpy. We assembled the spatial dataset by summing the transcriptome of all cells that share the same spatial barcode. Additionally, we increased the complexity of the dataset in two ways. First, we lowered the sequencing depth by performing binomial subsampling at rate 0.1. Second, we enriched the data by adding spots generated in silico from the individual cells of each spot but with different proportions. In particular, we enriched for cell types other than neurons by a factor of 5, as the original dataset is dominated by spots with mostly neurons. For evaluation, we applied the benchmarking pipeline previously used with the simulations. The main difference is that it relies on a subset of genes with zonation patterns (in the simulation, this is given by the sparsity pattern of the loading matrix from the PCA). For this, we took the intersection of the highly variable gene list and the list of genes highlighted in Fig. 5 of the sci-Space paper.

Automated pipeline for data exploration. The resulting model enables several types of downstream analyses for drawing hypotheses on both the spatial structure of an individual sample as well as the differences among pairs of conditions. In the following, we propose a standard pipeline for analysis of a single sample. This pipeline helped gain insight from a simple sample (a single lymph node or tumor section) and helped guide comparative analysis between samples (between lymph node sections or distinct areas of a tumor section). The pipeline consists of two parts. In the first part, we consider the data from a resolution of cell types. For each cell type, we report its proportion in every spot and then highlight cell types that tend to occur at specific niches (that is, not uniformly distributed across the tissue), using Geary's C autocorrelation statistic for the inferred proportions⁶⁰, as reported by Hotspot⁶¹ on the cell type proportion for each individual cell type.

The second part of the pipeline facilitates a more in-depth view, looking at variability within cell types, thus going beyond the functionality that is available in current pipelines. We start by selecting, for each type, spots that have a sufficiently high proportion of cells of that type.

Selecting informative thresholds for cell type proportions. We propose an automated way of estimating a cell-type-specific threshold for this procedure, but it may be also manually curated. Therefore, we chose a characteristic point of the autocorrelation curve, as a function of the thresholding. More precisely, for every cell type c , we apply the hard-thresholding operator T to the proportion vector β^c at different levels t_i (noted $T(\beta^c, t_i)$). We take t_i to be the empirical percentiles of the proportion across all spots. For each thresholded vector $T(\beta^c, t_i)$, we then calculate the autocorrelation metric from Hotspot⁶¹ for all of those thresholded proportions. The result is a curve that may be interpolated using splines using the scipy.interpolate function from SciPy⁶². We then analytically differentiate the spline and look for an inflection point (null second derivative).

Finding main axes of variation in the combined data. With this constrained view, our pipeline proceeds to report the main axes of variation in each cell type, using the spatial data. This analysis helps highlight and visualize the most dominant transcriptional programs in every cell type, as well as exploring their dependence on the cells' locations. The major question we would like to solve here is the following: within a single cell type, which directions of γ vary spatially? In other words, we want to find within-cell-type and spatially varying transcriptomic programs.

General matrix factorization methods (such as NMF, PCA or CCA) could be applied to find those components. However, they present several crucial limitations: (1) the samples are not re-weighted by the cell type proportion—this is relevant because γ_c is inferred at every spot, even when β^c is null; and (2) we are only interested in variations that are relevant with respect to the spatial coordinates λ .

To this end, we developed a weighted PCA scheme that uses the inferred cell states (in the spatial data) and accounts for the inferred cell type proportions as well as the spatial layout of the spots (following previous work on robust linear dimensionality reduction;⁶³ Supplementary Note 9).

Interpretation of the weighted PCA. For interpreting the spatially weighted PCA for every cell type, we first project the cell-type-specific cell state of the spatial data as well as the scRNA-seq data onto the two weighted principal components. We then use a two-dimensional (2D) color map to assign a color to the spots in the spatial data, according to the position of the (2D) projection of the cell-type-specific cell state. This is helpful to prioritize which cell type may have spatially consistent variations in cell states. To relate this to the functions of each axis of the principal components, we filter genes that are the most correlated with each Pearson correlation, and we ran gene set enrichment analysis using EnrichR⁴⁴ for the top 50 genes.

A post hoc recipe for differential expression with DestVI. Our model also provides a natural way to estimate and evaluate the significance of differences between conditions or between niches in the same tissue section. On the gene level, we can identify cell-type-specific DE, comparing different conditions or different tissue areas. This analysis draws directly from our probabilistic representation of the data, which allows for uncertainty quantification and hypothesis testing. Indeed, performing DE with a probabilistic model is always challenging but is crucial for making robust scientific discoveries. In our previous work, we used a purely Bayesian approach to DE⁴⁶. Because we apply MAP inference (and not fully Bayesian inference) to DestVI, we instead developed a hybrid approach to DE, where we sample from the adequate generative distribution and then derive a *P* value with a frequentist test. More precisely, for two bags of spots $(x_a)_{a \in A}$ and $(x_b)_{b \in B}$, and a cell type *c*, we operate as follows.

Characterizing cell-type-specific gene expression. We query the parameters of the generative distribution for that spot but only for the contribution from cell type *c*. We do this by embedding all the spots and keeping only the latent variable for the cell type of interest γ_s^c and then querying the decoder network $f(\gamma_s^c, c)$ for this particular γ_s^c and cell type *c*. This vector, along with p_g , fully characterize the transcriptome of a fictitious cell x_c of type *c* in spot *s*:

$$x_{sc} \sim \text{NegativeBinomial} \left(\hat{l} f(\gamma_s^c, c), p_g \right),$$

where \hat{l} is a fixed scalar equal to the average library size in the single-cell data.

Simulating cell-type-specific gene expression. Then, we sample multiple times per spot from the mean of the Poisson distribution underlying the negative binomial:

$$w_{sc} \sim \text{Gamma} \left(\hat{l} f(\gamma_s^c, c), p_g \right).$$

There are several considerations for sampling this way. First, as we underlined in our previous work for calculating gene correlations in totalVI, using the mean of the generative distribution provides all forms of biases. Second, sampling from the full negative binomial distribution introduces Poisson noise and reduces sensitivity of the method. Sampling from the gamma distribution is an intermediate, although post hoc, solution. This step plays the role of denoising and is especially important when there are very few spots in the bags *A* or *B*, because we can generate more data points.

Hypothesis testing. For every spot in every bag, we generate between $p=1$ and $p=180$ independent samples of the random variables w_{sc} and apply a two-sample Kolmogorov–Smirnov test, and we correct for multiple testing using the Benjamini–Hochberg procedure⁶⁵. Additionally, we tag genes as differentially expressed if two conditions are met: (1) the null hypothesis is rejected, after control of the FDR at level 0.05, and (2) the log-fold change (LFC) is greater than a data-driven threshold in absolute value. To calculate this threshold, we assume that a significant amount of LFCs will concentrate around zero (which correspond to the genes that are equally expressed), whereas DE genes will concentrate around other modes. Based on this assumption, we fit a three-component Gaussian mixture model to the LFC and keep the mean of the mode with largest absolute value, whose associated distributions should contain DE genes.

Extending simulations for assessing robustness to cell type mismatch. We added two additional cell types to our simulations whose parameter W_c was learned from our lymph node scRNA-seq data using sPCA (as per previous simulations). To simulate a cell type missing from the single-cell data, we simply filtered all the cells from this cell type out post hoc. To simulate a cell type missing from the spatial data, we cancelled all the contributions of that cell type to the transcriptome, inside of every spot.

Data pre-processing for the Seq-Scope liver analysis. We relied on publicly available scRNA-seq data from Tabula Muris Senis⁶⁶, filtering for the droplet-based liver tissue data. We annotated the data manually, based on scanpy and marker genes (Mup3 for hepatocytes, Clec4f for macrophages and Kdr for endothelial cells). We imported the spatial transcriptomics data at 10-μm resolution from the Seq-Scope paper¹⁹ and annotated at the spot level by Seurat.

Extending simulations for high-resolution data. To simulate high-resolution data, we have relied on the simulations with increased number of cell types (7) and lowered the temperature parameter *T* from 1 to 0.5 (lowering from 5–6 to 2–3 the empirical number of cell types per spot). We have also decreased the library size of the spatial transcriptomics data to be 50% of the library size of the scRNA-seq data by binomial downsampling.

Statistics and reproducibility. A unique section of a new lymph node has been used for all imaging (Fig. 3i,m and Supplementary Figs. 16 and 21). A unique section of the same tumor has been used for Figs. 4f and 5c and Supplementary Figs. 24–27. The result in Fig. 5e has been tested on 12 independent tumor sections.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw data discussed in this manuscript have been deposited in the National Center for Biotechnology Information's Gene Expression Omnibus under accession number GSE173778 (murine lymph node and tumor; spatial transcriptomics and scRNA-seq data). Processed sequencing data are available on our reproducibility repository (<https://github.com/romain-lopez/DestVI-reproducibility>).

Code availability

The code to reproduce the results in this manuscript is available on the GitHub repository (<https://github.com/romain-lopez/DestVI-reproducibility>) and has been deposited to Zenodo (<https://doi.org/10.5281/zenodo.4685952>). The reference implementation of DestVI, along with accompanying tutorials, is available via the scvi-tools codebase at <https://scvi-tools.org/>.

References

51. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
52. Ioffe, S. & Szegedy, C. Batch normalization: accelerating deep network training by reducing internal covariate shift. International Conference on Machine Learning (2015).
53. Ba, J. L., Kiros, J. R. & Hinton, G. E. Layer normalization. Preprint at <https://arxiv.org/abs/1607.06450> (2016).
54. Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S. & Vert, J.-P. A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat. Commun.* **9**, 284 (2018).
55. Zhang, X., Xu, C. & Yosef, N. Simulating multiple faceted variability in single cell RNA sequencing. *Nat. Commun.* **10**, 2611 (2019).
56. Choudhary, S. & Satija, R. Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol.* **18**, 27 (2022).
57. Zou, H., Hastie, T. & Tibshirani, R. Sparse principal component analysis. *J. Comput. Graph. Stat.* **15**, 265–286 (2006).
58. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
59. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
60. Geary, R. C. The contiguity ratio and statistical mapping. *The Incorporated Statistician* **5**, 115–127 (1954).
61. DeTomaso, D. & Yosef, N. Hotspot identifies informative gene modules across modalities of single-cell genomics. *Cell Syst.* (2021).
62. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
63. Koren, Y. & Carmel, L. Robust linear dimensionality reduction. *IEEE Transactions on Visualization and Computer Graphics* **10**, 459–470 (2004).
64. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinf.* **14**, 128 (2013).
65. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **57**, 289–300 (1995).
66. Tabula Muris Consortium. A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature* **583**, 590–595 (2020).

Acknowledgements

We would like to acknowledge A. Gayoso, G. Xing and J. Hong for their help integrating DestVI in the scvi-tools codebase. Thanks to Z. Steier for providing guidance on the annotation of the lymph node single-cell data. We acknowledge S. Itzkovitz for guidance on interpreting the liver results. We thank E. Davidson for the artwork. We are grateful for insightful conversations with A. Regev, D. Pe'er, Q. Morris, A. Battle, A. Weiner, E. Rahmani and M. Jones. Funding: N.Y. and R.L. were supported by the Chan Zuckerberg Biohub, the Chan-Zuckerberg Foundation Network under grant number 2019–02452 (N.Y.) and the National Institute of Mental Health under grant

number U19MH114821 (N.Y.). I.A. is an Eden and Steven Romick Professorial Chair, supported by Merck, the Chan Zuckerberg Initiative, the Howard Hughes Medical Institute International Scholar Award, European Research Council Consolidator Grant 724471-HemTree2.0, an SCA award of the Wolfson Foundation and Family Charitable Trust, the Thompson Family Foundation, a Melanoma Research Alliance Established Investigator Award (509044), the Israel Science Foundation (703/15), the Ernest and Bonnie Beutler Research Program for Excellence in Genomic Medicine, the Helen and Martin Kimmel award for innovative investigation, the NeuroMac DFG/Transregional Collaborative Research Center Grant, International Progressive MS Alliance/NMSS PA-1604 08459, the ISF Israel Precision Medicine Program (IPMP) 607/20 grant and an Adelis Foundation grant.

Author contributions

R.L., B.L., H.K.-S., I.A. and N.Y. designed the study and the experiments. B.L. performed the experimental procedures, with input from I.Y., O.G., F.R. and H.K.-S. M.K. and D.P prepared Visium and scRNA-seq libraries. A.J. processed scRNA-seq of the tumor data. B.L. and Y.A. contributed to microscopy analysis. E.D. assisted with RNA sequencing data processing and data upload. R.L. conceived the statistical model with input from

B.L., H.K.-S. and M.I.J. R.L implemented the DestVI software and applied the software to analyze the data, with input from A.W. and C.E. P.B. proposed the spatially aware extension of DestVI. I.A. and N.Y. supervised the work.

Competing interests

N.Y. is an advisor and/or has equity in Cellarity, Celsius Therapeutics and Rheos Medicine. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41587-022-01272-8>.

Correspondence and requests for materials should be addressed to Ido Amit or Nir Yosef.

Peer review information *Nature Biotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Software used for data collection are outlined in the manuscript. For flow cytometry BD Bioscience FACS DIVA (v8), for imaging Leica DMI8 wide-field inverted microscope with Leica LAS X, Andor Dragonfly spinning disk confocal on an inverted Leica DMI8 microscope with Leica Fusion, Zeiss LSM880 with Zen Blue software. Sequencing data was collected using Illumina NovaSeq Control Software v1.6 for NovaSeq6000 system.

Data analysis

We used Cell Ranger 4.0 and Space Ranger software v1.2 for processing the RNA data. For flow cytometry FlowJo v10.7.1; for imaging Leica LAS X, Imaris X64 v9 (Bitplane), QuPath 0.3, Fiji (NIH Image, Bethesda, MD).

Code developed as part of the pipeline is publicly available at the Github repository <https://github.com/romain-lopez/DestVI-reproducibility> and has been deposited to Zenodo <https://doi.org/10.5281/zenodo.4685952>. We used the following python packages:
`anndata==0.7.5, harmonypy==0.0.5, numba==0.52.0, scikit-learn==0.24.1, scipy==1.6.0, pytorch-lightning==1.2.6, scanorama==1.7, scanpy==1.7.0, scrublet==0.2.3, torch==1.7.1, umap-learn==0.4.6`

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The raw data discussed in this manuscript have been deposited in the National Center for Biotechnology Information's Gene Expression Omnibus and are accessible through accession number GSE173778 (murine lymph node and tumor; spatial transcriptomics and scRNA-seq data). The processed data are available on our reproducibility GitHub repository (<https://github.com/romain-lopez/DestVI-reproducibility>).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

No statistical tool was used to apriori choose sample size. In this study, we use two models, one is MS treated LN, one is MCA205 tumor. These two models are studied a lot in our previous studies (reference A and B). So here we use 2 control LN sections and 2 MS treated LN tissue sections, and 2 MCA205 tumor sections for Visium processing. We use 3 mouse LN for verified staining and 12 mouse MCA205 tumor for Hypoxia verified staining.

[A] Ronnie Blecher-Gonen. Single-Cell Analysis of Diverse Pathogen Responses Defines a Molecular Roadmap for Generating Antigen-Specific Immunity. *Cell Systems*. 2019;8: 109–121.e6.
 [B] Katzenelenbogen Y, Sheban F, Yalin A, Yofe I, Svetlichnyy D, Jaitin DA, et al. Coupled scRNA-Seq and Intracellular Protein Activity Reveal an Immunosuppressive Role of TREM2 in Cancer. *Cell*. 2020;182: 872–885.e19.

Data exclusions

Described in detail in Methods. Exclusion criteria were determined following initial assessment and QC of the data.

Replication

For Lymph node experiment, LNs from 2 MS treated mice and 2 PBS treated mice are for single cell RNA sequencing. 2 sections from Ms LN and 2 sections from PBS LN are for Visium processing. 3 mouse LN for verified staining inflammatory monocytes distribution is consistent through replication samples.
 For tumor experiment, 2 tumors from 2 mice are for single cell RNA sequencing. 2 sections from 1 tumor are for Visium processing.
 For hypoxia test, 1 h before sacrificing, Hypoxyprobe™-1 solution was injected at a dosage of 60 mg/kg. 12 Flash frozen samples are for further testing. The hypoxia distribution in MCA205 solid tumor is consistent through replication samples. For MCA205 staining for Visium quantification, we take 3 different sections from the same tissue block for Visium with different distance from the Section 2 which is for Visium processing, described details in Figure 4.

Randomization

The animal are randomly allocated into MS or PBS group. The LN sections for Visium processing are taken from the center line of the LNs. The MCA205 tumor section for Visium are taken from a proper size tumor to match Visium capture frame size 6.5mm x 6.5mm and make sure the boundary area is included in the section for Visium processing. The images for hypoxia test distribution are randomly sampling from the whole tissue section.

Blinding

LNs are blindly collected for single cell RNA sequencing. The analysis of the LN single cell RNA sequencing is blind. We recognized different LNs based on our single cell data and also based on our previous study. LNs for Visium are not blind because we want to put MS LN and PBS LN together on the same OCT block and sampled on the same Visium capture area.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

CD45 APC (1:100, 30-F11, eBioscience, 17-0451-82), CD11b FITC (1:200, M1/70, Biolegend, 101205), CD11b Biotin (1:100, M1/70, Biolegend, 101204), CD11b PE (1:100, M1/70, eBioscience, 12-0112-83), CD64 PE (1:100, X54-5/7.1, Biolegend, 139303), Ly6C FITC (1:100, HK1.4, Biolegend, 128005), B220 (1:100, RA3-6B2, Biolegend, 103208), CD3 Biotin (1:100, 17A2, Biolegend, 100243), MHCII I-A/I-E FITC (1:100, M5/114.15.2, Biolegend, 107606), F4/80 APC (1:100, BM8, eBioscience, 17-4801-82), NK1.1 PE (1:100, PK136, eBioscience, 12-5941-63), CD31 APC (1:100, MEC13.3, Biolegend, 102509), IFIT3 polyclonal antibody (1:500, Proteintech, 15201-1-AP), Arginase 1 (ARG1, 1:200, Novus, NB100-59740), pimonidazole antibody (Hypoxyprobe, 1:200, PAb2627AP), GPNMB (1:200, Abcam, ab234529), F4/80 (1:100, Abcam, ab6640), DAPI (1 µg/mL, ThermoFisher, D1306), Streptavidin APC (1:400, Biolegend, 405207), DyLight 550 conjugated donkey anti-goat IgG heavy and light chain antibody (1:400, Bethyl, A50-101D3), Cy2 conjugated AffiniPure donkey anti-rabbit IgG heavy and light chain antibody (1:200, Jackson ImmunoResearch, 711-225-152), DyLight 594 conjugated donkey anti-rat IgG heavy and light chain cross-adsorbed antibody (1:400, Bethyl, A110-337D4), Alexa Fluor 647 conjugated AffiniPure F(ab')2 fragment donkey anti-rabbit IgG heavy and light chain antibody (1:400, Jackson ImmunoResearch, 711-606-152), Goat anti-Rabbit IgG-heavy and light chain Antibody DyLight® 650 Conjugated (1:800, Bethyl, A120-101D5).

Validation

All antibodies used are commercially available. Validation of antibodies was provided by the manufacturer and their quality control data can be accessed through vendor's website. Further in our experiment, we also used mouse Fc-block in FACS experiments, and blocking solution in immuno-staining experiments. Moreover, we validated that antibodies used in the study were specific for unique cell types, and didn't exhibit any non-specific binding. For multiplex staining, we also conducted single antibody staining to check the specific binding in each antibody.

Eukaryotic cell linesPolicy information about [cell lines](#)

Cell line source(s)

MCA205 fibrosarcoma cell line

Authentication

We used the cell-line for in tumor model studies. 5 × 10⁵ MCA205 tumor cells were suspended in 100µl PBS and injected subcutaneous (s.c.) into 8-week female mice.

Mycoplasma contamination

Cell line was negative for Mycoplasma.

Commonly misidentified lines
(See [ICLAC](#) register)

N/A

Animals and other organismsPolicy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

C57BL/6J WT female mice were used in the study.

Wild animals

Study did not involve wild animals

Field-collected samples

Study did not involve samples collected in the field.

Ethics oversight

All animals were handled according to the regulations formulated by the Weizmann Institutional Animal Care and Use Committee approved from Weizmann Institute of Science.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

To achieve single-cell suspensions in the murine model, tumors were cut into small pieces, and suspended with RPMI-1640 supplemented with DNase I (0.1mg/ml, Sigma-Adrich) and collagenase IV (1mg/ml, Worthington). Tissues were homogenized by GentleMacs tissue homogenizer (Miltenyi Biotec), and incubated at 37°C for 10 min, with frequent agitation. This tissue dissociation procedure was performed twice, for each tumor.
To achieve single-cell suspensions from LN, tissues were digested in Isocoves modified Dulbeccos medium (IMDM; Sigma-Aldrich) medium. For mild dissociation, tissues were supplemented with Liberase-TL (100 µg/ml, Roche) and DNase I (100 µg/ml, Roche), and incubated with frequent agitation at 37°C for 20 min.
After all dissociation procedures, cells were washed with cold PBS, filtered through a 70 µm cell strainer, and centrifuged at 350g for 5 min at 4°C.

Instrument

Cell populations were sorted using either ARIA-III instrument (BD Biosciences), and analyzed using BD FACSDIVA software (BD Biosciences) and FlowJo software.

Software

Cell populations were sorted using ARIA-III instrument (BD Biosciences), and analyzed using BD FACS DIVA software (BD Biosciences) and FlowJo software.

Cell population abundance

Singe cells from LNs were loaded onto Next GEM Chip G at concentration around 1000/µL and yielded 14,989 cells after quality filtering and described in Figure 3C. Single cells from MCA205 tumor were enriched in CD45+ gate and their cell abundances were calculated and reported in Figures 4C.

Gating strategy

mouse: FSC-A/SSC-A -> FSC-H/FSC-W-> CD45+-> CD45+DAPI-

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.