

SOLAR: Fixed orthogonal label anchors for single-cell reference mapping

Additional file 1: Supplementary methods, figures and tables

Minqi Chen^{1,2}, Guanjie Zou^{4,5}, Minrui Chen³,
Hiroto Saigo^{3,5}, Koji Tsuda^{4,5}, Lunxu Liu^{1,2}

Contents

1	Supplementary study details	1
2	Supplementary methods	1
2.1	Reference–query construction and preprocessing	1
2.2	SOLAR optimization and controls	2
2.3	Comparator implementations	2
2.4	Metrics, populations, and aggregation	3
3	Supplementary tables	4
4	Supplementary figures	27

1 Supplementary study details

This Additional file 1 provides the reference–query protocols, implementation details, held-out results, output-width analysis, and anchor controls. All Tables S1–S9 and Figures S1–S5 cited here are part of this file. The machine-readable results are available from the SOLAR repository.

2 Supplementary methods

2.1 Reference–query construction and preprocessing

The primary evaluation uses held-out batches: one complete batch forms the query and all remaining batches form the reference. Three batches per real dataset were selected to span query sizes. Every evaluated query label occurs in the corresponding reference (Table S1). Seeds 40–44 index repeated stochastic fits on the same held-out split, not new biological replicates. A supplementary output-width analysis uses batch–label-stratified within-batch partitions, with 80% of each stratum assigned to the reference and 20% to the query for split seeds 40–44; a one-cell stratum remains in the reference.

Feature selection uses the outer reference before internal training/validation partitioning. For Supervised Orthogonal Label-Anchor Representation (SOLAR) and matched principal component analysis (Matched PCA), principal component analysis (PCA) is likewise fitted exclusively on the outer reference. Supervised contrastive learning on cells and their types (SCLSC) [1] instead operates on the normalized 2,000-gene expression matrix selected with Scanpy [2]. Query profiles use the corresponding fixed preprocessing parameters.

The common comparison uses 30-dimensional outputs for SOLAR, Matched PCA, single-cell variational inference (scVI), its annotation-aware extension scANVI, and SCLSC. The method-specific summaries use 128-dimensional outputs for SOLAR, 40-dimensional Matched PCA, 30-dimensional scVI and scANVI, and 16-dimensional SCLSC. Both SOLAR configurations consume 40 reference-fitted principal components. The same 30-dimensional scVI and scANVI representations appear in both reporting blocks. Tables S3 and S7 distinguish the roles of the configurations.

After reference and query roles are assigned, query labels are used only for evaluation and are excluded from feature selection, PCA fitting, optimization, internal validation, checkpoint selection, and mapping. Query expression profiles enter the fixed mapping operations only after training and model selection are complete. Cached full-input PCA50 is used only as the post hoc evaluation source for trustworthiness; it is not the Matched PCA baseline or the input used to train SOLAR.

2.2 SOLAR optimization and controls

The expression encoder contains two 512-unit hidden layers with batch normalization, ReLU, and dropout of 0.1, followed by a linear projection, layer normalization, and ℓ_2 normalization. AdamW [3] uses learning rate 10^{-4} , batch size 512, temperature 0.07, a same-label cosine offset of 0.2 subtracted before temperature scaling, gradient-norm limit 1.0, and zero weight decay. Training stops after 10 epochs without a validation-loss improvement of 10^{-4} , or after 50 epochs, and restores the best validation-loss checkpoint.

Repeated anchors contribute one indexed copy of the fixed class anchor per cell in a mini-batch. Unique anchors contribute one anchor per represented class; the no-anchor control uses the same encoder and supervised contrastive objective without anchor views. The supplementary within-batch output-width analysis compares separately trained 30-, 64-, and 128-dimensional SOLAR models (Table S4). The held-out configuration comparison in Table S8 also uses separately trained models. The available records do not establish identical internal-validation subsets across these configurations, so the observed differences combine output width with any differences introduced by training the models separately. Table S9 reports the 128-dimensional anchor comparisons across evaluated metrics and cell populations.

Algorithm 1 Reference training and direct query mapping with SOLAR

Require: Labeled reference $\{(x_i, y_i)\}_{i \in \mathcal{R}}$; query profiles $\{x_j\}_{j \in \mathcal{Q}}$; temperature τ ; margin m

Ensure: Trained encoder and reference/query embeddings

- 1: Sample $G \in \mathbb{R}^{d \times K}$ and factorize $G = UT$ with $U^\top U = I_K$
 - 2: Use column u_k of U as the fixed anchor for label k
 - 3: Divide the reference into optimization and validation subsets
 - 4: **for** each epoch **do**
 - 5: **for** each reference mini-batch **do**
 - 6: Encode and normalize expression views z_i
 - 7: Add indexed anchors u_{y_i} (repeated) or one per represented label (unique); add none in the no-anchor control
 - 8: Set $V = 2B$, $B + K_B$, or B for repeated, unique, or no anchors
 - 9: Form $C(a) = \{1, \dots, V\} \setminus \{a\}$ and $\text{Pos}(a) = \{j \in C(a) : \hat{y}_j = \hat{y}_a\}$
 - 10: Set an empty-positive focal term to zero; retain all V terms in the mean
 - 11: Evaluate the loss with cosine offset m and temperature τ
 - 12: Update encoder parameters; keep U fixed
 - 13: **end for**
 - 14: Evaluate validation loss and retain an improved checkpoint
 - 15: **if** validation loss has not improved for 10 epochs **then**
 - 16: **break**
 - 17: **end if**
 - 18: **end for**
 - 19: Restore the best encoder and encode all reference cells
 - 20: Encode query cells with the same preprocessing and encoder
 - 21: Keep all preprocessing and encoder parameters fixed during query mapping
 - 22: **return** reference and query embeddings
-

Because candidate and positive sets are indexed, duplicate-valued repeated anchors remain distinct views in the objective. The loss denominator contains all non-self views, whereas the numerator is averaged over same-label positive indices. The factor τ/τ_{base} is one because both temperatures are 0.07. In the no-anchor control, an empty-positive focal row contributes zero to the all-row mean but can still serve as a negative for other rows. Thus, this rule neither removes singleton cells nor implies zero total gradients for them.

2.3 Comparator implementations

SCLSC [1] uses the official multilayer perceptron and instance-type contrastive objective, with a 16-dimensional output in the method-specific summaries and a separately trained 30-dimensional output in the common comparison.

Class representatives are arithmetic means of expression vectors from the internal training subset. These expression-space means pass through the trainable encoder, so their positions change during training; they are not fixed orthogonal directions. The objective combines same-class squared Euclidean distances with different-class squared Euclidean hinge penalties, rather than SOLAR’s all-view temperature-scaled cosine softmax.

The SCLSC adapter uses a reference-only stratified 90:10 internal split, five fits per held-out batch, and early-stopping patience of 10. It retains the final stopped epoch, whereas SOLAR restores the best validation-loss checkpoint. Because reliable raw counts were unavailable across the three processed inputs, the adapter uses their normalized expression matrix and the shared reference-selected 2,000 genes. This adapts the publication’s input pipeline and does not reproduce its count filtering, analytic-Pearson-residual feature selection, normalization, and logarithmic transformation. The common readout fits a distance-weighted 15-nearest-neighbour classifier on all reference embeddings.

The scVI and scANVI comparators are trained on reference cells and use frozen 30-dimensional encoders for query mapping; no query-adaptive fine-tuning is performed. Table S2 summarizes all five methods in the comparison.

2.4 Metrics, populations, and aggregation

The evaluation distinguishes query-only annotation recovery, integration metrics on the joint reference and query population, and local-structure endpoints for specific populations. The annotation readout fits a classifier to all reference embeddings using the 15 nearest neighbours with distance weights and Euclidean distance (Minkowski $p = 2$). For the set C_Q of labels observed in the query truth, balanced accuracy is the mean class recall and macro-F1 is the mean of class-specific F1 scores over C_Q . Both range from 0 to 1 and are higher for better label recovery. Predictions of reference-only labels remain errors for the corresponding true query class.

Joint integration metrics. The evaluation uses a pinned single-cell integration benchmarking (scIB) backend at commit `e2a37e0ed63dc34b60aa535cc656400552af757a` [4]. Label average silhouette width (label ASW) is the cell-weighted mean silhouette of biological labels in the joint representation, transformed from s to $(s + 1)/2$; its range is $[0, 1]$ and larger values indicate stronger label separation. Batch ASW instead computes $1 - |s_i|$ for batch-label silhouettes within each eligible cell type, averages within type, and then averages type means. It ranges from 0 to 1, with higher values indicating less within-type batch segregation; it is not a raw metadata silhouette.

The integration local inverse Simpson’s index (iLISI) is the median graph-based local inverse Simpson diversity minus 1, not a score normalized to $[0, 1]$. The nominal range is $[0, B_{\text{joint}} - 1]$ for regular valid neighbourhoods and B_{joint} batch identities; larger values indicate greater local batch diversity. The pinned implementation may retain a negative sentinel for a degenerate neighbourhood, so this interval is not an unconditional bound. The evaluation requests 50% graph-LISI subsampling. The C++ implementation initializes that sampling from a system-generated random seed, so repeating the metric calculation may produce small numerical differences.

Local structure and pseudotime. Trustworthiness [5] measures whether neighbours in a target representation were also nearby in the source space. We use the scikit-learn rank-based score with Euclidean distance and $k = 15$; the range is $[0, 1]$, with larger values indicating stronger measured neighbourhood preservation. The source is the cached 50-component PCA representation of the processed scIB input. It is a post hoc evaluation source, not the reference-fitted Matched PCA target and not an input to SOLAR training or checkpoint selection. The held-out evaluation freezes at most 6,000 cell IDs per dataset, batch, and reference/query/joint population, using sampling seed 40 after sorting IDs. Every method, output dimension, and fitting seed uses that same sample. The endpoint is evaluated for immune and lung datasets; the pancreas rows are explicitly not applicable.

Immune-cell pseudotime neighbourhood smoothness is the Spearman correlation between the existing diffusion-pseudotime annotation and the mean annotation of 15 neighbours in the target embedding. Cells with NaN annotations are removed before 16 neighbours are requested; the first returned neighbour is discarded before the mean is taken. This neighbour is treated as the focal cell, so tied embedding vectors can make the procedure differ from explicit index-based self-exclusion. The range is $[-1, 1]$; larger values indicate stronger local agreement with the supplied annotation, not proof of a correct trajectory. The processed Figshare input supplies `dpt_pseudotime`; this analysis does not select a separate diffusion-pseudotime root, and the root identity is not recorded in the available metadata. A population with fewer than 16 annotated cells or an undefined correlation is NA. This endpoint is separate from scIB trajectory conservation: that metric chooses the annotated type with the lowest mean source pseudotime, restricts its root candidates to the largest target-graph component, chooses an extreme candidate by majority vote over the first three diffusion components, and runs target diffusion pseudotime. It reports $(\rho + 1)/2$, where ρ is the source–target pseudotime Spearman correlation, giving a $[0, 1]$ score when defined. In particular, failure to find an eligible root is not replaced with a pseudotime-smoothness value.

Raw metadata association. Within-label metadata silhouettes use study and sample ID for immune cells, and donor and patient group for lung cells. For each biological label with at least 10 cells and a mathematically valid multi-group silhouette, the raw silhouette is computed and then averaged across labels with cell-count weights. Each label is capped at 8,000 cells. The seed-0 capped cell selection is fixed for each dataset/batch/population/metadata field and shared across methods and fitting seeds. These raw scores range from -1 to 1 and describe residual covariate association; there is no uniformly desirable direction for all clinical or technical covariates. A single metadata group does not define a silhouette and is reported as NA, not zero.

Repeated fits, missingness, and paired summaries. For each fixed held-out batch, the tables require five finite, comparison-eligible scores (seeds 40–44) for a stochastic method and one for Matched PCA. A batch mean is withheld if that requirement is not met. Table S3 reports fitting-seed standard deviations (SDs) with divisor $5 - 1$; deterministic PCA rows have no seed SD. Across batches, summaries first average fits within a batch, then weight batches equally within a dataset and datasets equally. With complete three-dataset coverage, this equals the mean of nine batch means. SD across batch means describes between-batch variation, not seed variation or a standard error. Repeated fits on the same held-out batch are not independent biological replicates.

Paired differences use only batches for which both methods have complete, compatible batch means. Tables S7–S9 report applicable batch denominators, contributing fits, and reasons for unavailable values. Missing scores and out-of-scope populations are not replaced with zero. No significance or equivalence test is inferred from a descriptive SD or a signed difference.

3 Supplementary tables

Table S1: Datasets and held-out-batch composition.

Dataset	Query batch	Reference cells	Query cells	Ref. labels	Query labels	Coverage
Immune	10X	22,779	10,727	16	12	1.0
Immune	Oetjen_A	30,920	2,586	16	16	1.0
Immune	Villani	32,484	1,022	16	4	1.0
Lung	3	32,010	462	17	10	1.0
Lung	4	28,709	3,763	17	12	1.0
Lung	B1	30,332	2,140	17	11	1.0
Pancreas	fluidigmcl	15,744	638	14	13	1.0
Pancreas	inDrop2	14,658	1,724	14	14	1.0
Pancreas	inDrop3	12,777	3,605	14	14	1.0
Simulation 1	Batch1	9,189	2,908	7	6	1.0
Simulation 1	Batch3	9,977	2,120	7	5	1.0
Simulation 1	Batch6	11,140	957	7	4	1.0
Simulation 2	Batch1Sub3	18,108	1,210	4	2	1.0
Simulation 2	Batch3Sub1	17,388	1,930	4	2	1.0
Simulation 2	Batch3Sub4	18,833	485	4	1	1.0

Coverage is the fraction of query labels represented in the corresponding reference.

Table S2: Comparator implementations and query-data access.

Method	Software/source	Dimension	Training scope	Fits	Query handling
SOLAR	Public SOLAR v1.2	128 / 30	Reference only	5	Trained encoder
Matched PCA	Reference PCA	40 / 30	Reference-fitted transform	1	Fixed transform
scVI	scvi-tools	30	Reference only	5	Trained encoder
scANVI	scvi-tools	30	Reference only	5	Trained encoder
SCLSC	Official architecture and loss	16 / 30	Reference only; internal 90:10 split	5	Trained encoder

Table S3: All 90 held-out annotation rows. A: method-specific output configurations; B: common 30-dimensional outputs. Each stochastic entry is mean $\pm$ fitting-seed SD over five fits on one fixed batch; Matched PCA is deterministic (seed SD $-$). n is the number of compatible fits. The same 30-dimensional scVI and scANVI representations are reported in both blocks.

Block	Dataset / query batch	Method (output)	n	Macro-F1	Balanced accuracy
A	Immune / 10X	SOLAR (128D)	5	0.905 $\pm$ 0.004	0.894 $\pm$ 0.002
A	Immune / 10X	Matched PCA (40D)	1	0.869 (SD $-$)	0.852 (SD $-$)
A	Immune / 10X	scVI (30D)	5	0.857 $\pm$ 0.019	0.862 $\pm$ 0.016
A	Immune / 10X	scANVI (30D)	5	0.895 $\pm$ 0.007	0.891 $\pm$ 0.009
A	Immune / 10X	SCLSC (16D)	5	0.889 $\pm$ 0.006	0.876 $\pm$ 0.008
A	Immune / Oetjen_A	SOLAR (128D)	5	0.852 $\pm$ 0.003	0.876 $\pm$ 0.004
A	Immune / Oetjen_A	Matched PCA (40D)	1	0.878 (SD $-$)	0.891 (SD $-$)
A	Immune / Oetjen_A	scVI (30D)	5	0.848 $\pm$ 0.004	0.884 $\pm$ 0.002
A	Immune / Oetjen_A	scANVI (30D)	5	0.849 $\pm$ 0.008	0.890 $\pm$ 0.006
A	Immune / Oetjen_A	SCLSC (16D)	5	0.841 $\pm$ 0.002	0.880 $\pm$ 0.004
A	Immune / Villani	SOLAR (128D)	5	0.549 $\pm$ 0.018	0.637 $\pm$ 0.012
A	Immune / Villani	Matched PCA (40D)	1	0.463 (SD $-$)	0.574 (SD $-$)
A	Immune / Villani	scVI (30D)	5	0.705 $\pm$ 0.032	0.762 $\pm$ 0.030
A	Immune / Villani	scANVI (30D)	5	0.684 $\pm$ 0.018	0.740 $\pm$ 0.016
A	Immune / Villani	SCLSC (16D)	5	0.665 $\pm$ 0.049	0.723 $\pm$ 0.041
A	Lung / 3	SOLAR (128D)	5	0.886 $\pm$ 0.007	0.893 $\pm$ 0.007
A	Lung / 3	Matched PCA (40D)	1	0.828 (SD $-$)	0.864 (SD $-$)
A	Lung / 3	scVI (30D)	5	0.792 $\pm$ 0.012	0.835 $\pm$ 0.012
A	Lung / 3	scANVI (30D)	5	0.844 $\pm$ 0.020	0.864 $\pm$ 0.017
A	Lung / 3	SCLSC (16D)	5	0.843 $\pm$ 0.019	0.863 $\pm$ 0.016
A	Lung / 4	SOLAR (128D)	5	0.754 $\pm$ 0.024	0.849 $\pm$ 0.021
A	Lung / 4	Matched PCA (40D)	1	0.746 (SD $-$)	0.856 (SD $-$)
A	Lung / 4	scVI (30D)	5	0.619 $\pm$ 0.012	0.793 $\pm$ 0.026
A	Lung / 4	scANVI (30D)	5	0.699 $\pm$ 0.028	0.864 $\pm$ 0.012
A	Lung / 4	SCLSC (16D)	5	0.729 $\pm$ 0.024	0.850 $\pm$ 0.021
A	Lung / B1	SOLAR (128D)	5	0.879 $\pm$ 0.014	0.881 $\pm$ 0.021
A	Lung / B1	Matched PCA (40D)	1	0.840 (SD $-$)	0.851 (SD $-$)
A	Lung / B1	scVI (30D)	5	0.810 $\pm$ 0.005	0.837 $\pm$ 0.011
A	Lung / B1	scANVI (30D)	5	0.856 $\pm$ 0.006	0.893 $\pm$ 0.008
A	Lung / B1	SCLSC (16D)	5	0.844 $\pm$ 0.004	0.865 $\pm$ 0.007
A	Pancreas / fluidigm1	SOLAR (128D)	5	0.866 $\pm$ 0.045	0.845 $\pm$ 0.041
A	Pancreas / fluidigm1	Matched PCA (40D)	1	0.792 (SD $-$)	0.758 (SD $-$)
A	Pancreas / fluidigm1	scVI (30D)	5	0.863 $\pm$ 0.058	0.902 $\pm$ 0.037
A	Pancreas / fluidigm1	scANVI (30D)	5	0.894 $\pm$ 0.027	0.921 $\pm$ 0.018
A	Pancreas / fluidigm1	SCLSC (16D)	5	0.858 $\pm$ 0.040	0.843 $\pm$ 0.022
A	Pancreas / inDrop2	SOLAR (128D)	5	0.812 $\pm$ 0.022	0.838 $\pm$ 0.029
A	Pancreas / inDrop2	Matched PCA (40D)	1	0.805 (SD $-$)	0.816 (SD $-$)
A	Pancreas / inDrop2	scVI (30D)	5	0.863 $\pm$ 0.032	0.910 $\pm$ 0.036
A	Pancreas / inDrop2	scANVI (30D)	5	0.903 $\pm$ 0.026	0.948 $\pm$ 0.035
A	Pancreas / inDrop2	SCLSC (16D)	5	0.874 $\pm$ 0.030	0.918 $\pm$ 0.027
A	Pancreas / inDrop3	SOLAR (128D)	5	0.817 $\pm$ 0.034	0.827 $\pm$ 0.042

Table S3 (continued)

Block	Dataset / query batch	Method (output)	n	Macro-F1	Balanced accuracy
A	Pancreas / inDrop3	Matched PCA (40D)	1	0.799 (SD -)	0.848 (SD -)
A	Pancreas / inDrop3	scVI (30D)	5	0.907 $\pm$ 0.067	0.908 $\pm$ 0.071
A	Pancreas / inDrop3	scANVI (30D)	5	0.920 $\pm$ 0.063	0.931 $\pm$ 0.065
A	Pancreas / inDrop3	SCLSC (16D)	5	0.843 $\pm$ 0.042	0.855 $\pm$ 0.048
B	Immune / 10X	SOLAR (30D)	5	0.904 $\pm$ 0.007	0.892 $\pm$ 0.007
B	Immune / 10X	Matched PCA (30D)	1	0.862 (SD -)	0.841 (SD -)
B	Immune / 10X	scVI (30D)	5	0.857 $\pm$ 0.019	0.862 $\pm$ 0.016
B	Immune / 10X	scANVI (30D)	5	0.895 $\pm$ 0.007	0.891 $\pm$ 0.009
B	Immune / 10X	SCLSC (30D)	5	0.886 $\pm$ 0.019	0.870 $\pm$ 0.024
B	Immune / Oetjen_A	SOLAR (30D)	5	0.849 $\pm$ 0.003	0.874 $\pm$ 0.005
B	Immune / Oetjen_A	Matched PCA (30D)	1	0.874 (SD -)	0.882 (SD -)
B	Immune / Oetjen_A	scVI (30D)	5	0.848 $\pm$ 0.004	0.884 $\pm$ 0.002
B	Immune / Oetjen_A	scANVI (30D)	5	0.849 $\pm$ 0.008	0.890 $\pm$ 0.006
B	Immune / Oetjen_A	SCLSC (30D)	5	0.846 $\pm$ 0.012	0.885 $\pm$ 0.009
B	Immune / Villani	SOLAR (30D)	5	0.533 $\pm$ 0.023	0.626 $\pm$ 0.014
B	Immune / Villani	Matched PCA (30D)	1	0.454 (SD -)	0.567 (SD -)
B	Immune / Villani	scVI (30D)	5	0.705 $\pm$ 0.032	0.762 $\pm$ 0.030
B	Immune / Villani	scANVI (30D)	5	0.684 $\pm$ 0.018	0.740 $\pm$ 0.016
B	Immune / Villani	SCLSC (30D)	5	0.657 $\pm$ 0.043	0.716 $\pm$ 0.034
B	Lung / 3	SOLAR (30D)	5	0.881 $\pm$ 0.018	0.891 $\pm$ 0.017
B	Lung / 3	Matched PCA (30D)	1	0.828 (SD -)	0.857 (SD -)
B	Lung / 3	scVI (30D)	5	0.792 $\pm$ 0.012	0.835 $\pm$ 0.012
B	Lung / 3	scANVI (30D)	5	0.844 $\pm$ 0.020	0.864 $\pm$ 0.017
B	Lung / 3	SCLSC (30D)	5	0.839 $\pm$ 0.027	0.867 $\pm$ 0.014
B	Lung / 4	SOLAR (30D)	5	0.741 $\pm$ 0.027	0.846 $\pm$ 0.025
B	Lung / 4	Matched PCA (30D)	1	0.749 (SD -)	0.856 (SD -)
B	Lung / 4	scVI (30D)	5	0.619 $\pm$ 0.012	0.793 $\pm$ 0.026
B	Lung / 4	scANVI (30D)	5	0.699 $\pm$ 0.028	0.864 $\pm$ 0.012
B	Lung / 4	SCLSC (30D)	5	0.705 $\pm$ 0.035	0.827 $\pm$ 0.029
B	Lung / B1	SOLAR (30D)	5	0.877 $\pm$ 0.009	0.876 $\pm$ 0.017
B	Lung / B1	Matched PCA (30D)	1	0.836 (SD -)	0.841 (SD -)
B	Lung / B1	scVI (30D)	5	0.810 $\pm$ 0.005	0.837 $\pm$ 0.011
B	Lung / B1	scANVI (30D)	5	0.856 $\pm$ 0.006	0.893 $\pm$ 0.008
B	Lung / B1	SCLSC (30D)	5	0.841 $\pm$ 0.010	0.866 $\pm$ 0.021
B	Pancreas / fluidigmcl	SOLAR (30D)	5	0.865 $\pm$ 0.034	0.847 $\pm$ 0.034
B	Pancreas / fluidigmcl	Matched PCA (30D)	1	0.797 (SD -)	0.763 (SD -)
B	Pancreas / fluidigmcl	scVI (30D)	5	0.863 $\pm$ 0.058	0.902 $\pm$ 0.037
B	Pancreas / fluidigmcl	scANVI (30D)	5	0.894 $\pm$ 0.027	0.921 $\pm$ 0.018
B	Pancreas / fluidigmcl	SCLSC (30D)	5	0.850 $\pm$ 0.028	0.832 $\pm$ 0.023
B	Pancreas / inDrop2	SOLAR (30D)	5	0.822 $\pm$ 0.034	0.844 $\pm$ 0.037
B	Pancreas / inDrop2	Matched PCA (30D)	1	0.787 (SD -)	0.789 (SD -)
B	Pancreas / inDrop2	scVI (30D)	5	0.863 $\pm$ 0.032	0.910 $\pm$ 0.036
B	Pancreas / inDrop2	scANVI (30D)	5	0.903 $\pm$ 0.026	0.948 $\pm$ 0.035
B	Pancreas / inDrop2	SCLSC (30D)	5	0.860 $\pm$ 0.024	0.901 $\pm$ 0.015
B	Pancreas / inDrop3	SOLAR (30D)	5	0.827 $\pm$ 0.025	0.841 $\pm$ 0.031

Table S3 (continued)

Block	Dataset / query batch	Method (output)	n	Macro-F1	Balanced accuracy
B	Pancreas / inDrop3	Matched PCA (30D)	1	0.779 (SD –)	0.833 (SD –)
B	Pancreas / inDrop3	scVI (30D)	5	0.907 ± 0.067	0.908 ± 0.071
B	Pancreas / inDrop3	scANVI (30D)	5	0.920 ± 0.063	0.931 ± 0.065
B	Pancreas / inDrop3	SCLSC (30D)	5	0.849 ± 0.054	0.846 ± 0.053

Table S4: Representation-width sensitivity for separately trained within-batch SOLAR models. Values average five split seeds for each fitted output configuration.

Dataset	Dim.	label ASW	NMI	ARI	iLISI	PCR	Trajectory
Immune	30	0.854	0.766	0.395	1.714	0.729	0.643
Immune	64	0.851	0.756	0.377	1.522	0.721	0.731
Immune	128	0.847	0.742	0.347	1.392	0.722	0.712
Lung	30	0.852	0.813	0.544	1.375	0.673	–
Lung	64	0.850	0.806	0.523	1.156	0.670	–
Lung	128	0.855	0.804	0.512	1.065	0.670	–
Pancreas	30	0.919	0.753	0.399	2.143	0.927	–
Pancreas	64	0.919	0.739	0.365	1.736	0.925	–
Pancreas	128	0.921	0.738	0.353	1.384	0.925	–
Simulation 1	30	0.962	0.782	0.478	2.092	0.933	–
Simulation 1	64	0.960	0.778	0.482	1.743	0.932	–
Simulation 1	128	0.958	0.777	0.491	1.413	0.933	–
Simulation 2	30	0.982	0.687	0.389	1.940	0.437	–
Simulation 2	64	0.980	0.628	0.290	1.024	0.435	–
Simulation 2	128	0.979	0.613	0.264	0.100	0.439	–

NMI, normalized mutual information; ARI, adjusted Rand index; PCR, principal component regression. A dash indicates that the endpoint is outside the dataset scope.

Table S5: Complete held-out local-structure panel, including method-specific outputs, common 30-dimensional outputs, and 128-dimensional controls. A: method-specific configurations; B: common 30-dimensional outputs. Values are equally weighted means of compatible batch means; brackets give valid/applicable batches. Higher trustworthiness or pseudotime smoothness means stronger preservation of the measured relation. P: fewer than 16 annotated query cells; only immune Oetjen_A contributes query smoothness.

Dataset / measure	Block / target	Reference	Query	Joint	NA
Immune / Trustworthiness	A / SOLAR (128D)	0.911 [3/3]	0.948 [3/3]	0.918 [3/3]	–
Immune / Trustworthiness	A / Matched PCA (40D)	0.982 [3/3]	0.973 [3/3]	0.985 [3/3]	–
Immune / Trustworthiness	A / scVI (30D)	0.882 [3/3]	0.949 [3/3]	0.889 [3/3]	–
Immune / Trustworthiness	A / scANVI (30D)	0.888 [3/3]	0.951 [3/3]	0.895 [3/3]	–
Immune / Trustworthiness	A / SCLSC (16D)	0.873 [3/3]	0.913 [3/3]	0.885 [3/3]	–
Immune / Trustworthiness	B / SOLAR (30D)	0.886 [3/3]	0.943 [3/3]	0.895 [3/3]	–
Immune / Trustworthiness	B / Matched PCA (30D)	0.982 [3/3]	0.974 [3/3]	0.985 [3/3]	–
Immune / Trustworthiness	B / scVI (30D)	0.882 [3/3]	0.949 [3/3]	0.889 [3/3]	–
Immune / Trustworthiness	B / scANVI (30D)	0.888 [3/3]	0.951 [3/3]	0.895 [3/3]	–
Immune / Trustworthiness	B / SCLSC (30D)	0.873 [3/3]	0.914 [3/3]	0.885 [3/3]	–
Immune / Trustworthiness	Control / SOLAR unique (128D)	0.936 [3/3]	0.952 [3/3]	0.942 [3/3]	–
Immune / Trustworthiness	Control / SOLAR no anchor (128D)	0.932 [3/3]	0.950 [3/3]	0.939 [3/3]	–
Lung / Trustworthiness	A / SOLAR (128D)	0.945 [3/3]	0.933 [3/3]	0.947 [3/3]	–
Lung / Trustworthiness	A / Matched PCA (40D)	0.994 [3/3]	0.966 [3/3]	0.994 [3/3]	–
Lung / Trustworthiness	A / scVI (30D)	0.885 [3/3]	0.944 [3/3]	0.887 [3/3]	–
Lung / Trustworthiness	A / scANVI (30D)	0.925 [3/3]	0.947 [3/3]	0.925 [3/3]	–
Lung / Trustworthiness	A / SCLSC (16D)	0.921 [3/3]	0.892 [3/3]	0.921 [3/3]	–
Lung / Trustworthiness	B / SOLAR (30D)	0.928 [3/3]	0.930 [3/3]	0.930 [3/3]	–
Lung / Trustworthiness	B / Matched PCA (30D)	0.994 [3/3]	0.964 [3/3]	0.994 [3/3]	–
Lung / Trustworthiness	B / scVI (30D)	0.885 [3/3]	0.944 [3/3]	0.887 [3/3]	–
Lung / Trustworthiness	B / scANVI (30D)	0.925 [3/3]	0.947 [3/3]	0.925 [3/3]	–
Lung / Trustworthiness	B / SCLSC (30D)	0.923 [3/3]	0.893 [3/3]	0.923 [3/3]	–
Lung / Trustworthiness	Control / SOLAR unique (128D)	0.958 [3/3]	0.942 [3/3]	0.959 [3/3]	–
Lung / Trustworthiness	Control / SOLAR no anchor (128D)	0.955 [3/3]	0.942 [3/3]	0.956 [3/3]	–
Immune / Pseudotime smoothness	A / SOLAR (128D)	0.852 [3/3]	0.949 [1/3]	0.873 [3/3]	P
Immune / Pseudotime smoothness	A / Matched PCA (40D)	0.876 [3/3]	0.977 [1/3]	0.895 [3/3]	P
Immune / Pseudotime smoothness	A / scVI (30D)	0.864 [3/3]	0.969 [1/3]	0.880 [3/3]	P
Immune / Pseudotime smoothness	A / scANVI (30D)	0.868 [3/3]	0.970 [1/3]	0.886 [3/3]	P
Immune / Pseudotime smoothness	A / SCLSC (16D)	0.792 [3/3]	0.930 [1/3]	0.815 [3/3]	P
Immune / Pseudotime smoothness	B / SOLAR (30D)	0.837 [3/3]	0.947 [1/3]	0.861 [3/3]	P
Immune / Pseudotime smoothness	B / Matched PCA (30D)	0.876 [3/3]	0.977 [1/3]	0.894 [3/3]	P
Immune / Pseudotime smoothness	B / scVI (30D)	0.864 [3/3]	0.969 [1/3]	0.880 [3/3]	P
Immune / Pseudotime smoothness	B / scANVI (30D)	0.868 [3/3]	0.970 [1/3]	0.886 [3/3]	P
Immune / Pseudotime smoothness	B / SCLSC (30D)	0.794 [3/3]	0.934 [1/3]	0.818 [3/3]	P
Immune / Pseudotime smoothness	Control / SOLAR unique (128D)	0.857 [3/3]	0.953 [1/3]	0.878 [3/3]	P
Immune / Pseudotime smoothness	Control / SOLAR no anchor (128D)	0.854 [3/3]	0.953 [1/3]	0.875 [3/3]	P

Table S6: Complete held-out raw metadata-silhouette panel. Values are equally weighted means across compatible batch means; brackets give valid/applicable batches for that population. A denotes method-specific configurations, B denotes common 30-dimensional outputs, and controls are 128-dimensional. Raw covariate association has no universal higher-is-better direction. Reasons: C, no valid within-label multi-group silhouette.

Dataset / measure	Block / target	Reference	Query	Joint	NA
Immune / study silhouette	A / SOLAR (128D)	0.001 [3/3]	NA [0/3]	0.007 [3/3]	C
Immune / study silhouette	A / Matched PCA (40D)	0.191 [3/3]	NA [0/3]	0.202 [3/3]	C
Immune / study silhouette	A / scVI (30D)	-0.001 [3/3]	NA [0/3]	0.009 [3/3]	C
Immune / study silhouette	A / scANVI (30D)	-0.007 [3/3]	NA [0/3]	-0.003 [3/3]	C
Immune / study silhouette	A / SCLSC (16D)	-0.001 [3/3]	NA [0/3]	-0.018 [3/3]	C
Immune / study silhouette	B / SOLAR (30D)	-0.008 [3/3]	NA [0/3]	-0.002 [3/3]	C
Immune / study silhouette	B / Matched PCA (30D)	0.213 [3/3]	NA [0/3]	0.222 [3/3]	C
Immune / study silhouette	B / scVI (30D)	-0.001 [3/3]	NA [0/3]	0.009 [3/3]	C
Immune / study silhouette	B / scANVI (30D)	-0.007 [3/3]	NA [0/3]	-0.003 [3/3]	C
Immune / study silhouette	B / SCLSC (30D)	0.005 [3/3]	NA [0/3]	-0.011 [3/3]	C
Immune / study silhouette	Control / SOLAR unique (128D)	0.026 [3/3]	NA [0/3]	0.037 [3/3]	C
Immune / study silhouette	Control / SOLAR no anchor (128D)	0.022 [3/3]	NA [0/3]	0.032 [3/3]	C
Immune / sample-ID silhouette	A / SOLAR (128D)	0.001 [3/3]	NA [0/3]	0.007 [3/3]	C
Immune / sample-ID silhouette	A / Matched PCA (40D)	0.191 [3/3]	NA [0/3]	0.202 [3/3]	C
Immune / sample-ID silhouette	A / scVI (30D)	-0.001 [3/3]	NA [0/3]	0.009 [3/3]	C
Immune / sample-ID silhouette	A / scANVI (30D)	-0.007 [3/3]	NA [0/3]	-0.003 [3/3]	C
Immune / sample-ID silhouette	A / SCLSC (16D)	-0.001 [3/3]	NA [0/3]	-0.018 [3/3]	C
Immune / sample-ID silhouette	B / SOLAR (30D)	-0.008 [3/3]	NA [0/3]	-0.002 [3/3]	C
Immune / sample-ID silhouette	B / Matched PCA (30D)	0.213 [3/3]	NA [0/3]	0.222 [3/3]	C
Immune / sample-ID silhouette	B / scVI (30D)	-0.001 [3/3]	NA [0/3]	0.009 [3/3]	C
Immune / sample-ID silhouette	B / scANVI (30D)	-0.007 [3/3]	NA [0/3]	-0.003 [3/3]	C
Immune / sample-ID silhouette	B / SCLSC (30D)	0.005 [3/3]	NA [0/3]	-0.011 [3/3]	C
Immune / sample-ID silhouette	Control / SOLAR unique (128D)	0.026 [3/3]	NA [0/3]	0.037 [3/3]	C
Immune / sample-ID silhouette	Control / SOLAR no anchor (128D)	0.022 [3/3]	NA [0/3]	0.032 [3/3]	C
Lung / donor silhouette	A / SOLAR (128D)	-0.089 [3/3]	NA [0/3]	-0.095 [3/3]	C
Lung / donor silhouette	A / Matched PCA (40D)	0.069 [3/3]	NA [0/3]	0.054 [3/3]	C
Lung / donor silhouette	A / scVI (30D)	-0.050 [3/3]	NA [0/3]	-0.052 [3/3]	C
Lung / donor silhouette	A / scANVI (30D)	-0.057 [3/3]	NA [0/3]	-0.060 [3/3]	C
Lung / donor silhouette	A / SCLSC (16D)	-0.233 [3/3]	NA [0/3]	-0.246 [3/3]	C
Lung / donor silhouette	B / SOLAR (30D)	-0.101 [3/3]	NA [0/3]	-0.108 [3/3]	C
Lung / donor silhouette	B / Matched PCA (30D)	0.069 [3/3]	NA [0/3]	0.054 [3/3]	C
Lung / donor silhouette	B / scVI (30D)	-0.050 [3/3]	NA [0/3]	-0.052 [3/3]	C
Lung / donor silhouette	B / scANVI (30D)	-0.057 [3/3]	NA [0/3]	-0.060 [3/3]	C
Lung / donor silhouette	B / SCLSC (30D)	-0.220 [3/3]	NA [0/3]	-0.232 [3/3]	C
Lung / donor silhouette	Control / SOLAR unique (128D)	-0.041 [3/3]	NA [0/3]	-0.047 [3/3]	C
Lung / donor silhouette	Control / SOLAR no anchor (128D)	-0.049 [3/3]	NA [0/3]	-0.055 [3/3]	C
Lung / patient-group silhouette	A / SOLAR (128D)	0.250 [3/3]	NA [0/3]	0.236 [3/3]	C
Lung / patient-group silhouette	A / Matched PCA (40D)	0.308 [3/3]	NA [0/3]	0.309 [3/3]	C
Lung / patient-group silhouette	A / scVI (30D)	0.097 [3/3]	NA [0/3]	0.093 [3/3]	C
Lung / patient-group silhouette	A / scANVI (30D)	0.111 [3/3]	NA [0/3]	0.106 [3/3]	C
Lung / patient-group silhouette	A / SCLSC (16D)	0.264 [3/3]	NA [0/3]	0.234 [3/3]	C

Table S6 (continued)

Dataset / measure	Block / target	Reference	Query	Joint	NA
Lung / patient-group silhouette	B / SOLAR (30D)	0.243 [3/3]	NA [0/3]	0.229 [3/3]	C
Lung / patient-group silhouette	B / Matched PCA (30D)	0.324 [3/3]	NA [0/3]	0.325 [3/3]	C
Lung / patient-group silhouette	B / scVI (30D)	0.097 [3/3]	NA [0/3]	0.093 [3/3]	C
Lung / patient-group silhouette	B / scANVI (30D)	0.111 [3/3]	NA [0/3]	0.106 [3/3]	C
Lung / patient-group silhouette	B / SCLSC (30D)	0.261 [3/3]	NA [0/3]	0.231 [3/3]	C
Lung / patient-group silhouette	Control / SOLAR unique (128D)	0.232 [3/3]	NA [0/3]	0.219 [3/3]	C
Lung / patient-group silhouette	Control / SOLAR no anchor (128D)	0.236 [3/3]	NA [0/3]	0.224 [3/3]	C

Table S7: Full held-out cross-method results by metric and population. A: method-specific configurations; B: common 30-dimensional outputs. Each mean weights complete batch means within dataset, then datasets equally. SD describes the contributing batch means (not seed variation). D, B and F report valid/applicable datasets, complete batches, and contributing/expected fits. Codes are defined below; numerical and comparable seed coverage and exact reasons are in the companion CSV.

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
A	SOLAR 128D	ARI / J	0.404 $\pm$ 0.077	3/3	9/9	45/45	–
A	SOLAR 128D	label ASW / J	0.869 $\pm$ 0.036	3/3	9/9	45/45	–
A	SOLAR 128D	batch ASW / J	0.842 $\pm$ 0.020	3/3	9/9	45/45	–
A	SOLAR 128D	NMI / J	0.758 $\pm$ 0.034	3/3	9/9	45/45	–
A	SOLAR 128D	PCR batch / J	0.771 $\pm$ 0.118	3/3	9/9	45/45	–
A	SOLAR 128D	Balanced accuracy / Q	0.838 $\pm$ 0.079	3/3	9/9	45/45	–
A	SOLAR 128D	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
A	SOLAR 128D	Cell-cycle conservation / J	0.539 $\pm$ 0.179	3/3	9/9	45/45	–
A	SOLAR 128D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
A	SOLAR 128D	iLISI / J	1.268 $\pm$ 0.165	3/3	9/9	45/45	–
A	SOLAR 128D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
A	SOLAR 128D	Macro-F1 / Q	0.813 $\pm$ 0.109	3/3	9/9	45/45	–
A	SOLAR 128D	Pseudotime smoothness / J	0.873 $\pm$ 0.007	1/1	3/3	15/15	S
A	SOLAR 128D	Pseudotime smoothness / Q	0.949 (SD –)	1/1	1/3	5/15	P,S
A	SOLAR 128D	Pseudotime smoothness / R	0.852 $\pm$ 0.043	1/1	3/3	15/15	S
A	SOLAR 128D	Trajectory conservation / J	0.623 $\pm$ 0.095	1/1	3/3	15/15	T
A	SOLAR 128D	Trustworthiness / J	0.933 $\pm$ 0.018	2/2	6/6	30/30	S
A	SOLAR 128D	Trustworthiness / Q	0.941 $\pm$ 0.012	2/2	6/6	30/30	S
A	SOLAR 128D	Trustworthiness / R	0.928 $\pm$ 0.019	2/2	6/6	30/30	S
A	SOLAR 128D	Raw donor silhouette / J	-0.095 $\pm$ 0.003	1/1	3/3	15/15	S
A	SOLAR 128D	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SOLAR 128D	Raw donor silhouette / R	-0.089 $\pm$ 0.004	1/1	3/3	15/15	S
A	SOLAR 128D	Raw patient-group silhouette / J	0.236 $\pm$ 0.015	1/1	3/3	15/15	S
A	SOLAR 128D	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SOLAR 128D	Raw patient-group silhouette / R	0.250 $\pm$ 0.024	1/1	3/3	15/15	S
A	SOLAR 128D	Raw sample-ID silhouette / J	0.007 $\pm$ 0.023	1/1	3/3	15/15	S
A	SOLAR 128D	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SOLAR 128D	Raw sample-ID silhouette / R	0.001 $\pm$ 0.015	1/1	3/3	15/15	S
A	SOLAR 128D	Raw study silhouette / J	0.007 $\pm$ 0.023	1/1	3/3	15/15	S
A	SOLAR 128D	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SOLAR 128D	Raw study silhouette / R	0.001 $\pm$ 0.015	1/1	3/3	15/15	S
A	Matched PCA (40D)	ARI / J	0.279 $\pm$ 0.051	3/3	9/9	9/9	–
A	Matched PCA (40D)	label ASW / J	0.570 $\pm$ 0.020	3/3	9/9	9/9	–
A	Matched PCA (40D)	batch ASW / J	0.846 $\pm$ 0.022	3/3	9/9	9/9	–
A	Matched PCA (40D)	NMI / J	0.666 $\pm$ 0.029	3/3	9/9	9/9	–
A	Matched PCA (40D)	PCR batch / J	0.437 $\pm$ 0.297	3/3	9/9	9/9	–
A	Matched PCA (40D)	Balanced accuracy / Q	0.812 $\pm$ 0.097	3/3	9/9	9/9	–
A	Matched PCA (40D)	cLISI / J	0.987 $\pm$ 0.012	3/3	9/9	9/9	–
A	Matched PCA (40D)	Cell-cycle conservation / J	0.660 $\pm$ 0.228	3/3	9/9	9/9	–
A	Matched PCA (40D)	HVG overlap / J	NA (SD –)	0/3	0/9	0/9	H
A	Matched PCA (40D)	iLISI / J	0.297 $\pm$ 0.161	3/3	9/9	9/9	–

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
A	Matched PCA (40D)	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	9/9	–
A	Matched PCA (40D)	Macro-F1 / Q	0.780 $\pm$ 0.125	3/3	9/9	9/9	–
A	Matched PCA (40D)	Pseudotime smoothness / J	0.895 $\pm$ 0.002	1/1	3/3	3/3	S
A	Matched PCA (40D)	Pseudotime smoothness / Q	0.977 (SD –)	1/1	1/3	1/3	P,S
A	Matched PCA (40D)	Pseudotime smoothness / R	0.876 $\pm$ 0.034	1/1	3/3	3/3	S
A	Matched PCA (40D)	Trajectory conservation / J	0.871 $\pm$ 0.003	1/1	3/3	3/3	T
A	Matched PCA (40D)	Trustworthiness / J	0.989 $\pm$ 0.005	2/2	6/6	6/6	S
A	Matched PCA (40D)	Trustworthiness / Q	0.969 $\pm$ 0.008	2/2	6/6	6/6	S
A	Matched PCA (40D)	Trustworthiness / R	0.988 $\pm$ 0.007	2/2	6/6	6/6	S
A	Matched PCA (40D)	Raw donor silhouette / J	0.054 $\pm$ 0.007	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/3	C,S
A	Matched PCA (40D)	Raw donor silhouette / R	0.069 $\pm$ 0.015	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw patient-group silhouette / J	0.309 $\pm$ 0.002	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/3	C,S
A	Matched PCA (40D)	Raw patient-group silhouette / R	0.308 $\pm$ 0.006	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw sample-ID silhouette / J	0.202 $\pm$ 0.020	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/3	C,S
A	Matched PCA (40D)	Raw sample-ID silhouette / R	0.191 $\pm$ 0.031	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw study silhouette / J	0.202 $\pm$ 0.020	1/1	3/3	3/3	S
A	Matched PCA (40D)	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/3	C,S
A	Matched PCA (40D)	Raw study silhouette / R	0.191 $\pm$ 0.031	1/1	3/3	3/3	S
A	scVI 30D	ARI / J	0.353 $\pm$ 0.057	3/3	9/9	45/45	–
A	scVI 30D	label ASW / J	0.578 $\pm$ 0.031	3/3	9/9	45/45	–
A	scVI 30D	batch ASW / J	0.878 $\pm$ 0.020	3/3	9/9	45/45	–
A	scVI 30D	NMI / J	0.677 $\pm$ 0.009	3/3	9/9	45/45	–
A	scVI 30D	PCR batch / J	0.871 $\pm$ 0.069	3/3	9/9	45/45	–
A	scVI 30D	Balanced accuracy / Q	0.855 $\pm$ 0.053	3/3	9/9	45/45	–
A	scVI 30D	cLISI / J	0.911 $\pm$ 0.086	3/3	9/9	45/45	–
A	scVI 30D	Cell-cycle conservation / J	0.536 $\pm$ 0.072	3/3	9/9	45/45	–
A	scVI 30D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
A	scVI 30D	iLISI / J	2.125 $\pm$ 0.302	3/3	9/9	45/45	–
A	scVI 30D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
A	scVI 30D	Macro-F1 / Q	0.807 $\pm$ 0.091	3/3	9/9	45/45	–
A	scVI 30D	Pseudotime smoothness / J	0.880 $\pm$ 0.003	1/1	3/3	15/15	S
A	scVI 30D	Pseudotime smoothness / Q	0.969 (SD –)	1/1	1/3	5/15	P,S
A	scVI 30D	Pseudotime smoothness / R	0.864 $\pm$ 0.026	1/1	3/3	15/15	S
A	scVI 30D	Trajectory conservation / J	0.830 $\pm$ 0.027	1/1	3/3	15/15	T
A	scVI 30D	Trustworthiness / J	0.888 $\pm$ 0.009	2/2	6/6	30/30	S
A	scVI 30D	Trustworthiness / Q	0.946 $\pm$ 0.019	2/2	6/6	30/30	S
A	scVI 30D	Trustworthiness / R	0.884 $\pm$ 0.006	2/2	6/6	30/30	S
A	scVI 30D	Raw donor silhouette / J	-0.052 $\pm$ 0.004	1/1	3/3	15/15	S
A	scVI 30D	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	scVI 30D	Raw donor silhouette / R	-0.050 $\pm$ 0.003	1/1	3/3	15/15	S
A	scVI 30D	Raw patient-group silhouette / J	0.093 $\pm$ 0.002	1/1	3/3	15/15	S
A	scVI 30D	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
A	scVI 30D	Raw patient-group silhouette / R	0.097 ± 0.001	1/1	3/3	15/15	S
A	scVI 30D	Raw sample-ID silhouette / J	0.009 ± 0.022	1/1	3/3	15/15	S
A	scVI 30D	Raw sample-ID silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
A	scVI 30D	Raw sample-ID silhouette / R	-0.001 ± 0.009	1/1	3/3	15/15	S
A	scVI 30D	Raw study silhouette / J	0.009 ± 0.022	1/1	3/3	15/15	S
A	scVI 30D	Raw study silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
A	scVI 30D	Raw study silhouette / R	-0.001 ± 0.009	1/1	3/3	15/15	S
A	scANVI 30D	ARI / J	0.389 ± 0.082	3/3	9/9	45/45	-
A	scANVI 30D	label ASW / J	0.657 ± 0.018	3/3	9/9	45/45	-
A	scANVI 30D	batch ASW / J	0.867 ± 0.015	3/3	9/9	45/45	-
A	scANVI 30D	NMI / J	0.753 ± 0.046	3/3	9/9	45/45	-
A	scANVI 30D	PCR batch / J	0.836 ± 0.065	3/3	9/9	45/45	-
A	scANVI 30D	Balanced accuracy / Q	0.882 ± 0.061	3/3	9/9	45/45	-
A	scANVI 30D	cLISI / J	1.000 ± 0.000	3/3	9/9	45/45	-
A	scANVI 30D	Cell-cycle conservation / J	0.604 ± 0.159	3/3	9/9	45/45	-
A	scANVI 30D	HVG overlap / J	NA (SD -)	0/3	0/9	0/45	H
A	scANVI 30D	iLISI / J	1.902 ± 0.189	3/3	9/9	45/45	-
A	scANVI 30D	Isolated-label silhouette (source) / J	0.501 ± 0.010	3/3	9/9	45/45	-
A	scANVI 30D	Macro-F1 / Q	0.838 ± 0.087	3/3	9/9	45/45	-
A	scANVI 30D	Pseudotime smoothness / J	0.886 ± 0.001	1/1	3/3	15/15	S
A	scANVI 30D	Pseudotime smoothness / Q	0.970 (SD -)	1/1	1/3	5/15	P,S
A	scANVI 30D	Pseudotime smoothness / R	0.868 ± 0.030	1/1	3/3	15/15	S
A	scANVI 30D	Trajectory conservation / J	0.855 ± 0.003	1/1	3/3	15/15	T
A	scANVI 30D	Trustworthiness / J	0.910 ± 0.018	2/2	6/6	30/30	S
A	scANVI 30D	Trustworthiness / Q	0.949 ± 0.018	2/2	6/6	30/30	S
A	scANVI 30D	Trustworthiness / R	0.906 ± 0.021	2/2	6/6	30/30	S
A	scANVI 30D	Raw donor silhouette / J	-0.060 ± 0.003	1/1	3/3	15/15	S
A	scANVI 30D	Raw donor silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
A	scANVI 30D	Raw donor silhouette / R	-0.057 ± 0.002	1/1	3/3	15/15	S
A	scANVI 30D	Raw patient-group silhouette / J	0.106 ± 0.004	1/1	3/3	15/15	S
A	scANVI 30D	Raw patient-group silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
A	scANVI 30D	Raw patient-group silhouette / R	0.111 ± 0.008	1/1	3/3	15/15	S
A	scANVI 30D	Raw sample-ID silhouette / J	-0.003 ± 0.020	1/1	3/3	15/15	S
A	scANVI 30D	Raw sample-ID silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
A	scANVI 30D	Raw sample-ID silhouette / R	-0.007 ± 0.015	1/1	3/3	15/15	S
A	scANVI 30D	Raw study silhouette / J	-0.003 ± 0.020	1/1	3/3	15/15	S
A	scANVI 30D	Raw study silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
A	scANVI 30D	Raw study silhouette / R	-0.007 ± 0.015	1/1	3/3	15/15	S
A	SCLSC 16D	ARI / J	0.350 ± 0.111	3/3	9/9	45/45	-
A	SCLSC 16D	label ASW / J	0.950 ± 0.013	3/3	9/9	45/45	-
A	SCLSC 16D	batch ASW / J	0.739 ± 0.024	3/3	9/9	45/45	-
A	SCLSC 16D	NMI / J	0.756 ± 0.061	3/3	9/9	45/45	-
A	SCLSC 16D	PCR batch / J	0.787 ± 0.111	3/3	9/9	45/45	-
A	SCLSC 16D	Balanced accuracy / Q	0.853 ± 0.053	3/3	9/9	45/45	-
A	SCLSC 16D	cLISI / J	1.000 ± 0.000	3/3	9/9	45/45	-

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
A	SCLSC 16D	Cell-cycle conservation / J	0.501 $\pm$ 0.141	3/3	9/9	45/45	–
A	SCLSC 16D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
A	SCLSC 16D	iLISI / J	2.291 $\pm$ 0.545	3/3	9/9	45/45	–
A	SCLSC 16D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
A	SCLSC 16D	Macro-F1 / Q	0.821 $\pm$ 0.074	3/3	9/9	45/45	–
A	SCLSC 16D	Pseudotime smoothness / J	0.815 $\pm$ 0.002	1/1	3/3	15/15	S
A	SCLSC 16D	Pseudotime smoothness / Q	0.930 (SD –)	1/1	1/3	5/15	P,S
A	SCLSC 16D	Pseudotime smoothness / R	0.792 $\pm$ 0.039	1/1	3/3	15/15	S
A	SCLSC 16D	Trajectory conservation / J	0.524 $\pm$ 0.027	1/1	2/3	10/15	T
A	SCLSC 16D	Trustworthiness / J	0.903 $\pm$ 0.021	2/2	6/6	30/30	S
A	SCLSC 16D	Trustworthiness / Q	0.903 $\pm$ 0.024	2/2	6/6	30/30	S
A	SCLSC 16D	Trustworthiness / R	0.897 $\pm$ 0.027	2/2	6/6	30/30	S
A	SCLSC 16D	Raw donor silhouette / J	-0.246 $\pm$ 0.006	1/1	3/3	15/15	S
A	SCLSC 16D	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SCLSC 16D	Raw donor silhouette / R	-0.233 $\pm$ 0.005	1/1	3/3	15/15	S
A	SCLSC 16D	Raw patient-group silhouette / J	0.234 $\pm$ 0.054	1/1	3/3	15/15	S
A	SCLSC 16D	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SCLSC 16D	Raw patient-group silhouette / R	0.264 $\pm$ 0.061	1/1	3/3	15/15	S
A	SCLSC 16D	Raw sample-ID silhouette / J	-0.018 $\pm$ 0.025	1/1	3/3	15/15	S
A	SCLSC 16D	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SCLSC 16D	Raw sample-ID silhouette / R	-0.001 $\pm$ 0.004	1/1	3/3	15/15	S
A	SCLSC 16D	Raw study silhouette / J	-0.018 $\pm$ 0.025	1/1	3/3	15/15	S
A	SCLSC 16D	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
A	SCLSC 16D	Raw study silhouette / R	-0.001 $\pm$ 0.004	1/1	3/3	15/15	S
B	SOLAR 30D	ARI / J	0.441 $\pm$ 0.077	3/3	9/9	45/45	–
B	SOLAR 30D	label ASW / J	0.873 $\pm$ 0.036	3/3	9/9	45/45	–
B	SOLAR 30D	batch ASW / J	0.836 $\pm$ 0.020	3/3	9/9	45/45	–
B	SOLAR 30D	NMI / J	0.774 $\pm$ 0.032	3/3	9/9	45/45	–
B	SOLAR 30D	PCR batch / J	0.777 $\pm$ 0.115	3/3	9/9	45/45	–
B	SOLAR 30D	Balanced accuracy / Q	0.837 $\pm$ 0.082	3/3	9/9	45/45	–
B	SOLAR 30D	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
B	SOLAR 30D	Cell-cycle conservation / J	0.527 $\pm$ 0.171	3/3	9/9	45/45	–
B	SOLAR 30D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
B	SOLAR 30D	iLISI / J	1.740 $\pm$ 0.361	3/3	9/9	45/45	–
B	SOLAR 30D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
B	SOLAR 30D	Macro-F1 / Q	0.811 $\pm$ 0.114	3/3	9/9	45/45	–
B	SOLAR 30D	Pseudotime smoothness / J	0.861 $\pm$ 0.006	1/1	3/3	15/15	S
B	SOLAR 30D	Pseudotime smoothness / Q	0.947 (SD –)	1/1	1/3	5/15	P,S
B	SOLAR 30D	Pseudotime smoothness / R	0.837 $\pm$ 0.046	1/1	3/3	15/15	S
B	SOLAR 30D	Trajectory conservation / J	0.654 $\pm$ 0.075	1/1	3/3	15/15	T
B	SOLAR 30D	Trustworthiness / J	0.912 $\pm$ 0.022	2/2	6/6	30/30	S
B	SOLAR 30D	Trustworthiness / Q	0.936 $\pm$ 0.013	2/2	6/6	30/30	S
B	SOLAR 30D	Trustworthiness / R	0.907 $\pm$ 0.023	2/2	6/6	30/30	S
B	SOLAR 30D	Raw donor silhouette / J	-0.108 $\pm$ 0.007	1/1	3/3	15/15	S
B	SOLAR 30D	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
B	SOLAR 30D	Raw donor silhouette / R	-0.101 $\pm$ 0.007	1/1	3/3	15/15	S
B	SOLAR 30D	Raw patient-group silhouette / J	0.229 $\pm$ 0.017	1/1	3/3	15/15	S
B	SOLAR 30D	Raw patient-group silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SOLAR 30D	Raw patient-group silhouette / R	0.243 $\pm$ 0.025	1/1	3/3	15/15	S
B	SOLAR 30D	Raw sample-ID silhouette / J	-0.002 $\pm$ 0.027	1/1	3/3	15/15	S
B	SOLAR 30D	Raw sample-ID silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SOLAR 30D	Raw sample-ID silhouette / R	-0.008 $\pm$ 0.018	1/1	3/3	15/15	S
B	SOLAR 30D	Raw study silhouette / J	-0.002 $\pm$ 0.027	1/1	3/3	15/15	S
B	SOLAR 30D	Raw study silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SOLAR 30D	Raw study silhouette / R	-0.008 $\pm$ 0.018	1/1	3/3	15/15	S
B	Matched PCA (30D)	ARI / J	0.287 $\pm$ 0.058	3/3	9/9	9/9	-
B	Matched PCA (30D)	label ASW / J	0.574 $\pm$ 0.021	3/3	9/9	9/9	-
B	Matched PCA (30D)	batch ASW / J	0.831 $\pm$ 0.022	3/3	9/9	9/9	-
B	Matched PCA (30D)	NMI / J	0.669 $\pm$ 0.032	3/3	9/9	9/9	-
B	Matched PCA (30D)	PCR batch / J	0.437 $\pm$ 0.298	3/3	9/9	9/9	-
B	Matched PCA (30D)	Balanced accuracy / Q	0.803 $\pm$ 0.096	3/3	9/9	9/9	-
B	Matched PCA (30D)	cLISI / J	0.989 $\pm$ 0.010	3/3	9/9	9/9	-
B	Matched PCA (30D)	Cell-cycle conservation / J	0.660 $\pm$ 0.227	3/3	9/9	9/9	-
B	Matched PCA (30D)	HVG overlap / J	NA (SD -)	0/3	0/9	0/9	H
B	Matched PCA (30D)	iLISI / J	0.304 $\pm$ 0.168	3/3	9/9	9/9	-
B	Matched PCA (30D)	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	9/9	-
B	Matched PCA (30D)	Macro-F1 / Q	0.774 $\pm$ 0.127	3/3	9/9	9/9	-
B	Matched PCA (30D)	Pseudotime smoothness / J	0.894 $\pm$ 0.001	1/1	3/3	3/3	S
B	Matched PCA (30D)	Pseudotime smoothness / Q	0.977 (SD -)	1/1	1/3	1/3	P,S
B	Matched PCA (30D)	Pseudotime smoothness / R	0.876 $\pm$ 0.031	1/1	3/3	3/3	S
B	Matched PCA (30D)	Trajectory conservation / J	0.873 $\pm$ 0.003	1/1	3/3	3/3	T
B	Matched PCA (30D)	Trustworthiness / J	0.989 $\pm$ 0.005	2/2	6/6	6/6	S
B	Matched PCA (30D)	Trustworthiness / Q	0.969 $\pm$ 0.008	2/2	6/6	6/6	S
B	Matched PCA (30D)	Trustworthiness / R	0.988 $\pm$ 0.007	2/2	6/6	6/6	S
B	Matched PCA (30D)	Raw donor silhouette / J	0.054 $\pm$ 0.007	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw donor silhouette / Q	NA (SD -)	0/1	0/3	0/3	C,S
B	Matched PCA (30D)	Raw donor silhouette / R	0.069 $\pm$ 0.013	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw patient-group silhouette / J	0.325 $\pm$ 0.002	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw patient-group silhouette / Q	NA (SD -)	0/1	0/3	0/3	C,S
B	Matched PCA (30D)	Raw patient-group silhouette / R	0.324 $\pm$ 0.008	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw sample-ID silhouette / J	0.222 $\pm$ 0.028	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw sample-ID silhouette / Q	NA (SD -)	0/1	0/3	0/3	C,S
B	Matched PCA (30D)	Raw sample-ID silhouette / R	0.213 $\pm$ 0.033	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw study silhouette / J	0.222 $\pm$ 0.028	1/1	3/3	3/3	S
B	Matched PCA (30D)	Raw study silhouette / Q	NA (SD -)	0/1	0/3	0/3	C,S
B	Matched PCA (30D)	Raw study silhouette / R	0.213 $\pm$ 0.033	1/1	3/3	3/3	S
B	scVI 30D	ARI / J	0.353 $\pm$ 0.057	3/3	9/9	45/45	-
B	scVI 30D	label ASW / J	0.578 $\pm$ 0.031	3/3	9/9	45/45	-
B	scVI 30D	batch ASW / J	0.878 $\pm$ 0.020	3/3	9/9	45/45	-
B	scVI 30D	NMI / J	0.677 $\pm$ 0.009	3/3	9/9	45/45	-

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
B	scVI 30D	PCR batch / J	0.871 $\pm$ 0.069	3/3	9/9	45/45	–
B	scVI 30D	Balanced accuracy / Q	0.855 $\pm$ 0.053	3/3	9/9	45/45	–
B	scVI 30D	cLISI / J	0.911 $\pm$ 0.086	3/3	9/9	45/45	–
B	scVI 30D	Cell-cycle conservation / J	0.536 $\pm$ 0.072	3/3	9/9	45/45	–
B	scVI 30D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
B	scVI 30D	iLISI / J	2.125 $\pm$ 0.302	3/3	9/9	45/45	–
B	scVI 30D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
B	scVI 30D	Macro-F1 / Q	0.807 $\pm$ 0.091	3/3	9/9	45/45	–
B	scVI 30D	Pseudotime smoothness / J	0.880 $\pm$ 0.003	1/1	3/3	15/15	S
B	scVI 30D	Pseudotime smoothness / Q	0.969 (SD –)	1/1	1/3	5/15	P,S
B	scVI 30D	Pseudotime smoothness / R	0.864 $\pm$ 0.026	1/1	3/3	15/15	S
B	scVI 30D	Trajectory conservation / J	0.830 $\pm$ 0.027	1/1	3/3	15/15	T
B	scVI 30D	Trustworthiness / J	0.888 $\pm$ 0.009	2/2	6/6	30/30	S
B	scVI 30D	Trustworthiness / Q	0.946 $\pm$ 0.019	2/2	6/6	30/30	S
B	scVI 30D	Trustworthiness / R	0.884 $\pm$ 0.006	2/2	6/6	30/30	S
B	scVI 30D	Raw donor silhouette / J	-0.052 $\pm$ 0.004	1/1	3/3	15/15	S
B	scVI 30D	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scVI 30D	Raw donor silhouette / R	-0.050 $\pm$ 0.003	1/1	3/3	15/15	S
B	scVI 30D	Raw patient-group silhouette / J	0.093 $\pm$ 0.002	1/1	3/3	15/15	S
B	scVI 30D	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scVI 30D	Raw patient-group silhouette / R	0.097 $\pm$ 0.001	1/1	3/3	15/15	S
B	scVI 30D	Raw sample-ID silhouette / J	0.009 $\pm$ 0.022	1/1	3/3	15/15	S
B	scVI 30D	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scVI 30D	Raw sample-ID silhouette / R	-0.001 $\pm$ 0.009	1/1	3/3	15/15	S
B	scVI 30D	Raw study silhouette / J	0.009 $\pm$ 0.022	1/1	3/3	15/15	S
B	scVI 30D	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scVI 30D	Raw study silhouette / R	-0.001 $\pm$ 0.009	1/1	3/3	15/15	S

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
B	scANVI 30D	ARI / J	0.389 $\pm$ 0.082	3/3	9/9	45/45	–
B	scANVI 30D	label ASW / J	0.657 $\pm$ 0.018	3/3	9/9	45/45	–
B	scANVI 30D	batch ASW / J	0.867 $\pm$ 0.015	3/3	9/9	45/45	–
B	scANVI 30D	NMI / J	0.753 $\pm$ 0.046	3/3	9/9	45/45	–
B	scANVI 30D	PCR batch / J	0.836 $\pm$ 0.065	3/3	9/9	45/45	–
B	scANVI 30D	Balanced accuracy / Q	0.882 $\pm$ 0.061	3/3	9/9	45/45	–
B	scANVI 30D	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
B	scANVI 30D	Cell-cycle conservation / J	0.604 $\pm$ 0.159	3/3	9/9	45/45	–
B	scANVI 30D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
B	scANVI 30D	iLISI / J	1.902 $\pm$ 0.189	3/3	9/9	45/45	–
B	scANVI 30D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
B	scANVI 30D	Macro-F1 / Q	0.838 $\pm$ 0.087	3/3	9/9	45/45	–
B	scANVI 30D	Pseudotime smoothness / J	0.886 $\pm$ 0.001	1/1	3/3	15/15	S
B	scANVI 30D	Pseudotime smoothness / Q	0.970 (SD –)	1/1	1/3	5/15	P,S
B	scANVI 30D	Pseudotime smoothness / R	0.868 $\pm$ 0.030	1/1	3/3	15/15	S
B	scANVI 30D	Trajectory conservation / J	0.855 $\pm$ 0.003	1/1	3/3	15/15	T
B	scANVI 30D	Trustworthiness / J	0.910 $\pm$ 0.018	2/2	6/6	30/30	S
B	scANVI 30D	Trustworthiness / Q	0.949 $\pm$ 0.018	2/2	6/6	30/30	S
B	scANVI 30D	Trustworthiness / R	0.906 $\pm$ 0.021	2/2	6/6	30/30	S
B	scANVI 30D	Raw donor silhouette / J	-0.060 $\pm$ 0.003	1/1	3/3	15/15	S
B	scANVI 30D	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scANVI 30D	Raw donor silhouette / R	-0.057 $\pm$ 0.002	1/1	3/3	15/15	S
B	scANVI 30D	Raw patient-group silhouette / J	0.106 $\pm$ 0.004	1/1	3/3	15/15	S
B	scANVI 30D	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scANVI 30D	Raw patient-group silhouette / R	0.111 $\pm$ 0.008	1/1	3/3	15/15	S
B	scANVI 30D	Raw sample-ID silhouette / J	-0.003 $\pm$ 0.020	1/1	3/3	15/15	S
B	scANVI 30D	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scANVI 30D	Raw sample-ID silhouette / R	-0.007 $\pm$ 0.015	1/1	3/3	15/15	S
B	scANVI 30D	Raw study silhouette / J	-0.003 $\pm$ 0.020	1/1	3/3	15/15	S
B	scANVI 30D	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
B	scANVI 30D	Raw study silhouette / R	-0.007 $\pm$ 0.015	1/1	3/3	15/15	S
B	SCLSC 30D	ARI / J	0.353 $\pm$ 0.120	3/3	9/9	45/45	–
B	SCLSC 30D	label ASW / J	0.948 $\pm$ 0.013	3/3	9/9	45/45	–
B	SCLSC 30D	batch ASW / J	0.743 $\pm$ 0.021	3/3	9/9	45/45	–
B	SCLSC 30D	NMI / J	0.757 $\pm$ 0.064	3/3	9/9	45/45	–
B	SCLSC 30D	PCR batch / J	0.783 $\pm$ 0.115	3/3	9/9	45/45	–
B	SCLSC 30D	Balanced accuracy / Q	0.845 $\pm$ 0.054	3/3	9/9	45/45	–
B	SCLSC 30D	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
B	SCLSC 30D	Cell-cycle conservation / J	0.491 $\pm$ 0.126	3/3	9/9	45/45	–
B	SCLSC 30D	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
B	SCLSC 30D	iLISI / J	2.241 $\pm$ 0.508	3/3	9/9	45/45	–
B	SCLSC 30D	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
B	SCLSC 30D	Macro-F1 / Q	0.815 $\pm$ 0.078	3/3	9/9	45/45	–
B	SCLSC 30D	Pseudotime smoothness / J	0.818 $\pm$ 0.006	1/1	3/3	15/15	S
B	SCLSC 30D	Pseudotime smoothness / Q	0.934 (SD –)	1/1	1/3	5/15	P,S

Table S7 (continued)

Block	Method/output	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
B	SCLSC 30D	Pseudotime smoothness / R	0.794 $\pm$ 0.036	1/1	3/3	15/15	S
B	SCLSC 30D	Trajectory conservation / J	0.435 $\pm$ 0.032	1/1	2/3	10/15	T
B	SCLSC 30D	Trustworthiness / J	0.904 $\pm$ 0.023	2/2	6/6	30/30	S
B	SCLSC 30D	Trustworthiness / Q	0.903 $\pm$ 0.024	2/2	6/6	30/30	S
B	SCLSC 30D	Trustworthiness / R	0.898 $\pm$ 0.030	2/2	6/6	30/30	S
B	SCLSC 30D	Raw donor silhouette / J	-0.232 $\pm$ 0.014	1/1	3/3	15/15	S
B	SCLSC 30D	Raw donor silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SCLSC 30D	Raw donor silhouette / R	-0.220 $\pm$ 0.010	1/1	3/3	15/15	S
B	SCLSC 30D	Raw patient-group silhouette / J	0.231 $\pm$ 0.040	1/1	3/3	15/15	S
B	SCLSC 30D	Raw patient-group silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SCLSC 30D	Raw patient-group silhouette / R	0.261 $\pm$ 0.044	1/1	3/3	15/15	S
B	SCLSC 30D	Raw sample-ID silhouette / J	-0.011 $\pm$ 0.040	1/1	3/3	15/15	S
B	SCLSC 30D	Raw sample-ID silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SCLSC 30D	Raw sample-ID silhouette / R	0.005 $\pm$ 0.020	1/1	3/3	15/15	S
B	SCLSC 30D	Raw study silhouette / J	-0.011 $\pm$ 0.040	1/1	3/3	15/15	S
B	SCLSC 30D	Raw study silhouette / Q	NA (SD -)	0/1	0/3	0/15	C,S
B	SCLSC 30D	Raw study silhouette / R	0.005 $\pm$ 0.020	1/1	3/3	15/15	S

R: reference; Q: query; J: joint; cLISI, cell-type local inverse Simpson's index. The source-space isolated-label silhouette is not a target-embedding score. Codes: S, outside the dataset scope; H, highly variable gene overlap disabled for embedding outputs; T, no source diffusion-pseudotime annotation or eligible target-graph root; P, fewer than 16 pseudotime-annotated cells; C, no biological label with at least 10 cells and a valid multi-group covariate silhouette. A batch without all five compatible seed scores is withheld; among immune queries, only Oetjen_A has pseudotime coverage.

Table S8: Full held-out output-configuration contrasts. L and R are the first and second named representations. For each contrast, fits are first averaged within a batch; differences are then calculated for batches with complete results for both configurations and averaged with equal dataset weights. D/B indicate dataset/batch coverage; F(L)/F(R) indicate contributing/expected fits.

Contrast	Metric / population	L	R	Δ	D	B	F(L)	F(R)
SOLAR30 minus128	ARI / J	0.441	0.404	+0.036	3/3	9/9	45/45	45/45
SOLAR30 minus128	label ASW / J	0.873	0.869	+0.004	3/3	9/9	45/45	45/45
SOLAR30 minus128	batch ASW / J	0.836	0.842	-0.006	3/3	9/9	45/45	45/45
SOLAR30 minus128	NMI / J	0.774	0.758	+0.016	3/3	9/9	45/45	45/45
SOLAR30 minus128	PCR batch / J	0.777	0.771	+0.005	3/3	9/9	45/45	45/45
SOLAR30 minus128	Balanced accuracy / Q	0.837	0.838	0.000	3/3	9/9	45/45	45/45
SOLAR30 minus128	cLISI / J	1.000	1.000	+0.000	3/3	9/9	45/45	45/45
SOLAR30 minus128	Cell-cycle conservation / J	0.527	0.539	-0.012	3/3	9/9	45/45	45/45
SOLAR30 minus128	HVG overlap / J	NA	NA	NA	0/3	0/9	0/45	0/45
SOLAR30 minus128	iLISI / J	1.740	1.268	+0.472	3/3	9/9	45/45	45/45
SOLAR30 minus128	Isolated-label silhouette (source) / J	0.501	0.501	+0.000	3/3	9/9	45/45	45/45
SOLAR30 minus128	Macro-F1 / Q	0.811	0.813	-0.002	3/3	9/9	45/45	45/45
SOLAR30 minus128	Pseudotime smoothness / J	0.861	0.873	-0.012	1/1	3/3	15/15	15/15
SOLAR30 minus128	Pseudotime smoothness / Q	0.947	0.949	-0.002	1/1	1/3	5/15	5/15
SOLAR30 minus128	Pseudotime smoothness / R	0.837	0.852	-0.014	1/1	3/3	15/15	15/15
SOLAR30 minus128	Trajectory conservation / J	0.654	0.623	+0.031	1/1	3/3	15/15	15/15
SOLAR30 minus128	Trustworthiness / J	0.912	0.933	-0.020	2/2	6/6	30/30	30/30
SOLAR30 minus128	Trustworthiness / Q	0.936	0.941	-0.004	2/2	6/6	30/30	30/30
SOLAR30 minus128	Trustworthiness / R	0.907	0.928	-0.021	2/2	6/6	30/30	30/30
SOLAR30 minus128	Raw donor silhouette / J	-0.108	-0.095	-0.012	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw donor silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SOLAR30 minus128	Raw donor silhouette / R	-0.101	-0.089	-0.012	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw patient-group silhouette / J	0.229	0.236	-0.007	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw patient-group silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SOLAR30 minus128	Raw patient-group silhouette / R	0.243	0.250	-0.006	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw sample-ID silhouette / J	-0.002	0.007	-0.009	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw sample-ID silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SOLAR30 minus128	Raw sample-ID silhouette / R	-0.008	0.001	-0.009	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw study silhouette / J	-0.002	0.007	-0.009	1/1	3/3	15/15	15/15
SOLAR30 minus128	Raw study silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SOLAR30 minus128	Raw study silhouette / R	-0.008	0.001	-0.009	1/1	3/3	15/15	15/15
SCLSC30 minus16	ARI / J	0.353	0.350	+0.003	3/3	9/9	45/45	45/45
SCLSC30 minus16	label ASW / J	0.948	0.950	-0.002	3/3	9/9	45/45	45/45
SCLSC30 minus16	batch ASW / J	0.743	0.739	+0.004	3/3	9/9	45/45	45/45
SCLSC30 minus16	NMI / J	0.757	0.756	+0.001	3/3	9/9	45/45	45/45
SCLSC30 minus16	PCR batch / J	0.783	0.787	-0.004	3/3	9/9	45/45	45/45
SCLSC30 minus16	Balanced accuracy / Q	0.845	0.853	-0.007	3/3	9/9	45/45	45/45
SCLSC30 minus16	cLISI / J	1.000	1.000	+0.000	3/3	9/9	45/45	45/45
SCLSC30 minus16	Cell-cycle conservation / J	0.491	0.501	-0.010	3/3	9/9	45/45	45/45
SCLSC30 minus16	HVG overlap / J	NA	NA	NA	0/3	0/9	0/45	0/45
SCLSC30 minus16	iLISI / J	2.241	2.291	-0.049	3/3	9/9	45/45	45/45

Table S8 (continued)

Contrast	Metric / population	L	R	Δ	D	B	F(L)	F(R)
SCLSC30 minus16	Isolated-label silhouette (source) / J	0.501	0.501	+0.000	3/3	9/9	45/45	45/45
SCLSC30 minus16	Macro-F1 / Q	0.815	0.821	-0.006	3/3	9/9	45/45	45/45
SCLSC30 minus16	Pseudotime smoothness / J	0.818	0.815	+0.003	1/1	3/3	15/15	15/15
SCLSC30 minus16	Pseudotime smoothness / Q	0.934	0.930	+0.004	1/1	1/3	5/15	5/15
SCLSC30 minus16	Pseudotime smoothness / R	0.794	0.792	+0.003	1/1	3/3	15/15	15/15
SCLSC30 minus16	Trajectory conservation / J	0.435	0.524	-0.089	1/1	2/3	10/15	10/15
SCLSC30 minus16	Trustworthiness / J	0.904	0.903	+0.001	2/2	6/6	30/30	30/30
SCLSC30 minus16	Trustworthiness / Q	0.903	0.903	+0.001	2/2	6/6	30/30	30/30
SCLSC30 minus16	Trustworthiness / R	0.898	0.897	+0.001	2/2	6/6	30/30	30/30
SCLSC30 minus16	Raw donor silhouette / J	-0.232	-0.246	+0.014	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw donor silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SCLSC30 minus16	Raw donor silhouette / R	-0.220	-0.233	+0.013	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw patient-group silhouette / J	0.231	0.234	-0.003	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw patient-group silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SCLSC30 minus16	Raw patient-group silhouette / R	0.261	0.264	-0.002	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw sample-ID silhouette / J	-0.011	-0.018	+0.007	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw sample-ID silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SCLSC30 minus16	Raw sample-ID silhouette / R	0.005	-0.001	+0.006	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw study silhouette / J	-0.011	-0.018	+0.007	1/1	3/3	15/15	15/15
SCLSC30 minus16	Raw study silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
SCLSC30 minus16	Raw study silhouette / R	0.005	-0.001	+0.006	1/1	3/3	15/15	15/15

The companion source tables give per-batch reasons and numerical/comparable seed counts. Pseudotime query coverage is one of three immune batches; trustworthiness spans immune and lung; raw covariate panels are dataset-specific; HVG overlap is disabled. Root-selection failures withhold incomplete trajectory batches rather than averaging fewer seeds.

Table S9: Full paired 128-dimensional anchor contrasts across the evaluated metrics and populations. L and R are repeated anchors and the named control. Positive signed differences are not uniformly favourable: raw metadata silhouettes have no universal desirable direction. D/B and F(L)/F(R) report dataset/batch and fit coverage.

Contrast	Metric / population	L	R	Δ	D	B	F(L)	F(R)
Repeated minus none128	ARI / J	0.404	0.392	+0.012	3/3	9/9	45/45	45/45
Repeated minus none128	label ASW / J	0.869	0.845	+0.024	3/3	9/9	45/45	45/45
Repeated minus none128	batch ASW / J	0.842	0.859	-0.017	3/3	9/9	45/45	45/45
Repeated minus none128	NMI / J	0.758	0.754	+0.004	3/3	9/9	45/45	45/45
Repeated minus none128	PCR batch / J	0.771	0.748	+0.023	3/3	9/9	45/45	45/45
Repeated minus none128	Balanced accuracy / Q	0.838	0.834	+0.004	3/3	9/9	45/45	45/45
Repeated minus none128	cLISI / J	1.000	1.000	0.000	3/3	9/9	45/45	45/45
Repeated minus none128	Cell-cycle conservation / J	0.539	0.570	-0.032	3/3	9/9	45/45	45/45
Repeated minus none128	HVG overlap / J	NA	NA	NA	0/3	0/9	0/45	0/45
Repeated minus none128	iLISI / J	1.268	0.887	+0.381	3/3	9/9	45/45	45/45
Repeated minus none128	Isolated-label silhouette (source) / J	0.501	0.501	+0.000	3/3	9/9	45/45	45/45
Repeated minus none128	Macro-F1 / Q	0.813	0.806	+0.007	3/3	9/9	45/45	45/45
Repeated minus none128	Pseudotime smoothness / J	0.873	0.875	-0.003	1/1	3/3	15/15	15/15
Repeated minus none128	Pseudotime smoothness / Q	0.949	0.953	-0.004	1/1	1/3	5/15	5/15
Repeated minus none128	Pseudotime smoothness / R	0.852	0.854	-0.002	1/1	3/3	15/15	15/15
Repeated minus none128	Trajectory conservation / J	0.623	0.685	-0.062	1/1	3/3	15/15	15/15
Repeated minus none128	Trustworthiness / J	0.933	0.947	-0.015	2/2	6/6	30/30	30/30
Repeated minus none128	Trustworthiness / Q	0.941	0.946	-0.006	2/2	6/6	30/30	30/30
Repeated minus none128	Trustworthiness / R	0.928	0.943	-0.015	2/2	6/6	30/30	30/30
Repeated minus none128	Raw donor silhouette / J	-0.095	-0.055	-0.040	1/1	3/3	15/15	15/15
Repeated minus none128	Raw donor silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus none128	Raw donor silhouette / R	-0.089	-0.049	-0.040	1/1	3/3	15/15	15/15
Repeated minus none128	Raw patient-group silhouette / J	0.236	0.224	+0.012	1/1	3/3	15/15	15/15
Repeated minus none128	Raw patient-group silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus none128	Raw patient-group silhouette / R	0.250	0.236	+0.014	1/1	3/3	15/15	15/15
Repeated minus none128	Raw sample-ID silhouette / J	0.007	0.032	-0.025	1/1	3/3	15/15	15/15
Repeated minus none128	Raw sample-ID silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus none128	Raw sample-ID silhouette / R	0.001	0.022	-0.020	1/1	3/3	15/15	15/15
Repeated minus none128	Raw study silhouette / J	0.007	0.032	-0.025	1/1	3/3	15/15	15/15
Repeated minus none128	Raw study silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus none128	Raw study silhouette / R	0.001	0.022	-0.020	1/1	3/3	15/15	15/15
Repeated minus unique128	ARI / J	0.404	0.390	+0.014	3/3	9/9	45/45	45/45
Repeated minus unique128	label ASW / J	0.869	0.844	+0.025	3/3	9/9	45/45	45/45
Repeated minus unique128	batch ASW / J	0.842	0.868	-0.026	3/3	9/9	45/45	45/45
Repeated minus unique128	NMI / J	0.758	0.754	+0.004	3/3	9/9	45/45	45/45
Repeated minus unique128	PCR batch / J	0.771	0.751	+0.020	3/3	9/9	45/45	45/45
Repeated minus unique128	Balanced accuracy / Q	0.838	0.838	-0.001	3/3	9/9	45/45	45/45
Repeated minus unique128	cLISI / J	1.000	1.000	0.000	3/3	9/9	45/45	45/45
Repeated minus unique128	Cell-cycle conservation / J	0.539	0.571	-0.033	3/3	9/9	45/45	45/45
Repeated minus unique128	HVG overlap / J	NA	NA	NA	0/3	0/9	0/45	0/45
Repeated minus unique128	iLISI / J	1.268	0.841	+0.427	3/3	9/9	45/45	45/45
Repeated minus unique128	Isolated-label silhouette (source) / J	0.501	0.501	+0.000	3/3	9/9	45/45	45/45

Table S9 (continued)

Contrast	Metric / population	L	R	Δ	D	B	F(L)	F(R)
Repeated minus unique128	Macro-F1 / Q	0.813	0.813	+0.000	3/3	9/9	45/45	45/45
Repeated minus unique128	Pseudotime smoothness / J	0.873	0.878	-0.005	1/1	3/3	15/15	15/15
Repeated minus unique128	Pseudotime smoothness / Q	0.949	0.953	-0.004	1/1	1/3	5/15	5/15
Repeated minus unique128	Pseudotime smoothness / R	0.852	0.857	-0.005	1/1	3/3	15/15	15/15
Repeated minus unique128	Trajectory conservation / J	0.623	0.762	-0.139	1/1	3/3	15/15	15/15
Repeated minus unique128	Trustworthiness / J	0.933	0.951	-0.018	2/2	6/6	30/30	30/30
Repeated minus unique128	Trustworthiness / Q	0.941	0.947	-0.007	2/2	6/6	30/30	30/30
Repeated minus unique128	Trustworthiness / R	0.928	0.947	-0.019	2/2	6/6	30/30	30/30
Repeated minus unique128	Raw donor silhouette / J	-0.095	-0.047	-0.048	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw donor silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus unique128	Raw donor silhouette / R	-0.089	-0.041	-0.048	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw patient-group silhouette / J	0.236	0.219	+0.017	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw patient-group silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus unique128	Raw patient-group silhouette / R	0.250	0.232	+0.018	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw sample-ID silhouette / J	0.007	0.037	-0.031	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw sample-ID silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus unique128	Raw sample-ID silhouette / R	0.001	0.026	-0.025	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw study silhouette / J	0.007	0.037	-0.031	1/1	3/3	15/15	15/15
Repeated minus unique128	Raw study silhouette / Q	NA	NA	NA	0/1	0/3	0/15	0/15
Repeated minus unique128	Raw study silhouette / R	0.001	0.026	-0.025	1/1	3/3	15/15	15/15

The companion source tables give per-batch reasons and numerical/comparable seed counts. Pseudotime query coverage is one of three immune batches; trustworthiness spans immune and lung; raw covariate panels are dataset-specific; HVG overlap is disabled. Root-selection failures withhold incomplete trajectory batches rather than averaging fewer seeds.

Table S9: (continued). Individual-arm summaries for the 128-dimensional anchor controls; coverage may differ from the paired subset shown in Table S9.

128D arm	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
repeated	ARI / J	0.404 $\pm$ 0.077	3/3	9/9	45/45	–
repeated	label ASW / J	0.869 $\pm$ 0.036	3/3	9/9	45/45	–
repeated	batch ASW / J	0.842 $\pm$ 0.020	3/3	9/9	45/45	–
repeated	NMI / J	0.758 $\pm$ 0.034	3/3	9/9	45/45	–
repeated	PCR batch / J	0.771 $\pm$ 0.118	3/3	9/9	45/45	–
repeated	Balanced accuracy / Q	0.838 $\pm$ 0.079	3/3	9/9	45/45	–
repeated	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
repeated	Cell-cycle conservation / J	0.539 $\pm$ 0.179	3/3	9/9	45/45	–
repeated	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
repeated	iLISI / J	1.268 $\pm$ 0.165	3/3	9/9	45/45	–
repeated	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
repeated	Macro-F1 / Q	0.813 $\pm$ 0.109	3/3	9/9	45/45	–
repeated	Pseudotime smoothness / J	0.873 $\pm$ 0.007	1/1	3/3	15/15	S
repeated	Pseudotime smoothness / Q	0.949 (SD –)	1/1	1/3	5/15	P,S
repeated	Pseudotime smoothness / R	0.852 $\pm$ 0.043	1/1	3/3	15/15	S
repeated	Trajectory conservation / J	0.623 $\pm$ 0.095	1/1	3/3	15/15	T
repeated	Trustworthiness / J	0.933 $\pm$ 0.018	2/2	6/6	30/30	S
repeated	Trustworthiness / Q	0.941 $\pm$ 0.012	2/2	6/6	30/30	S
repeated	Trustworthiness / R	0.928 $\pm$ 0.019	2/2	6/6	30/30	S
repeated	Raw donor silhouette / J	-0.095 $\pm$ 0.003	1/1	3/3	15/15	S
repeated	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
repeated	Raw donor silhouette / R	-0.089 $\pm$ 0.004	1/1	3/3	15/15	S
repeated	Raw patient-group silhouette / J	0.236 $\pm$ 0.015	1/1	3/3	15/15	S
repeated	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
repeated	Raw patient-group silhouette / R	0.250 $\pm$ 0.024	1/1	3/3	15/15	S
repeated	Raw sample-ID silhouette / J	0.007 $\pm$ 0.023	1/1	3/3	15/15	S
repeated	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
repeated	Raw sample-ID silhouette / R	0.001 $\pm$ 0.015	1/1	3/3	15/15	S
repeated	Raw study silhouette / J	0.007 $\pm$ 0.023	1/1	3/3	15/15	S
repeated	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
repeated	Raw study silhouette / R	0.001 $\pm$ 0.015	1/1	3/3	15/15	S
unique	ARI / J	0.390 $\pm$ 0.076	3/3	9/9	45/45	–
unique	label ASW / J	0.844 $\pm$ 0.039	3/3	9/9	45/45	–
unique	batch ASW / J	0.868 $\pm$ 0.014	3/3	9/9	45/45	–
unique	NMI / J	0.754 $\pm$ 0.035	3/3	9/9	45/45	–
unique	PCR batch / J	0.751 $\pm$ 0.138	3/3	9/9	45/45	–
unique	Balanced accuracy / Q	0.838 $\pm$ 0.084	3/3	9/9	45/45	–
unique	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
unique	Cell-cycle conservation / J	0.571 $\pm$ 0.188	3/3	9/9	45/45	–
unique	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
unique	iLISI / J	0.841 $\pm$ 0.175	3/3	9/9	45/45	–
unique	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
unique	Macro-F1 / Q	0.813 $\pm$ 0.118	3/3	9/9	45/45	–
unique	Pseudotime smoothness / J	0.878 $\pm$ 0.007	1/1	3/3	15/15	S

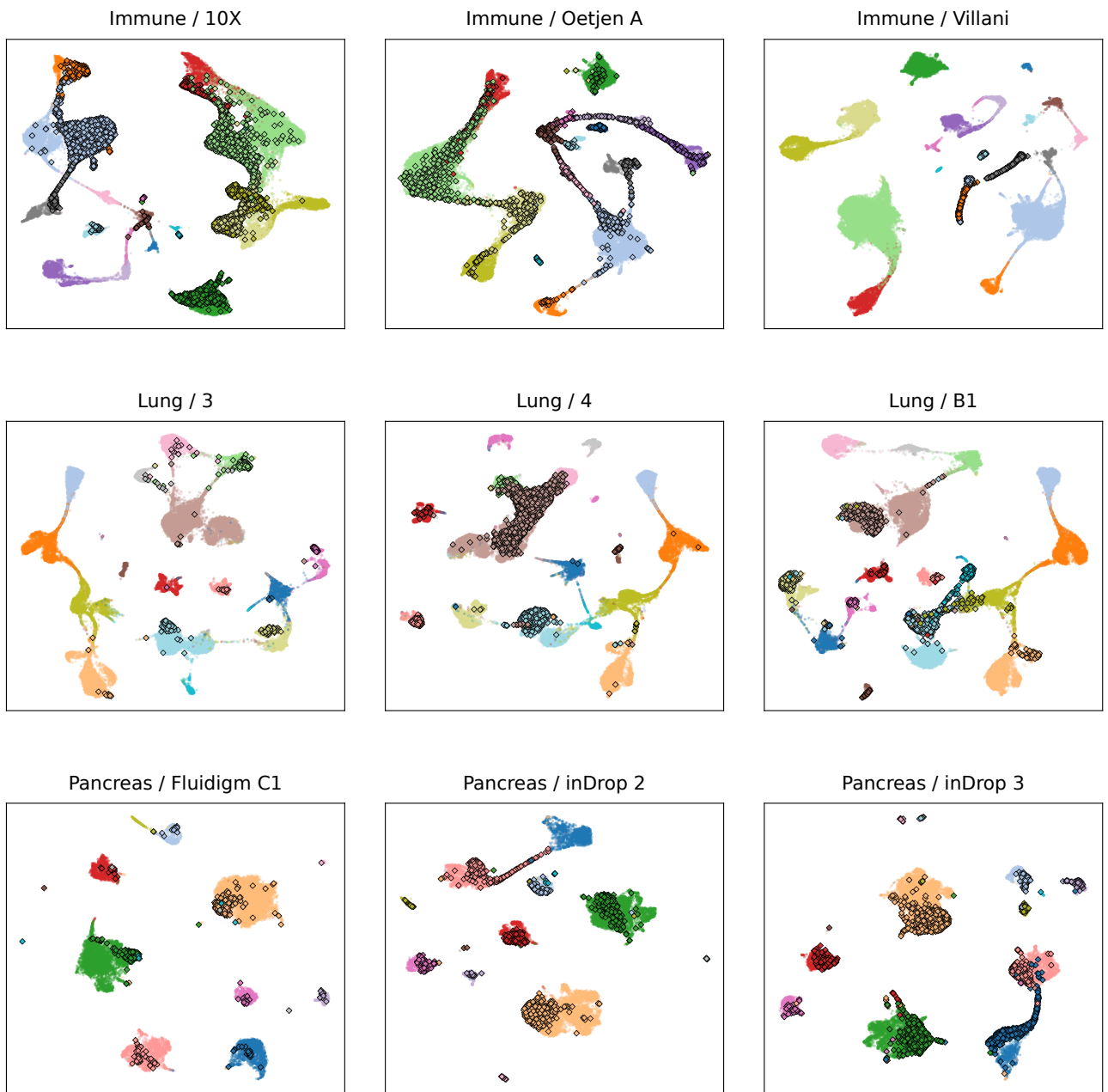
Table S9 (continued)

128D arm	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
unique	Pseudotime smoothness / Q	0.953 (SD $-$)	1/1	1/3	5/15	P,S
unique	Pseudotime smoothness / R	0.857 ± 0.043	1/1	3/3	15/15	S
unique	Trajectory conservation / J	0.762 ± 0.105	1/1	3/3	15/15	T
unique	Trustworthiness / J	0.951 ± 0.010	2/2	6/6	30/30	S
unique	Trustworthiness / Q	0.947 ± 0.010	2/2	6/6	30/30	S
unique	Trustworthiness / R	0.947 ± 0.012	2/2	6/6	30/30	S
unique	Raw donor silhouette / J	-0.047 ± 0.011	1/1	3/3	15/15	S
unique	Raw donor silhouette / Q	NA (SD $-$)	0/1	0/3	0/15	C,S
unique	Raw donor silhouette / R	-0.041 ± 0.008	1/1	3/3	15/15	S
unique	Raw patient-group silhouette / J	0.219 ± 0.012	1/1	3/3	15/15	S
unique	Raw patient-group silhouette / Q	NA (SD $-$)	0/1	0/3	0/15	C,S
unique	Raw patient-group silhouette / R	0.232 ± 0.018	1/1	3/3	15/15	S
unique	Raw sample-ID silhouette / J	0.037 ± 0.018	1/1	3/3	15/15	S
unique	Raw sample-ID silhouette / Q	NA (SD $-$)	0/1	0/3	0/15	C,S
unique	Raw sample-ID silhouette / R	0.026 ± 0.006	1/1	3/3	15/15	S
unique	Raw study silhouette / J	0.037 ± 0.018	1/1	3/3	15/15	S
unique	Raw study silhouette / Q	NA (SD $-$)	0/1	0/3	0/15	C,S
unique	Raw study silhouette / R	0.026 ± 0.006	1/1	3/3	15/15	S

Table S9 (continued)

128D arm	Metric / population	Mean $\pm$ batch SD	D	B	F	NA
none	ARI / J	0.392 $\pm$ 0.074	3/3	9/9	45/45	–
none	label ASW / J	0.845 $\pm$ 0.033	3/3	9/9	45/45	–
none	batch ASW / J	0.859 $\pm$ 0.012	3/3	9/9	45/45	–
none	NMI / J	0.754 $\pm$ 0.035	3/3	9/9	45/45	–
none	PCR batch / J	0.748 $\pm$ 0.142	3/3	9/9	45/45	–
none	Balanced accuracy / Q	0.834 $\pm$ 0.083	3/3	9/9	45/45	–
none	cLISI / J	1.000 $\pm$ 0.000	3/3	9/9	45/45	–
none	Cell-cycle conservation / J	0.570 $\pm$ 0.185	3/3	9/9	45/45	–
none	HVG overlap / J	NA (SD –)	0/3	0/9	0/45	H
none	iLISI / J	0.887 $\pm$ 0.187	3/3	9/9	45/45	–
none	Isolated-label silhouette (source) / J	0.501 $\pm$ 0.010	3/3	9/9	45/45	–
none	Macro-F1 / Q	0.806 $\pm$ 0.115	3/3	9/9	45/45	–
none	Pseudotime smoothness / J	0.875 $\pm$ 0.005	1/1	3/3	15/15	S
none	Pseudotime smoothness / Q	0.953 (SD –)	1/1	1/3	5/15	P,S
none	Pseudotime smoothness / R	0.854 $\pm$ 0.043	1/1	3/3	15/15	S
none	Trajectory conservation / J	0.685 $\pm$ 0.114	1/1	3/3	15/15	T
none	Trustworthiness / J	0.947 $\pm$ 0.011	2/2	6/6	30/30	S
none	Trustworthiness / Q	0.946 $\pm$ 0.010	2/2	6/6	30/30	S
none	Trustworthiness / R	0.943 $\pm$ 0.013	2/2	6/6	30/30	S
none	Raw donor silhouette / J	-0.055 $\pm$ 0.009	1/1	3/3	15/15	S
none	Raw donor silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
none	Raw donor silhouette / R	-0.049 $\pm$ 0.006	1/1	3/3	15/15	S
none	Raw patient-group silhouette / J	0.224 $\pm$ 0.012	1/1	3/3	15/15	S
none	Raw patient-group silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
none	Raw patient-group silhouette / R	0.236 $\pm$ 0.018	1/1	3/3	15/15	S
none	Raw sample-ID silhouette / J	0.032 $\pm$ 0.019	1/1	3/3	15/15	S
none	Raw sample-ID silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
none	Raw sample-ID silhouette / R	0.022 $\pm$ 0.005	1/1	3/3	15/15	S
none	Raw study silhouette / J	0.032 $\pm$ 0.019	1/1	3/3	15/15	S
none	Raw study silhouette / Q	NA (SD –)	0/1	0/3	0/15	C,S
none	Raw study silhouette / R	0.022 $\pm$ 0.005	1/1	3/3	15/15	S

4 Supplementary figures



Small dot = reference cell (alpha 0.35). Diamond = held-out query cell.
Colour = cell-type label (per-panel palette). Descriptive only, not quantitative evidence.

Figure S1. Joint reference–query UMAP overview for the nine real held-out batches. The 128-dimensional repeated-anchor SOLAR model with seed 40 is shown.

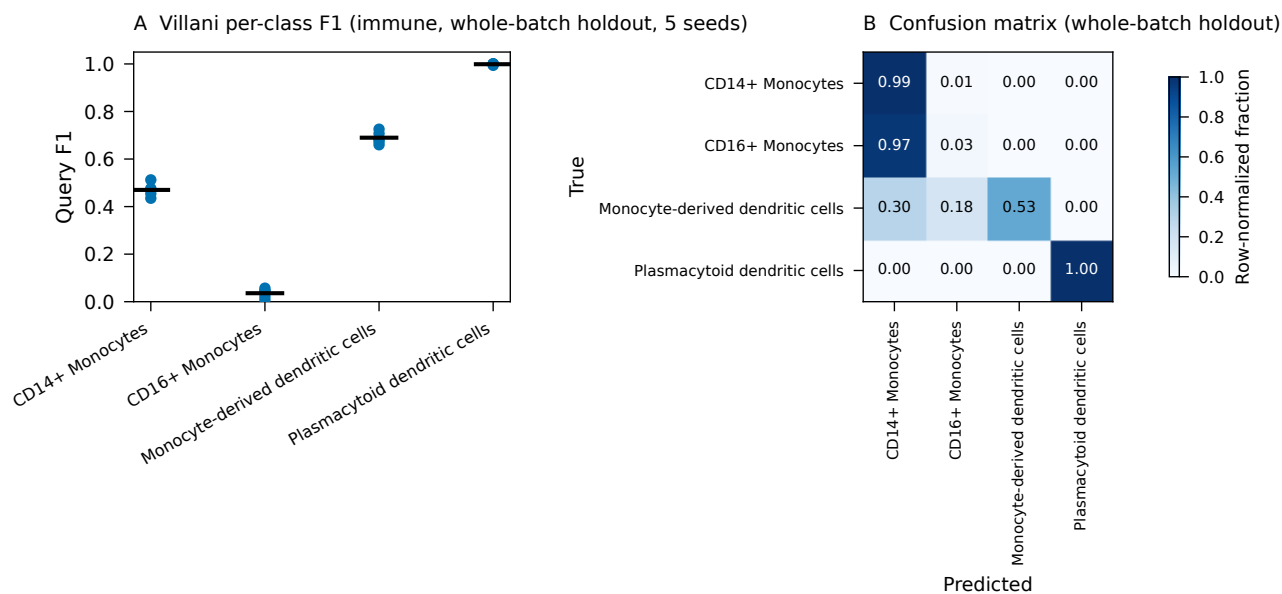


Figure S2. Error analysis for the held-out Villani batch. Per-class query F1 and the row-normalized confusion matrix for the 128-dimensional repeated-anchor SOLAR representation show errors concentrated in CD16+ monocytes despite complete reference-label coverage.

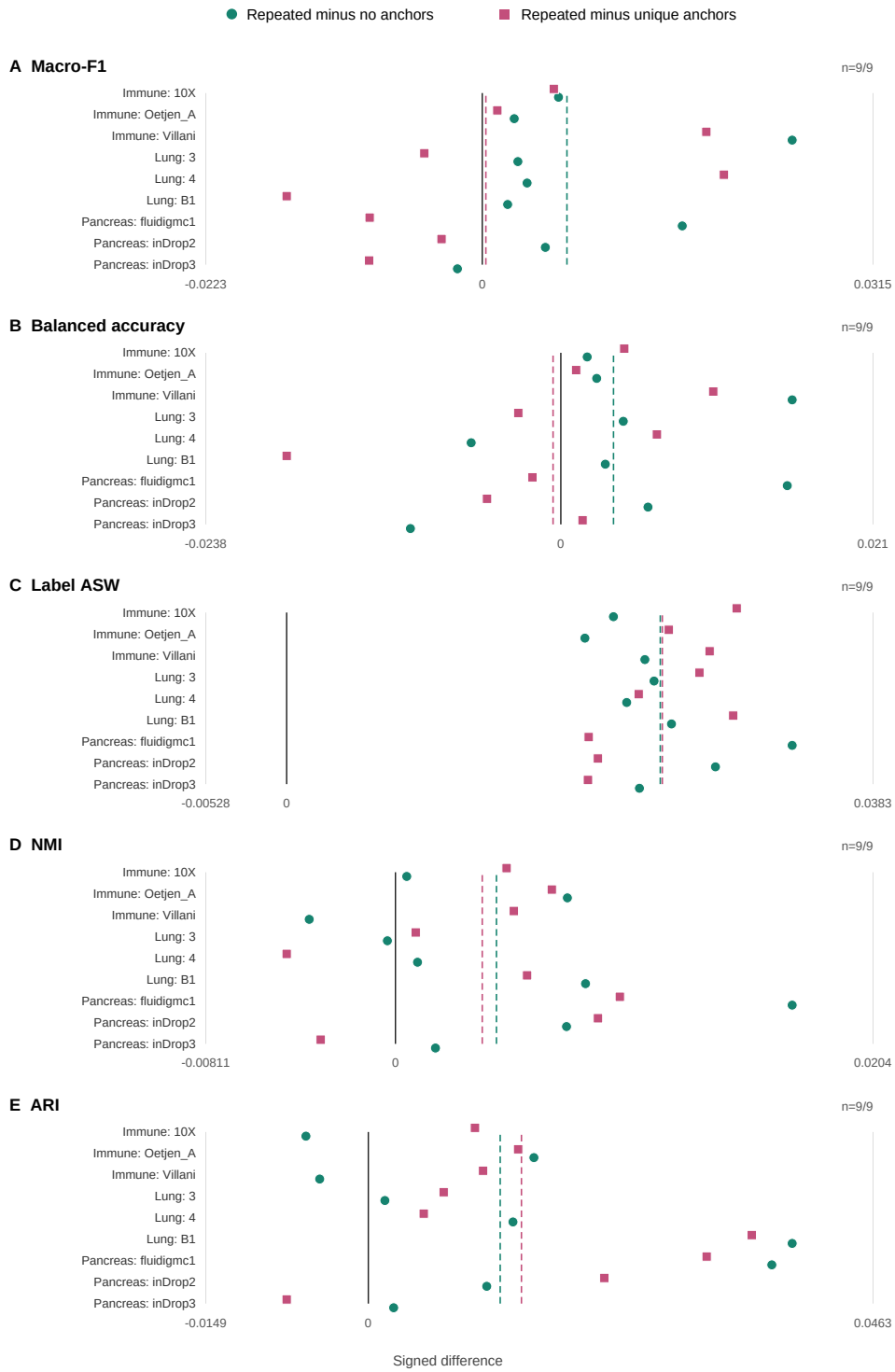


Figure S3. Metric-level 128-dimensional anchor ablation (panels A–E). Points show repeated-anchor SOLAR minus the no-anchor or unique-anchor control after averaging five fits within each held-out batch. Macro-F1 and balanced accuracy use query cells; label ASW, normalized mutual information (NMI), and adjusted Rand index (ARI) use the joint reference–query population. The solid vertical line marks zero, and coloured dashed lines show equal-dataset means of the batch differences.

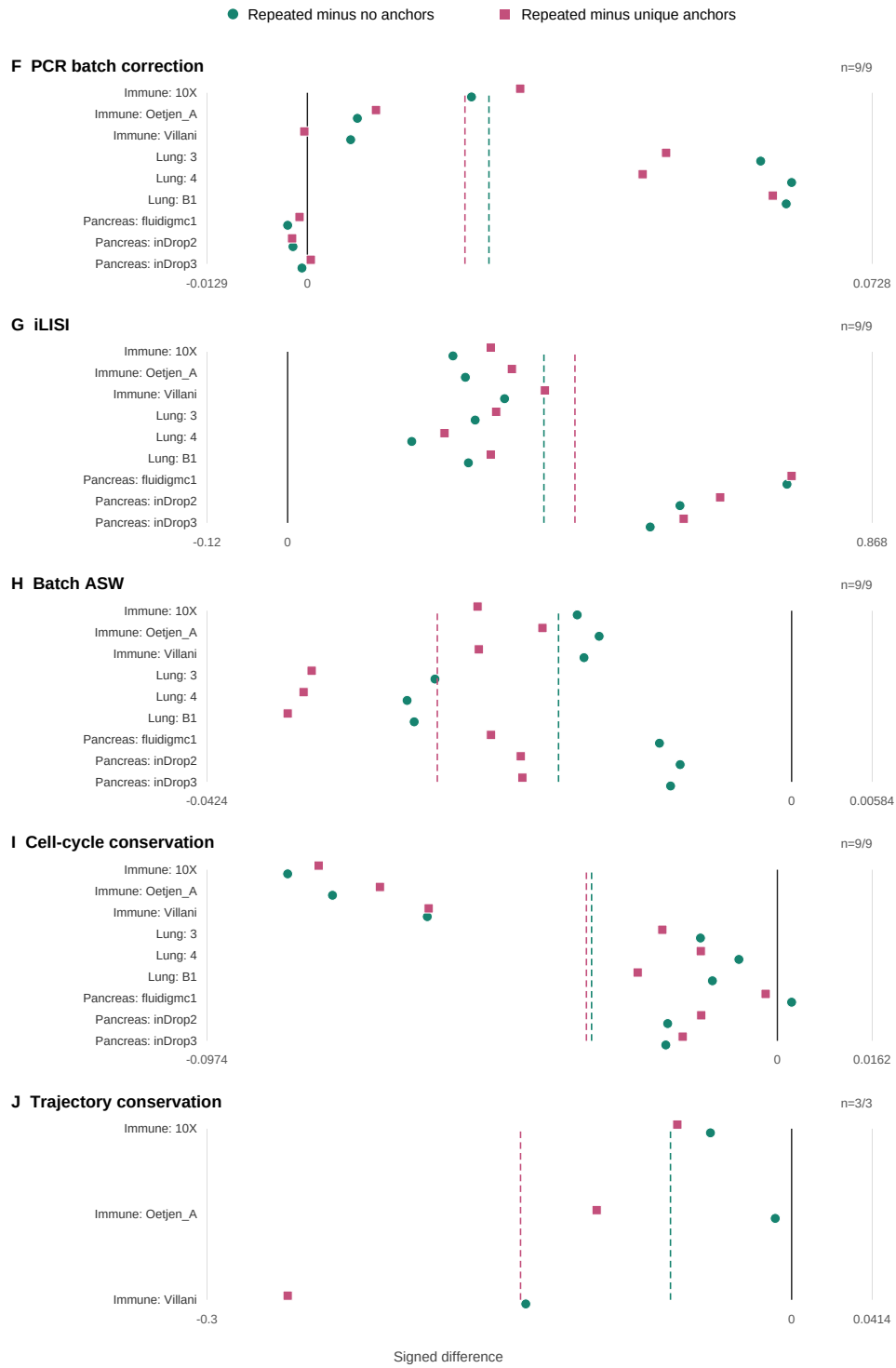


Figure S3 (continued; panels F–J). PCR batch correction, iLISI, batch ASW, cell-cycle conservation, and trajectory conservation use the joint reference–query population. Trajectory conservation applies to the three immune batches; the other displayed metrics cover all nine held-out batches. Table S9 reports evaluated metric–population combinations, including local-structure and metadata endpoints not plotted here.

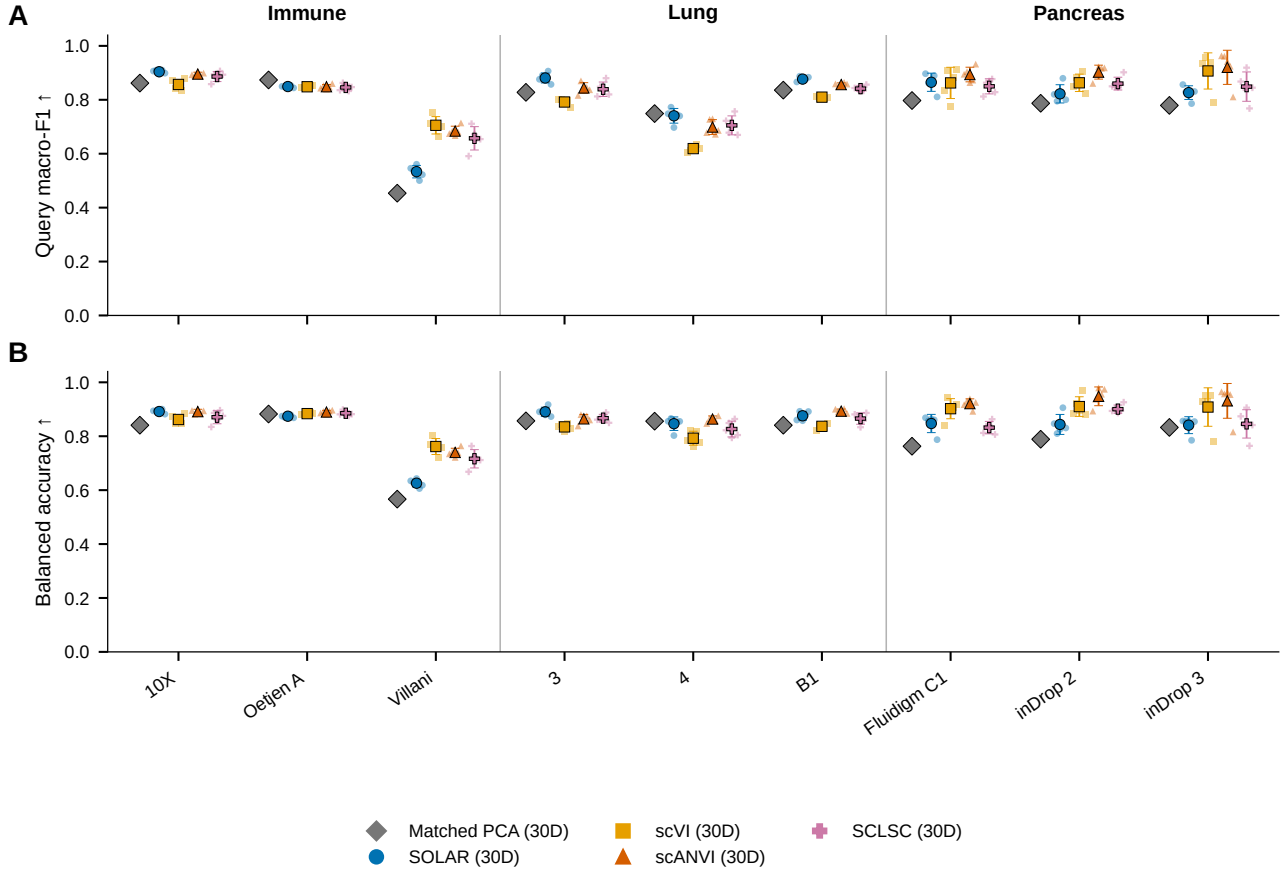


Figure S4. Query-label recovery with common 30-dimensional outputs. A shows macro-F1 and B balanced accuracy across the nine held-out batches using the reference-trained, distance-weighted 15-nearest-neighbour readout. Small points are the five fitting-seed values within a batch; large points and bars are their mean and sample SD. Matched PCA is deterministic and therefore has no fitting-seed SD. SOLAR uses 40 reference-fitted principal components as input before producing its 30-dimensional output. Tables S3 and S7 report the corresponding annotation and representation results.

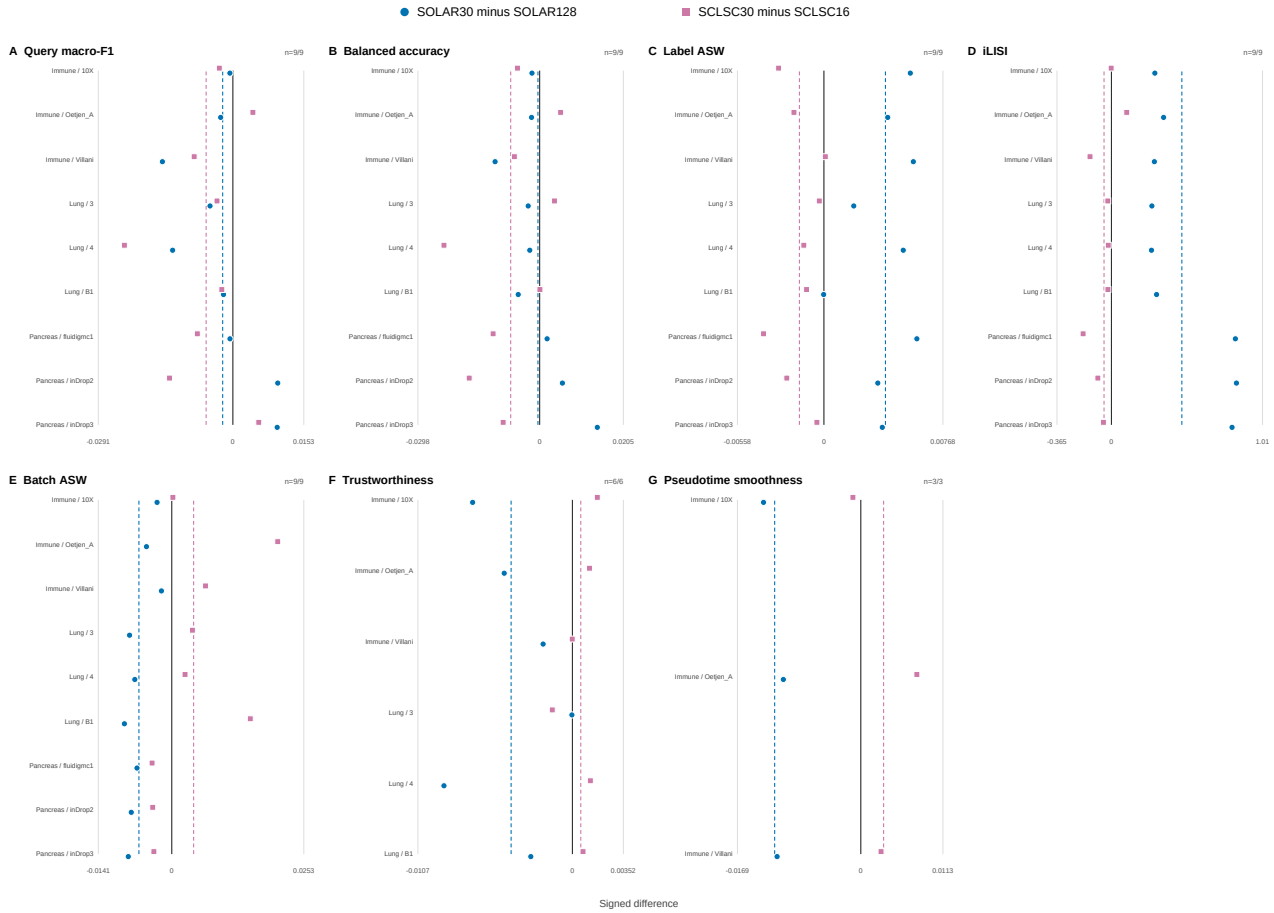


Figure S5. Held-out output-configuration comparisons. Points show SOLAR with 30-dimensional outputs minus SOLAR with 128-dimensional outputs, and SCLSC with 30-dimensional outputs minus SCLSC with 16-dimensional outputs. The five fits are first averaged within each batch, and contrasts use batches with complete estimates for both configurations. Macro-F1, balanced accuracy, and trustworthiness use query cells; label ASW, iLISI, batch ASW, and pseudotime smoothness use joint populations. Dashed lines show equal-dataset means of the batch differences. Trustworthiness applies to six immune/lung batches, pseudotime smoothness to three immune batches, and the other displayed endpoints to all nine. Table S8 reports the complete coverage.

References

- [1] Yusri Dwi Heryanto, Yao-zhong Zhang, and Seiya Imoto. Predicting cell types with supervised contrastive learning on cells and their types. *Scientific Reports*, 14:430, 2024.
- [2] F Alexander Wolf, Philipp Angerer, and Fabian J Theis. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology*, 19:15, 2018.
- [3] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *International Conference on Learning Representations*, 2019:1–19, 2019.
- [4] Malte D Luecken, Maren Büttner, Kridsakorn Chaichoompu, Anna Danese, Marta Interlandi, Michaela F Müller, Daniel C Strobl, Luke Zappia, Martin Dugas, Maria Colomé-Tatché, et al. Benchmarking atlas-level data integration in single-cell genomics. *Nature methods*, 19(1):41–50, 2022.
- [5] Jarkko Venna and Samuel Kaski. Neighborhood preservation in nonlinear projection methods: an experimental study. In *Artificial Neural Networks—ICANN 2001*, volume 2130 of *Lecture Notes in Computer Science*, pages 485–491, 2001.