nature methods

Supplementary information

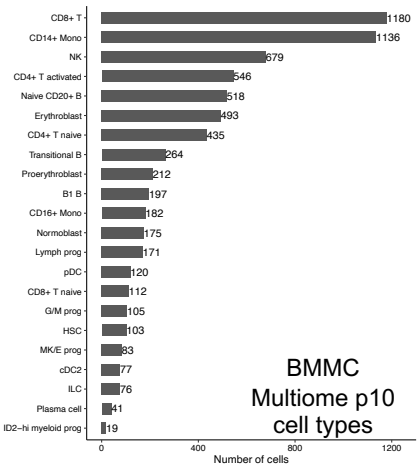
https://doi.org/10.1038/s41592-025-02737-9

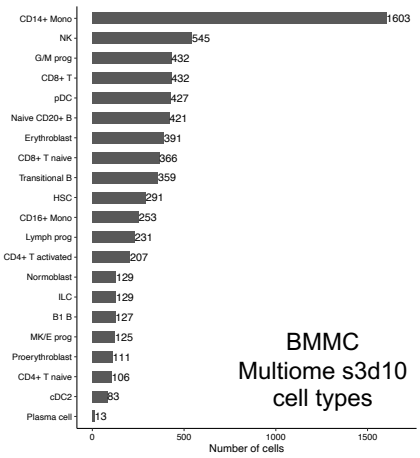
Benchmarking single-cell multi-modal data integrations

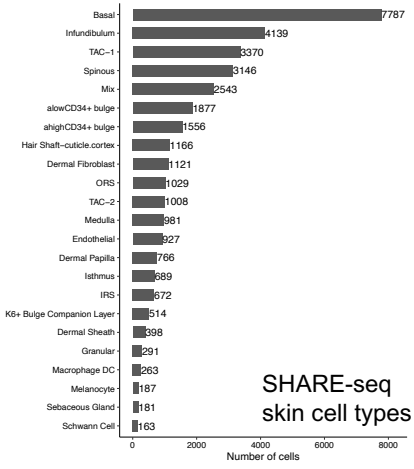
In the format provided by the authors and unedited

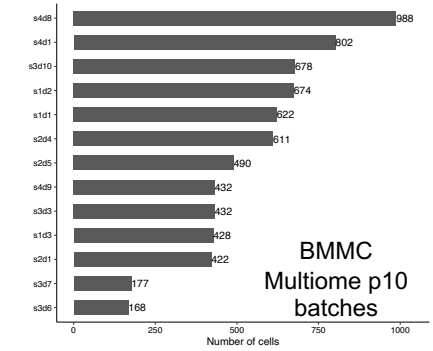
Benchmarking single-cell multi-modal data integrations

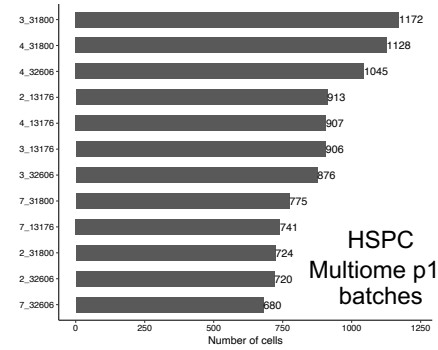
Supplementary Information

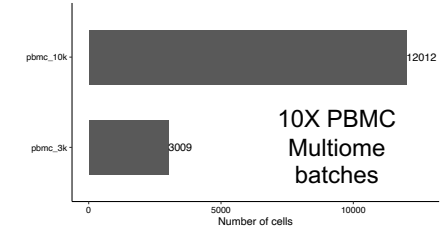


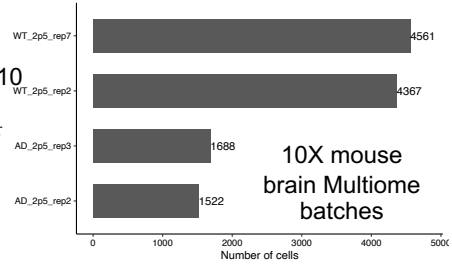




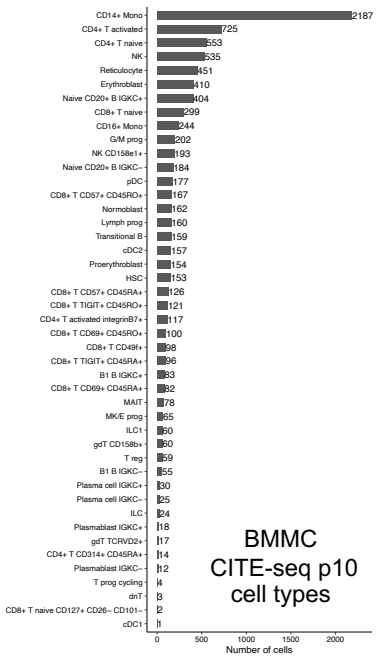




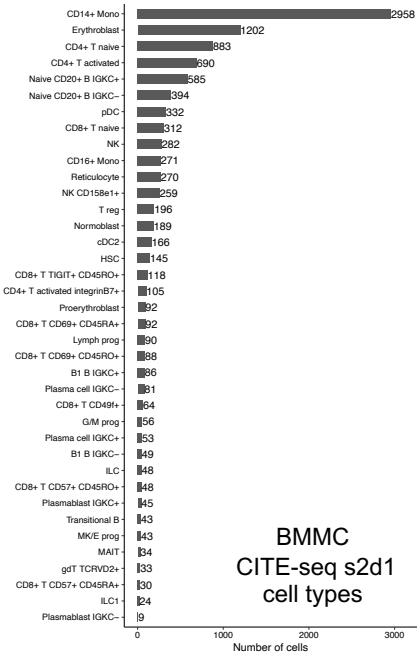




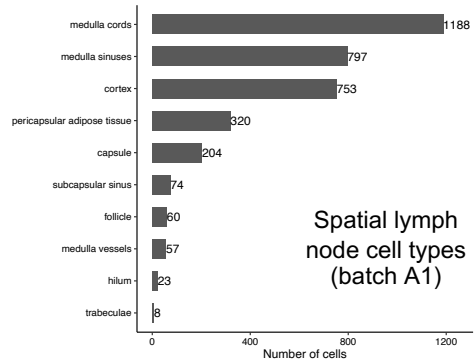
Supplementary Fig. 1. Summary of cell number in batches and cell types of scRNA and scATAC multimodal datasets used in the SCMMIB study. The y axis of boxplot shows the name of batch or cell type, and the x axis represented the number of cells.



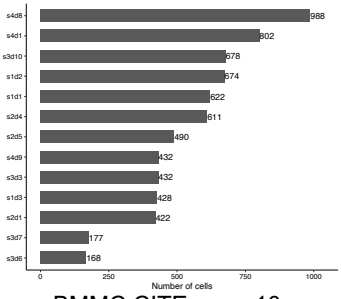
BMMC CITE-seq p10 cell types



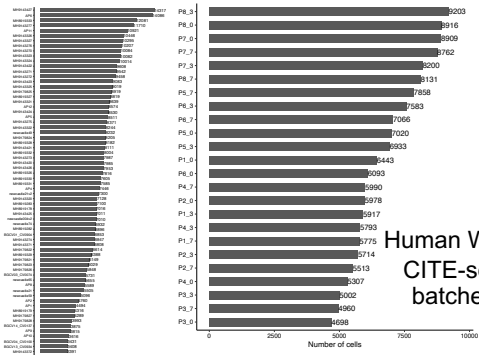
BMMC CITE-seq s2d1 cell types



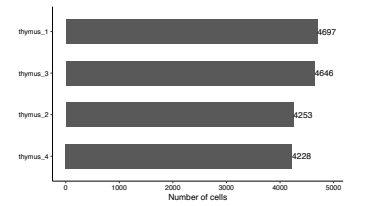
Spatial lymph node cell types (batch A1)



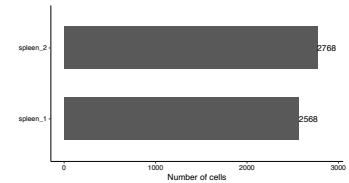
BMMC CITE-seq p10 batches



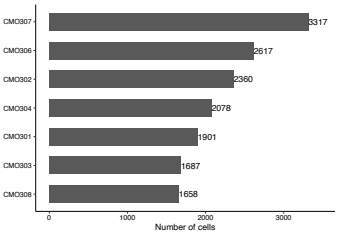
Human WBC CITE-seq batches



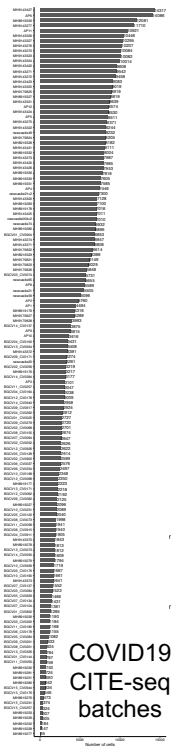
Thymus spatial batches



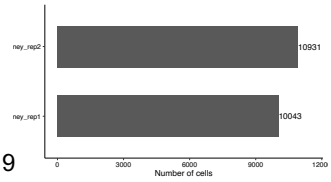
Spleen SPOTS batches



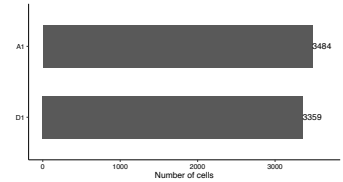
10X NSCLC CITE-seq batches



COVID19 CITE-seq batches

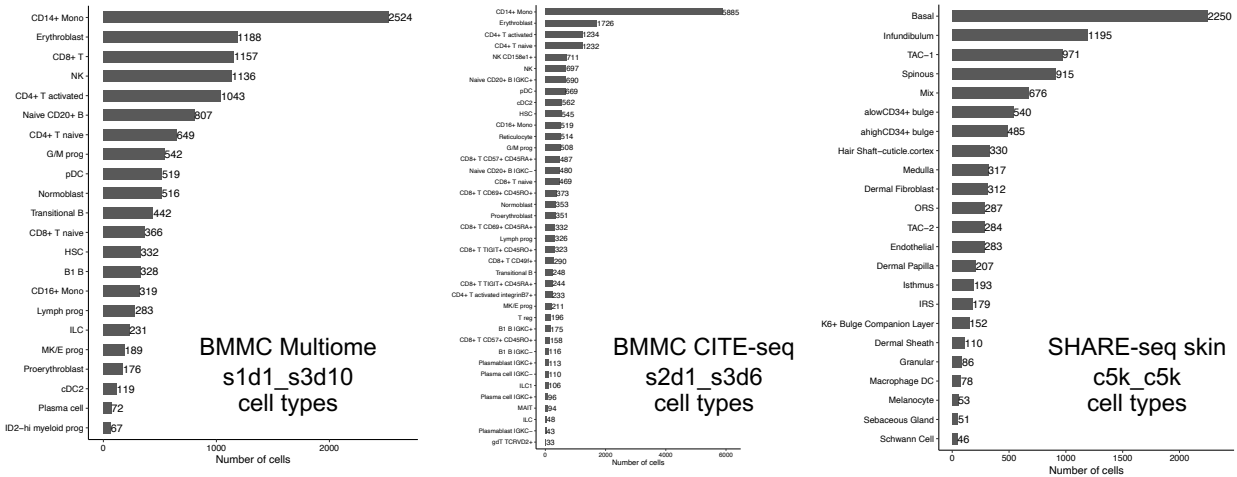


10X kidney cancer CITE-seq batches

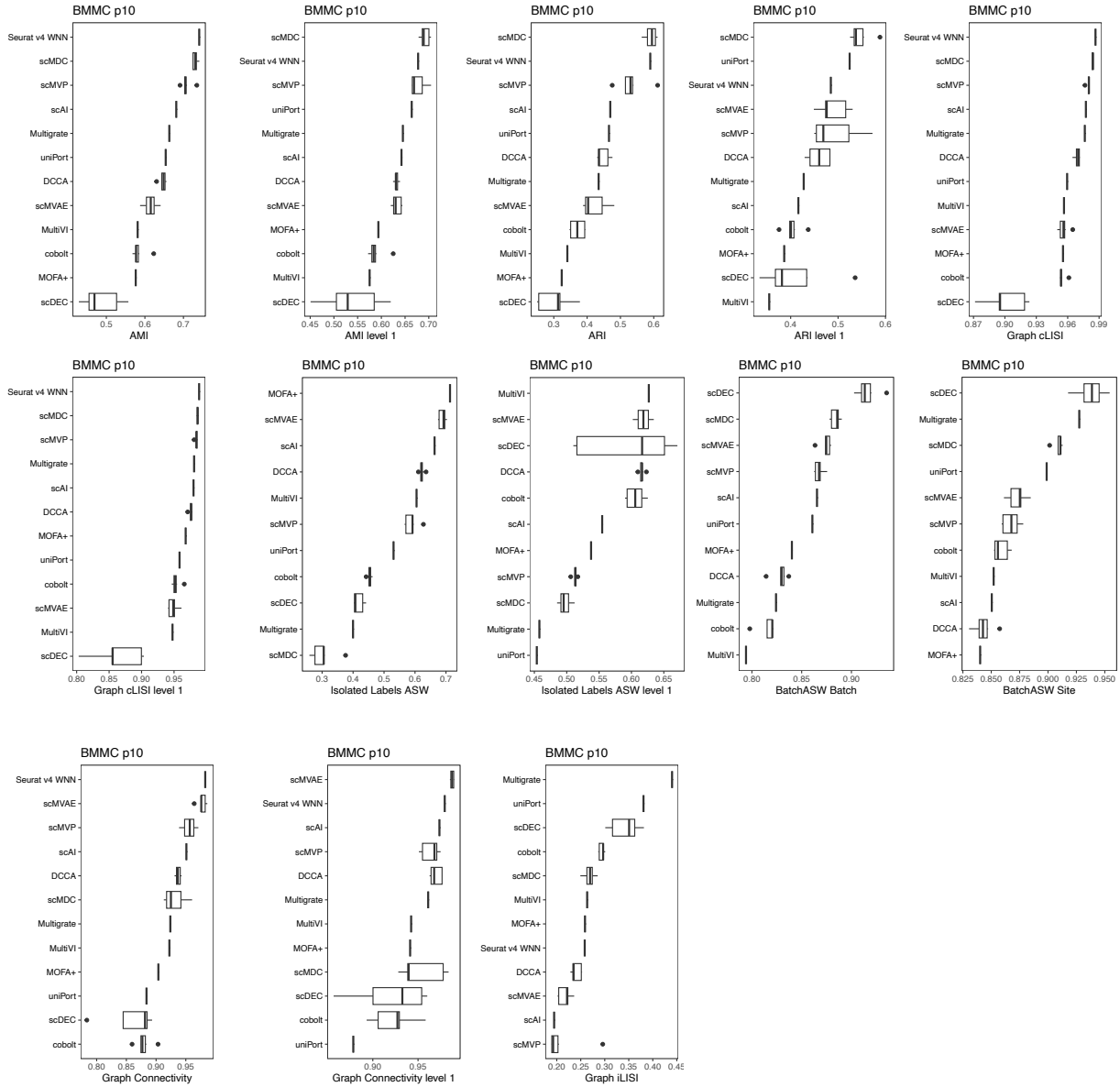


Lymph node spatial batches

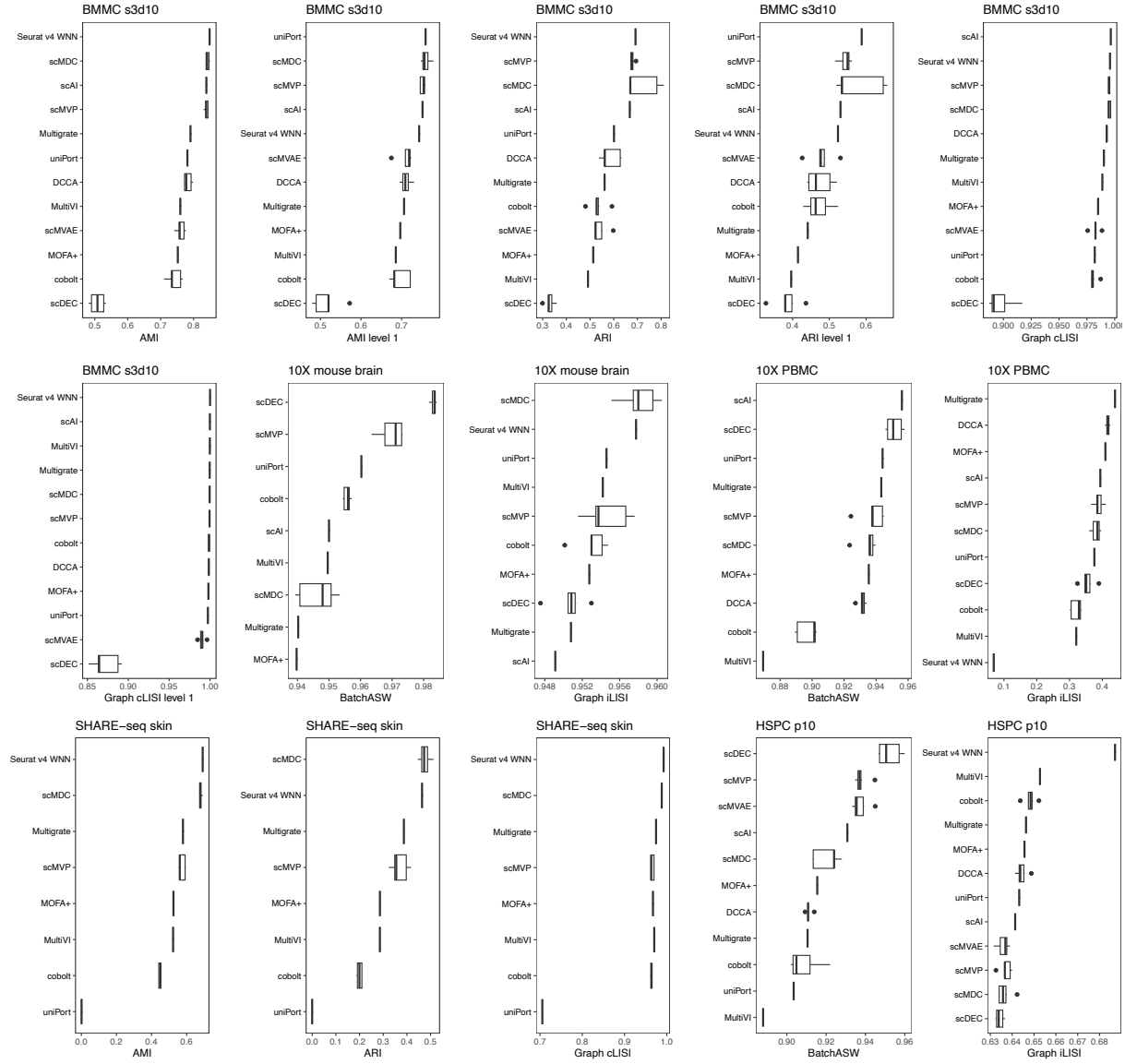
Supplementary Fig. 2. Summary of cell number in batches and cell types of scRNA and ADT datasets as well as spatial scRNA and ADT datasets used in the SCMMIB study. The y axis of boxplot shows the name of batch or cell type, and the x axis represented the number of cells.



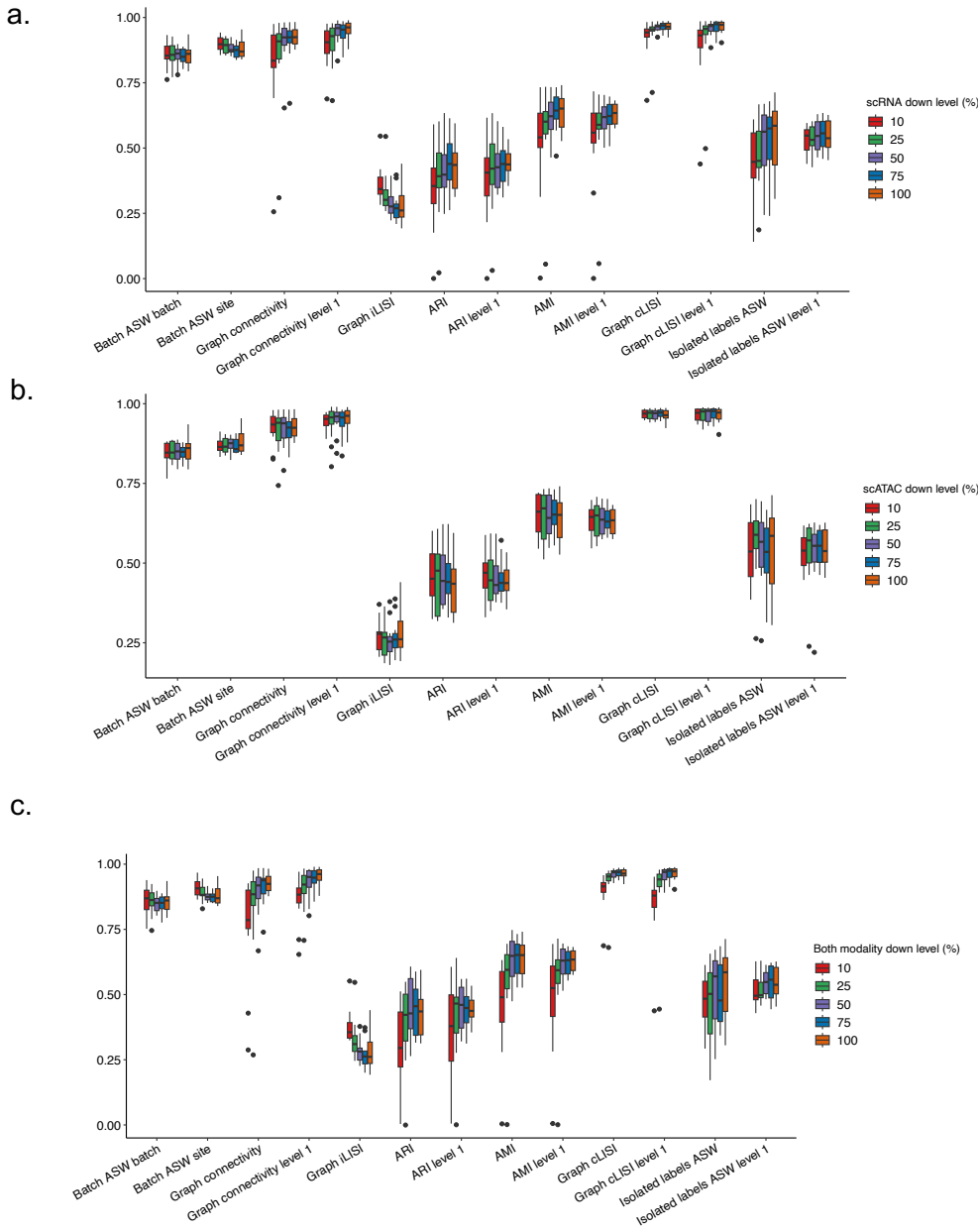
Supplementary Fig. 3. Summary of cell number in cell types of mosaic integration simulation datasets used in the SCMMIB study. The y axis of boxplot shows the name of cell type, and the x axis represented the number of cells.



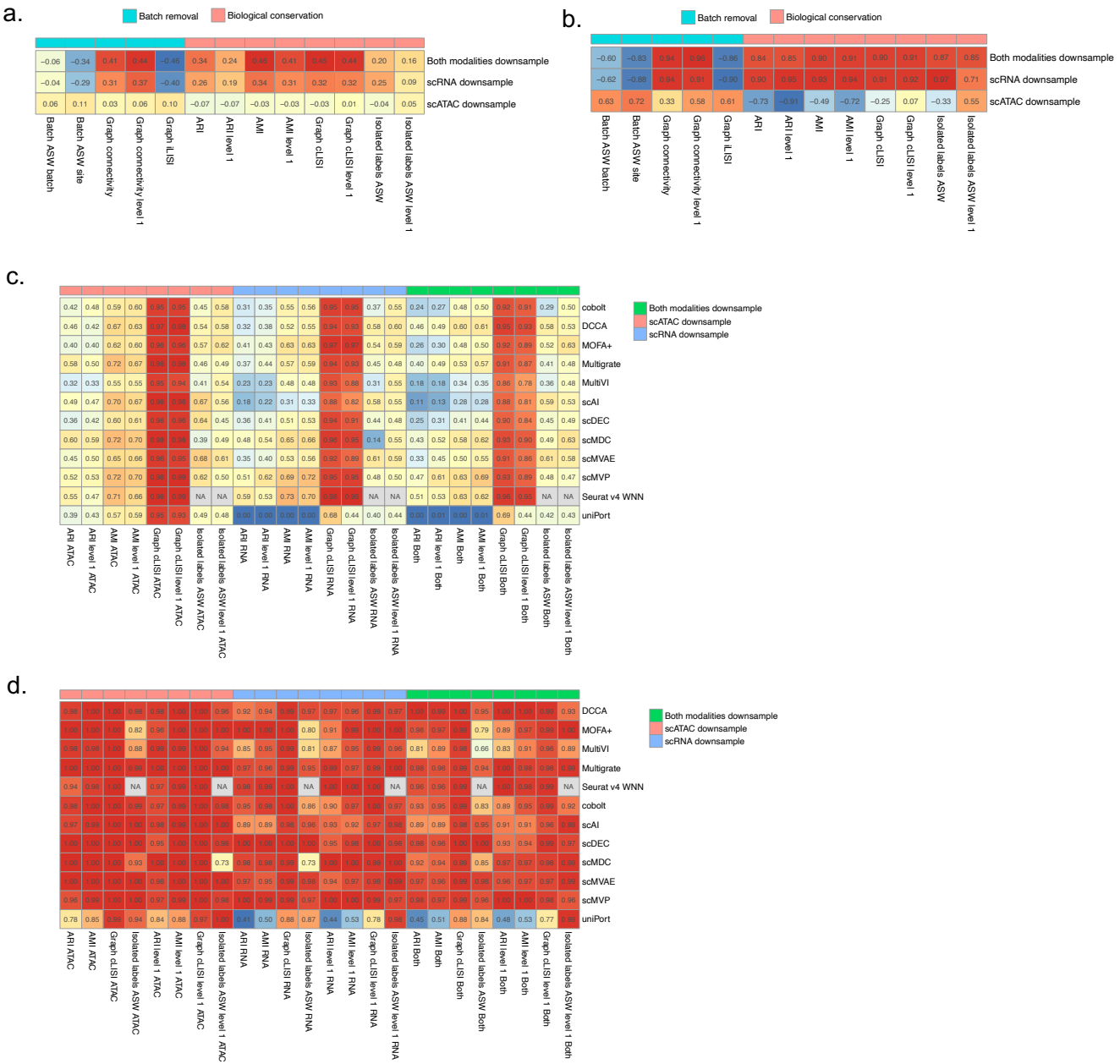
Supplementary Fig. 4. Boxplot of accuracy metrics for paired scRNA and scATAC integration algorithms in BMMC Multiome p10 dataset, corresponding to Fig. 2b and Extended Data Fig. 4b, Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



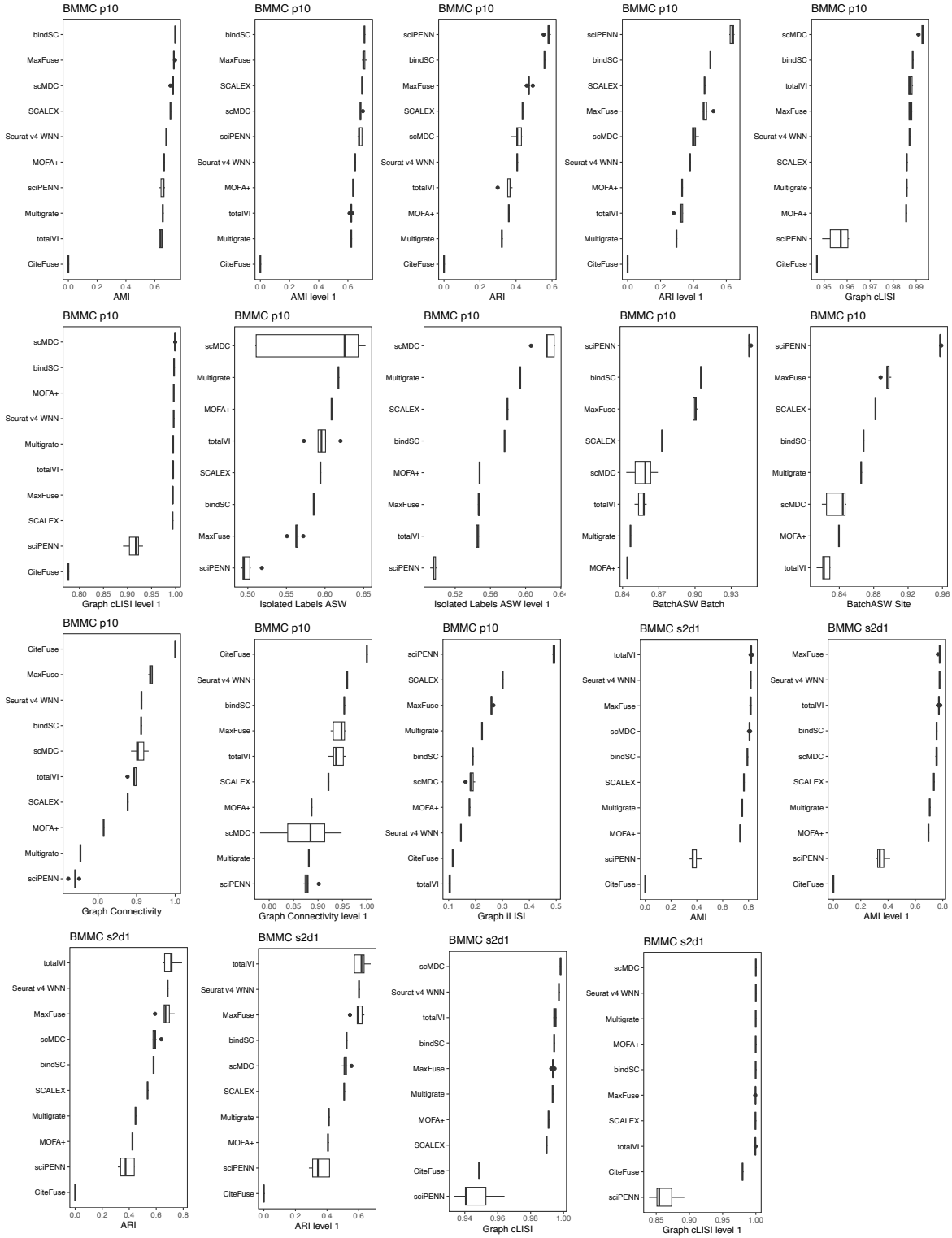
Supplementary Fig. 5. Boxplot of accuracy metrics for paired scRNA and scATAC integration algorithms in other datasets except for BMMC Multiome p10, corresponding to Fig. 2b and Extended Data Fig. 4b. The dataset names are labelled at the top of each boxplot. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



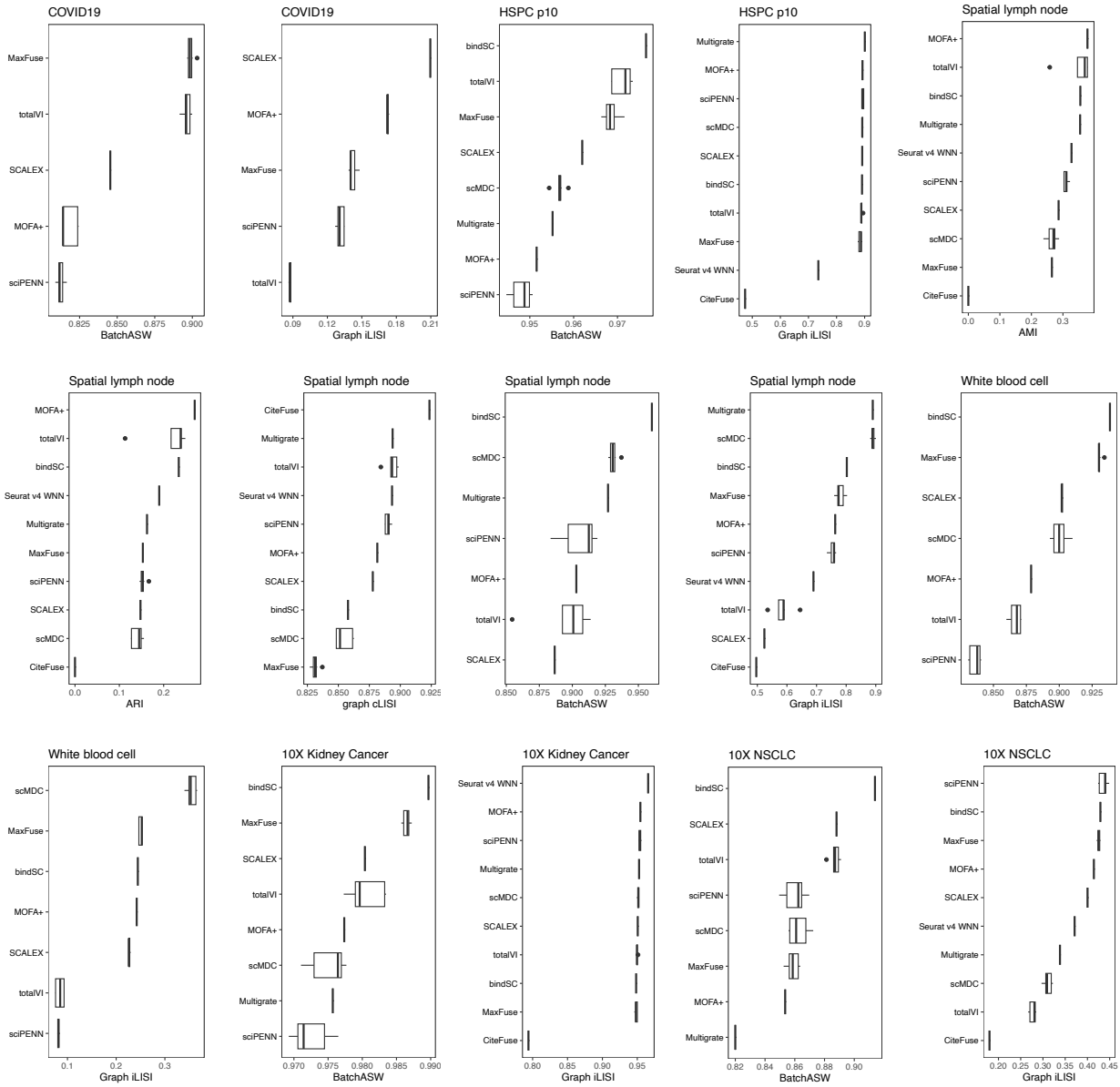
Supplementary Fig. 6. Boxplot of the biological conservation and batch removal metric values for all paired scRNA and scATAC integration methods at all downsampling levels. This figure shows the metric scores distribution for simulations of **(a)** scRNA modality downsampling, **(b)** scATAC modality downsampling and **(c)** both modalities downsampling. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



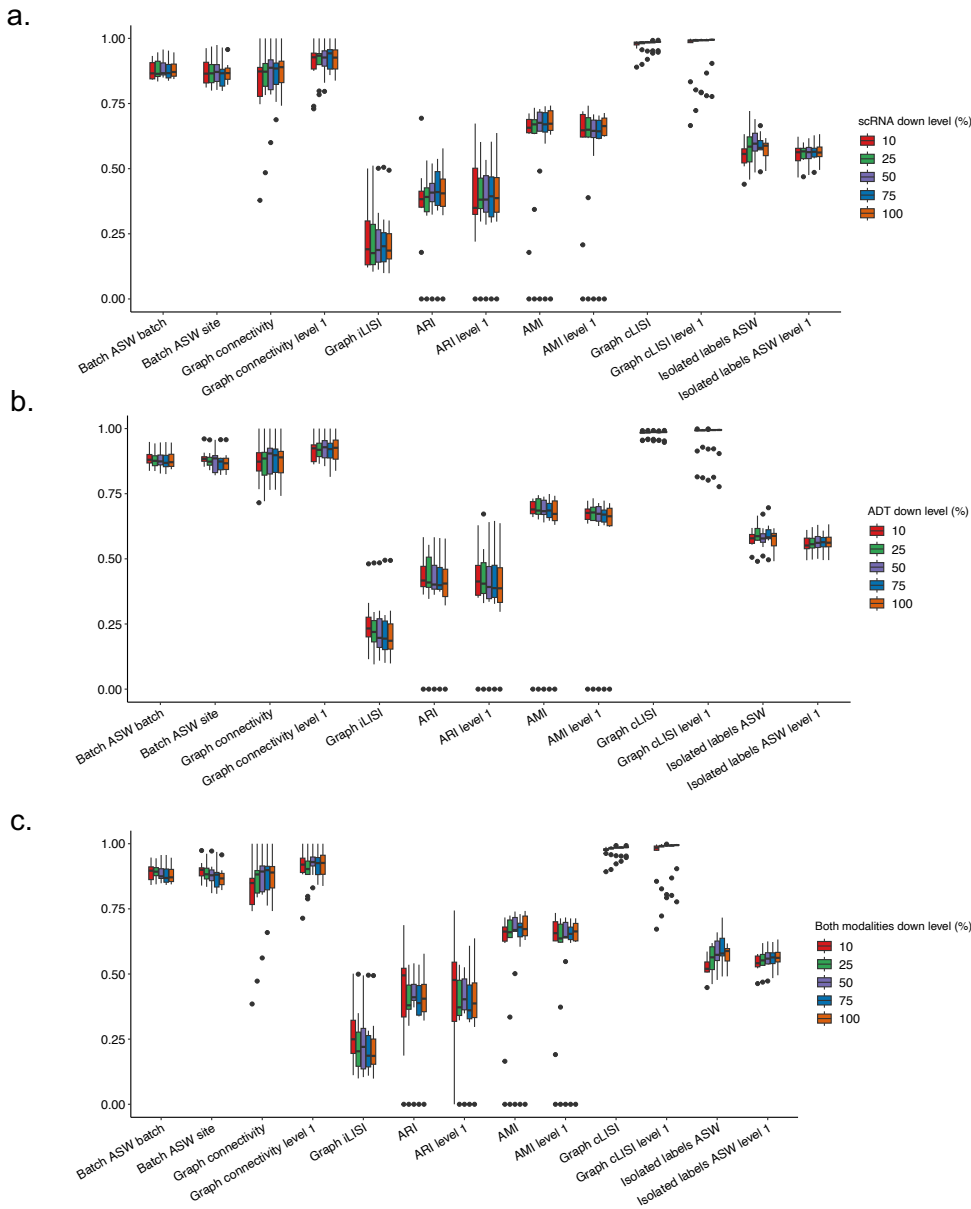
Supplementary Fig. 7. Details of the paired scRNA and scATAC integration robustness metrics in Fig. 2c and Extended Data Fig. 4d. (**a-b**). Correlations between metric values and downsample gradients. The color and number in heatmaps shows Pearson's r correlation between (**a)** metric values or (**b**) average metric values with the downsampling gradients. (**c**) Detail metric values for all paired scRNA and scATAC integration algorithms for BMMC p10 dataset at the 10% sequencing depth. (**d**) Detail downsample AUC values for all paired scRNA and scATAC integration algorithms.



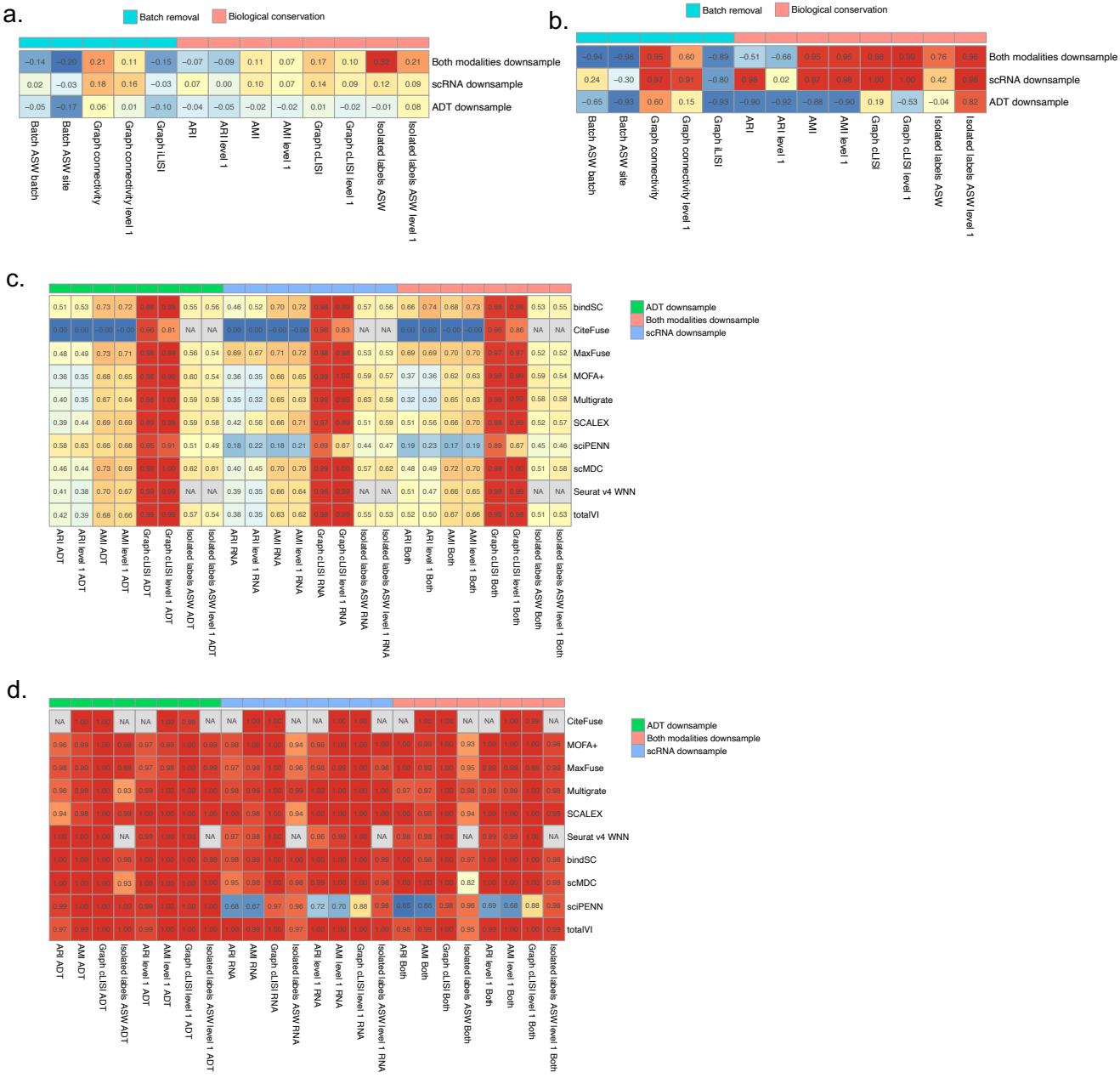
Supplementary Fig. 8. Boxplot of accuracy metrics for paired scRNA and ADT integration algorithms in BMMC CITE-seq p10 and s2d1 datasets, corresponding to Fig. 3b and Extended Data Fig. 5b. The dataset names are labelled at the top of each boxplot. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



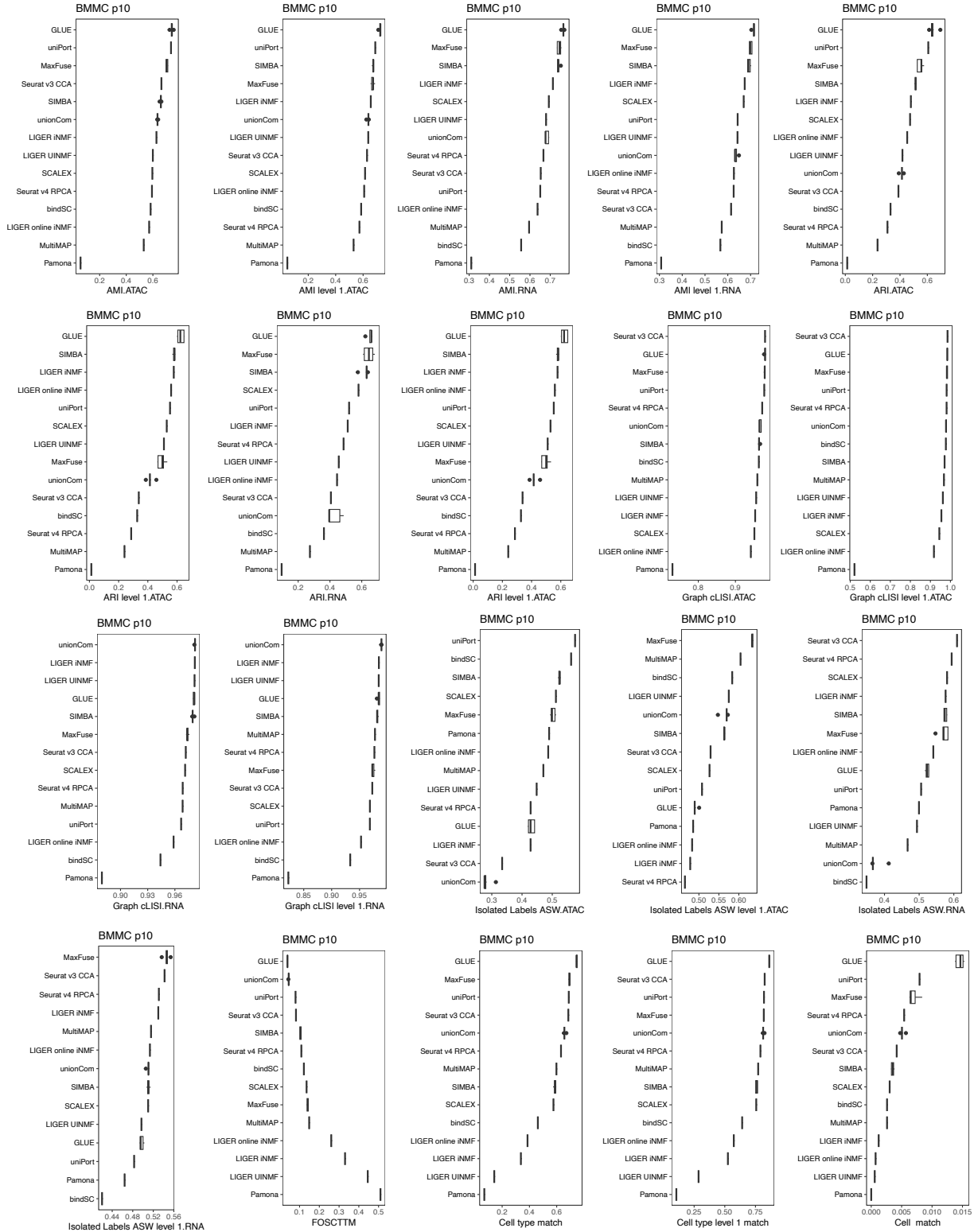
Supplementary Fig. 9 Boxplot of accuracy metrics for paired scRNA and ADT integration algorithms in other datasets except for BMMC CITE-seq datasets, corresponding to Fig. 3b and Extended Data Fig. 5b. The dataset names are labelled at the top of each boxplot. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



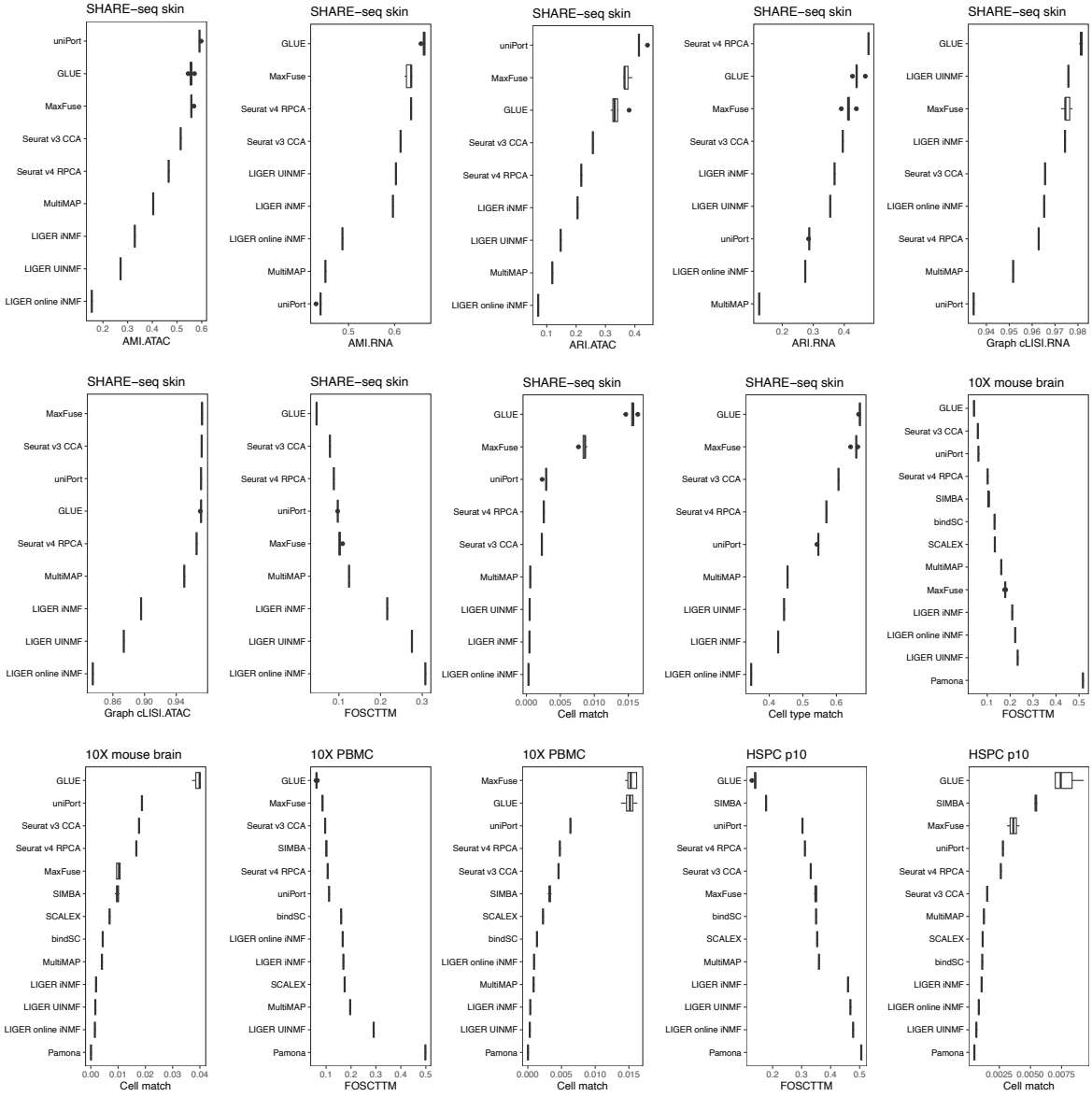
Supplementary Fig. 10. Boxplot of the biological conservation and batch removal metric values for all paired scRNA and ADT integration methods at all downsampling levels. This figure shows the metric scores distribution for simulations of **(a)** scRNA modality downsampling, **(b)** ADT modality downsampling and **(c)** both modalities downsampling. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



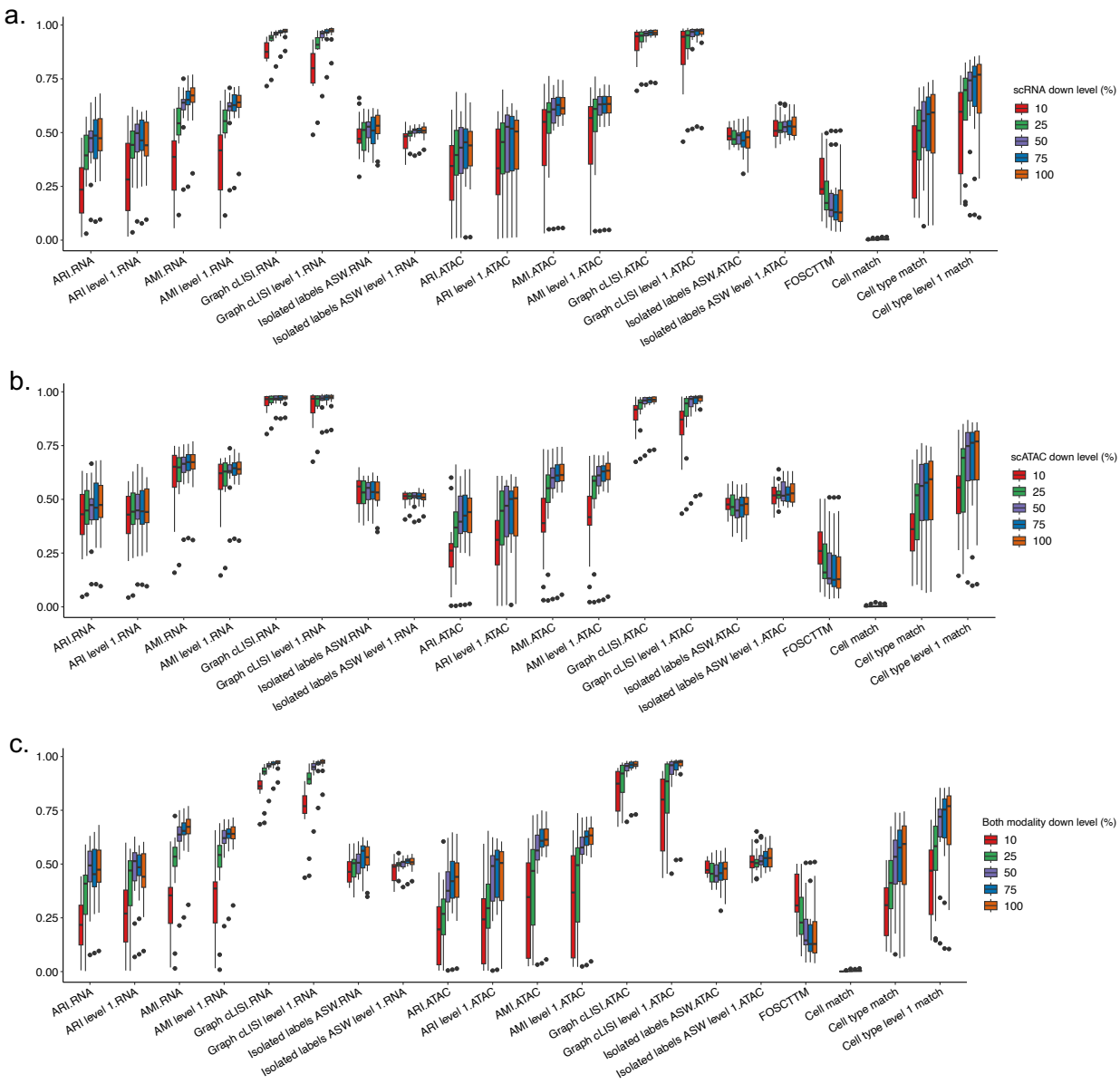
Supplementary Fig. 11. Details of the paired scRNA and ADT integration robustness metrics in Fig. 3d and Extended Data Fig. 5c. (**a-b**). Correlations between metric values and downsample gradients. The color and number in heatmaps shows Pearson's r correlation between (**a)** metric values or (**b**) average metric values with the downsampling gradients. (**c**) Detail metric values for all paired scRNA and ADT integration algorithms for BMMC p10 dataset at the 10% sequencing depth. (**d**) Detail downsample AUC values for all paired scRNA and ADT integration algorithms.



Supplementary Fig.12. Boxplot of accuracy metrics for unpaired scRNA and scATAC diagonal integration algorithms in BMMC Multiome p10 dataset, corresponding to Fig. 4c and Extended Data Fig. 6b. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 × IQR (interquartile range).

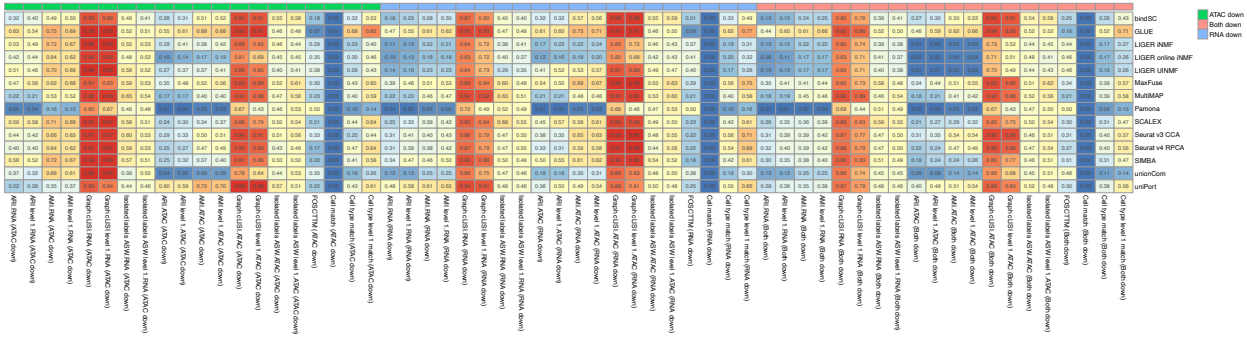


Supplementary Fig. 13. Boxplot of accuracy metrics for unpaired scRNA and scATAC diagonal integration algorithms in other datasets except for BMMC Multiome p10 dataset, corresponding to Fig. 4c and Extended Data Fig. 6b. The dataset names are labelled at the top of each boxplot. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).

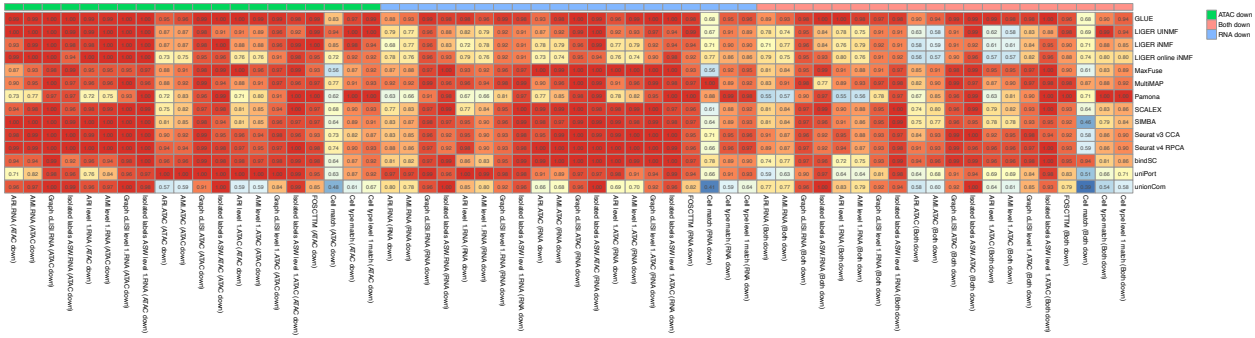


Supplementary Fig. 14. Boxplot of the embedding accuracy and cell alignment accuracy metric values for all unpaired scRNA and scATAC diagonal integration methods at all downsampling levels. This figure shows the metric scores distribution for simulations of **(a)** scRNA modality downsampling, **(b)** scATAC modality downsampling and **(c)** both modalities downsampling. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).

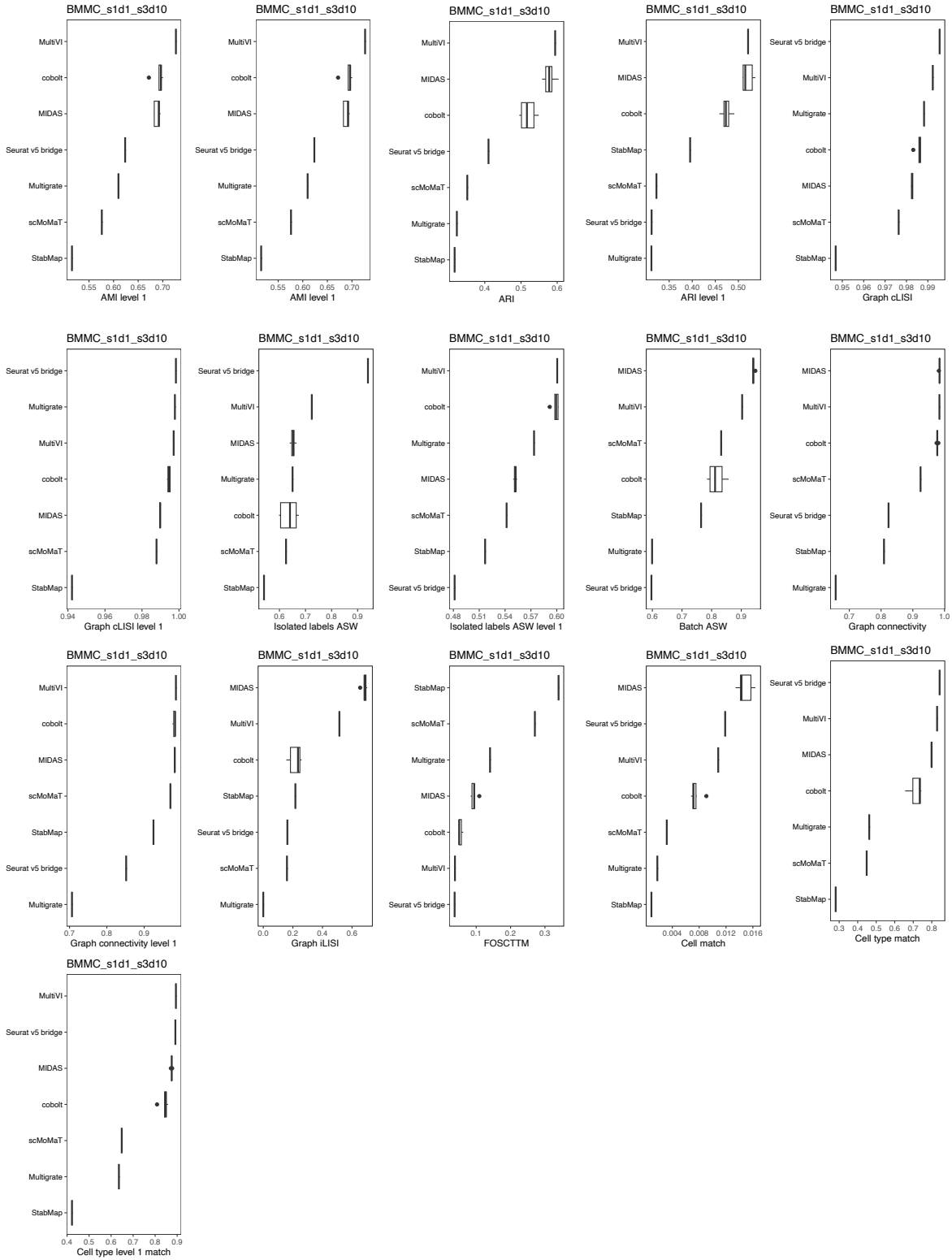
a.



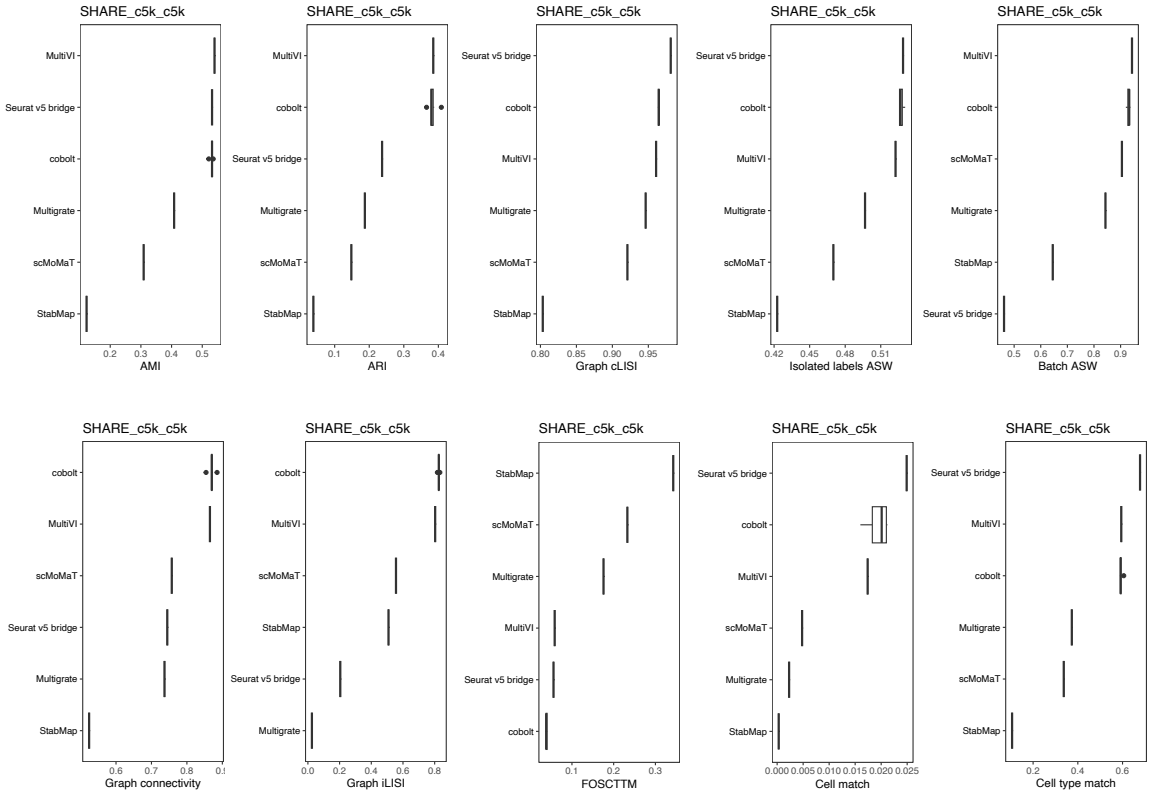
b.



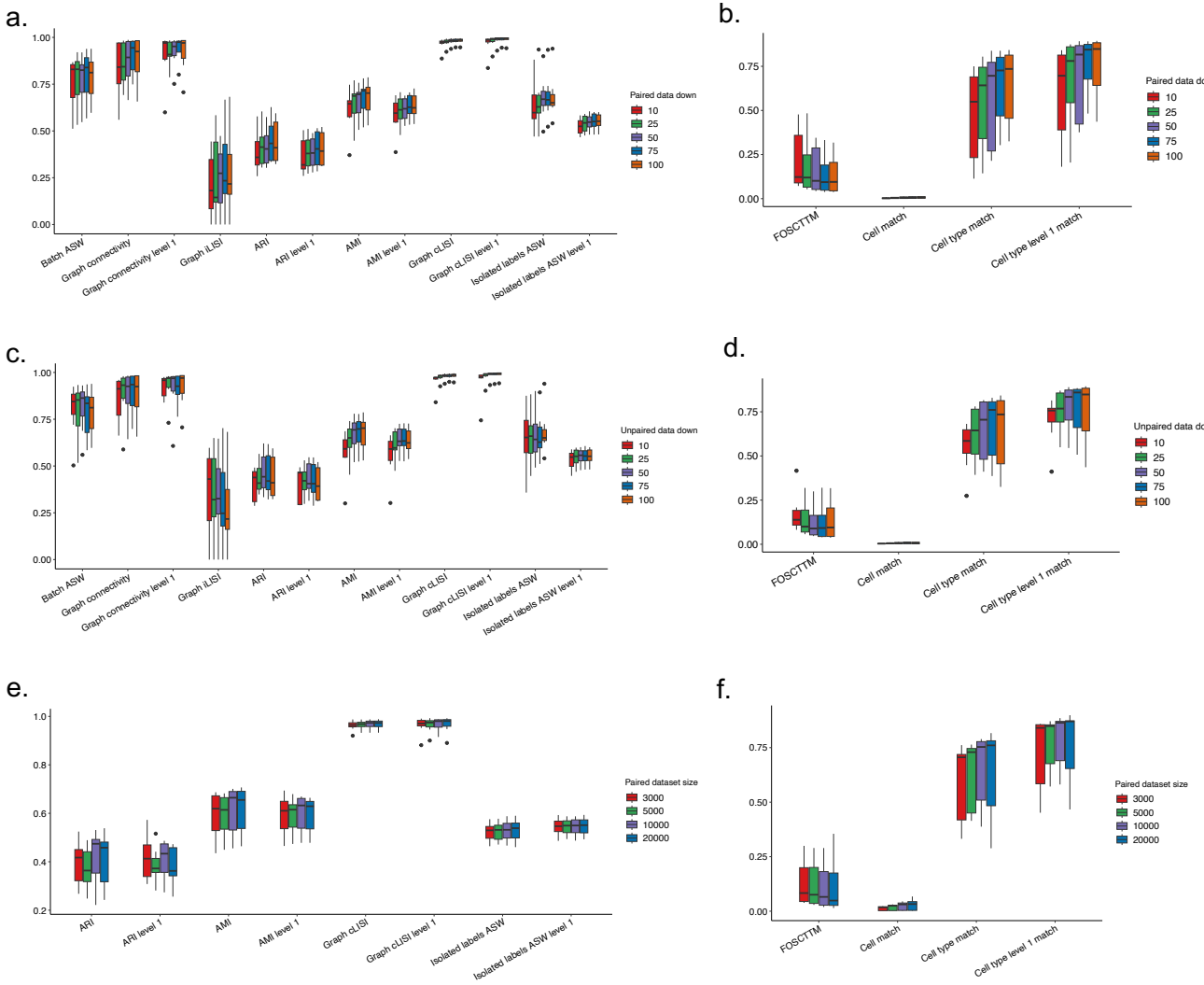
Supplementary Fig. 15. Details of the unpaired scRNA and scATAC diagonal robustness metrics in Fig. 4f and Extended Data Fig. 6c. (**a**) Detail metric values for all unpaired scRNA and scATAC diagonal integration algorithms for BMMC p10 dataset at the 10% sequencing depth. (**b**) Detail downsample AUC values for all unpaired scRNA and scATAC diagonal integration algorithms.



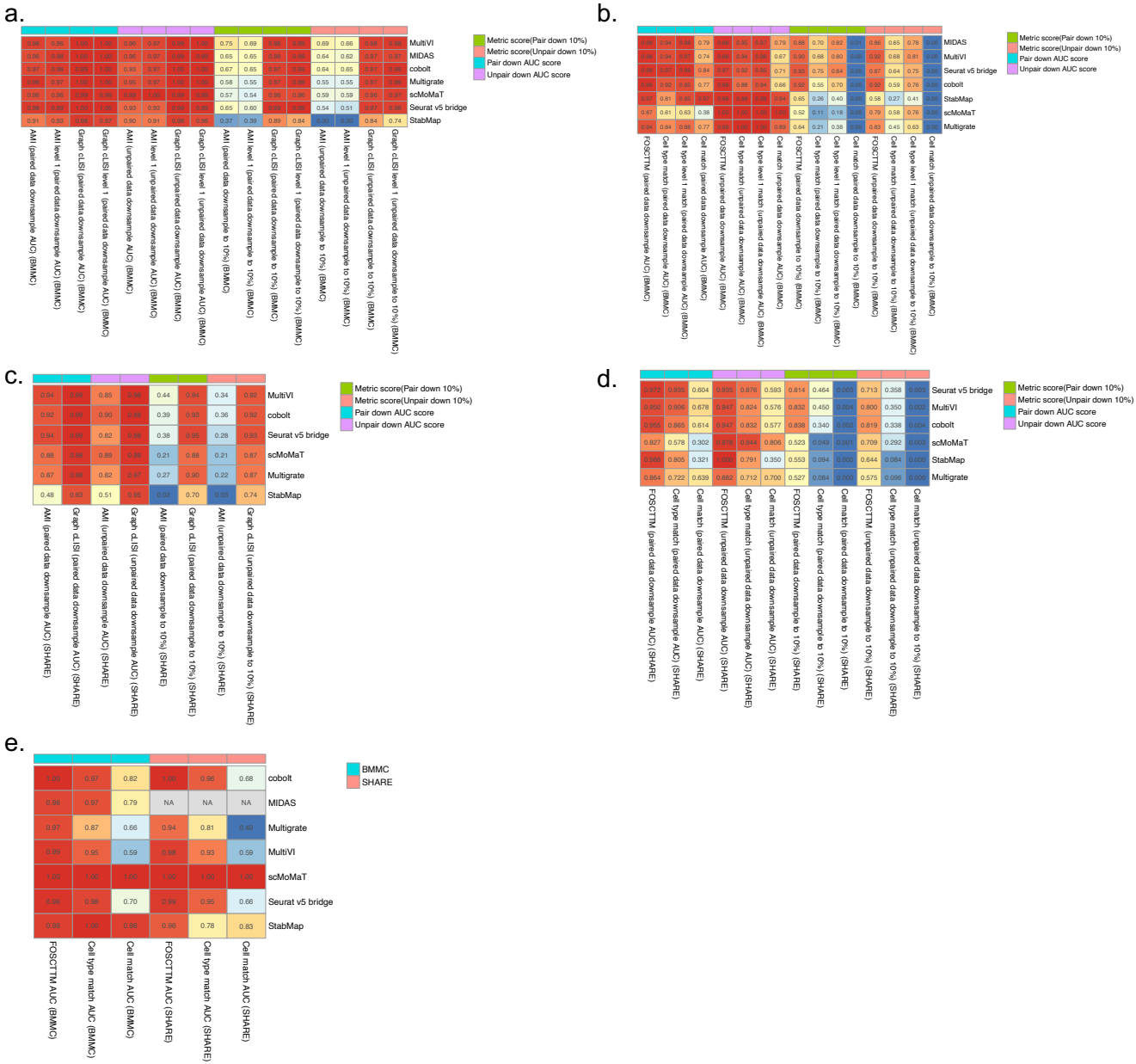
Supplementary Fig.16. Boxplot of the embedding accuracy and cell alignment accuracy metrics for unpaired scRNA and scATAC mosaic integration algorithms in BMMC Multiome s1d1_s3d10 dataset, corresponding to Fig. 5b and Extended Data Fig. 7b. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



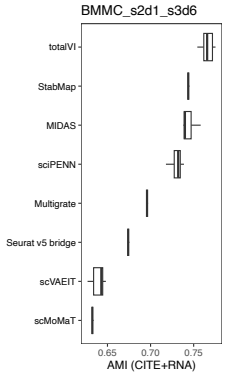
Supplementary Fig.17. Boxplot of the embedding accuracy and cell alignment accuracy metrics for unpaired scRNA and scATAC mosaic integration algorithms in SHARE-seq skin c5k_c5k dataset, corresponding to Fig. 5b and Extended Data Fig. 7b. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).

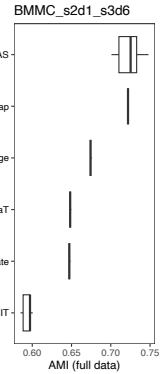


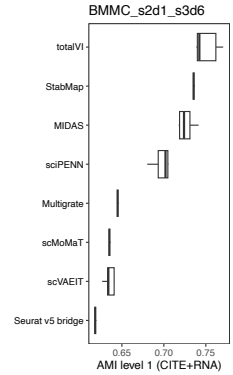
Supplementary Fig. 18. Boxplot of the embedding accuracy and cell alignment accuracy metric values for all unpaired scRNA and scATAC mosaic integration methods at (**a-b**) paired dataset downsampling levels, (**c-d**) unpaired dataset downsampling levels and (**e-f**) paired dataset sizes. The embedding accuracy and cell alignment accuracy metrics of different paired dataset sizes in (**e**) and (**f**) were calculated with unpaired cells only. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 × IQR (interquartile range).

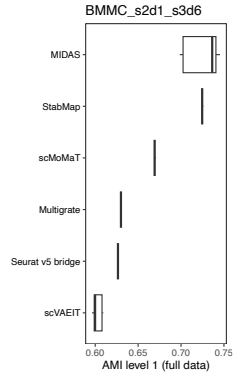


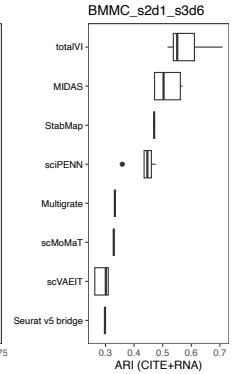
Supplementary Fig. 19. Details of the unpaired scRNA and scATAC mosaic integration robustness metrics in Fig. 5e. (**a-b**). Detail robustness scores of (**a**) embedding accuracy metrics and (**b**) cell alignment accuracy metrics for unpaired scRNA and scATAC mosaic integration algorithms in downsampling paired and unpaired multi-omic cells of BMMC Multiome s1d1_s3d10 dataset. (**c-d**). Detail robustness scores of embedding accuracy metrics and cell alignment accuracy metrics for all unpaired scRNA and scATAC mosaic integration algorithms in downsampling paired and unpaired multi-omic cells for SHARE-seq skin c5k_c5k dataset. (**e**) Detail AUCs of unpaired dataset cell alignment metrics in gradient paired dataset sizes for both BMMC and SHARE-seq datasets.

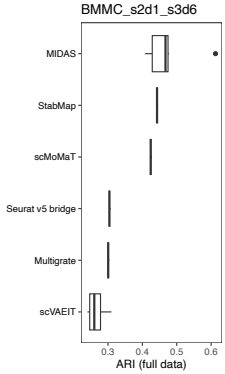

BMMC_s2d1_s3d6
totalVI
StabMap
MIDAS
sciPENN
Multigrate
Seurat v5 bridge
scVAEIT
scMoMaT
0.65
0.70
0.75
AMI (CITE+RNA)

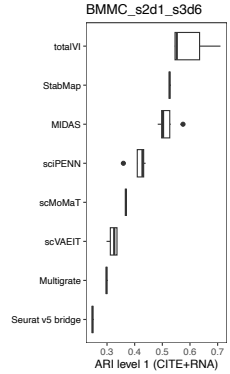

BMMC_s2d1_s3d6
MIDAS
StabMap
Seurat v5 bridge
scMoMaT
Multigrate
scVAEIT
0.60
0.65
0.70
0.75
AMI (full data)

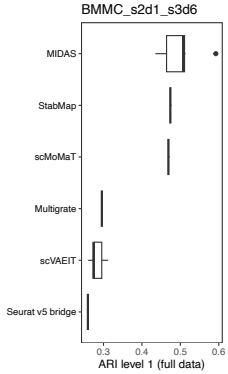

BMMC_s2d1_s3d6
totalVI
StabMap
MIDAS
sciPENN
Multigrate
scMoMaT
scVAEIT
Seurat v5 bridge
0.65
0.70
0.75
AMI level 1 (CITE+RNA)

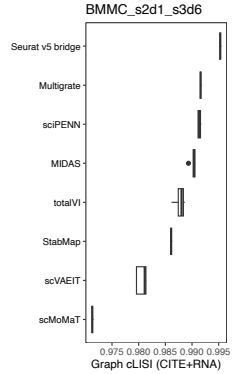

BMMC_s2d1_s3d6
MIDAS
StabMap
scMoMaT
Multigrate
Seurat v5 bridge
scVAEIT
0.60
0.65
0.70
0.75
AMI level 1 (full data)

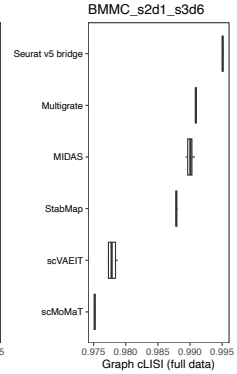

BMMC_s2d1_s3d6
totalVI
MIDAS
StabMap
sciPENN
Multigrate
scMoMaT
scVAEIT
Seurat v5 bridge
0.3
0.4
0.5
0.6
0.7
ARI (CITE+RNA)

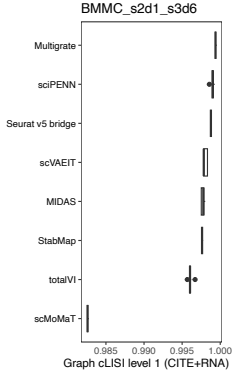

BMMC_s2d1_s3d6
MIDAS
StabMap
scMoMaT
Seurat v5 bridge
Multigrate
scVAEIT
0.3
0.4
0.5
0.6
ARI (full data)

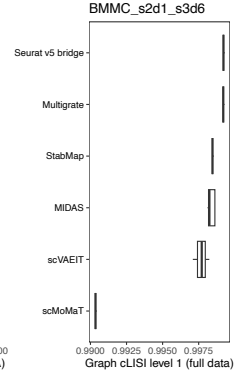

BMMC_s2d1_s3d6
totalVI
StabMap
MIDAS
sciPENN
scMoMaT
scVAEIT
Multigrate
Seurat v5 bridge
0.3
0.4
0.5
0.6
0.7
ARI level 1 (CITE+RNA)

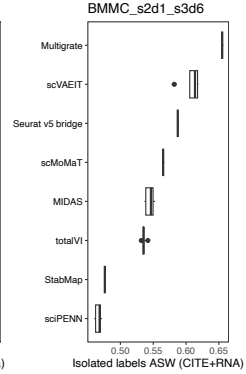

BMMC_s2d1_s3d6
MIDAS
StabMap
scMoMaT
Multigrate
scVAEIT
Seurat v5 bridge
0.3
0.4
0.5
0.6
ARI level 1 (full data)

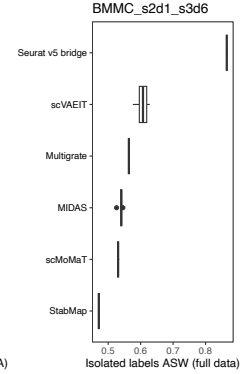

BMMC_s2d1_s3d6
Seurat v5 bridge
Multigrate
sciPENN
MIDAS
totalVI
StabMap
scVAEIT
scMoMaT
0.975
0.980
0.985
0.990
0.995
Graph cLISI (CITE+RNA)

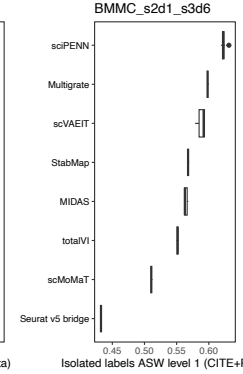

BMMC_s2d1_s3d6
Seurat v5 bridge
Multigrate
MIDAS
StabMap
scVAEIT
scMoMaT
0.975
0.980
0.985
0.990
0.995
Graph cLISI (full data)

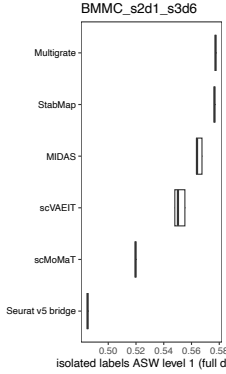

BMMC_s2d1_s3d6
Multigrate
sciPENN
Seurat v5 bridge
scVAEIT
MIDAS
StabMap
totalVI
scMoMaT
0.985
0.990
0.995
1.000
Graph cLISI level 1 (CITE+RNA)

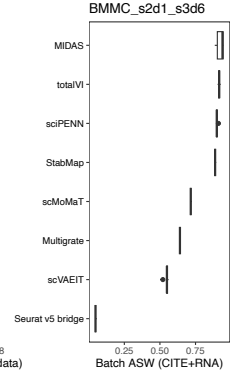

BMMC_s2d1_s3d6
Seurat v5 bridge
Multigrate
StabMap
MIDAS
scVAEIT
scMoMaT
0.9900
0.9925
0.9950
0.9975
Graph cLISI level 1 (full data)

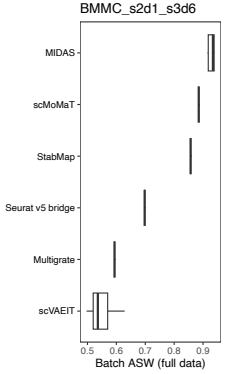

BMMC_s2d1_s3d6
Multigrate
scVAEIT
Seurat v5 bridge
scMoMaT
MIDAS
totalVI
StabMap
sciPENN
0.50
0.55
0.60
0.65
Isolated labels ASW (CITE+RNA)

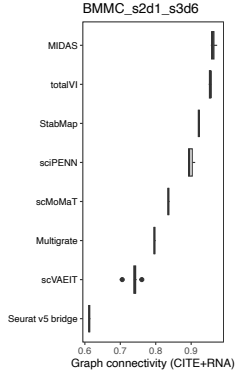

BMMC_s2d1_s3d6
Seurat v5 bridge
scVAEIT
Multigrate
MIDAS
scMoMaT
StabMap
0.5
0.6
0.7
0.8
Isolated labels ASW (full data)

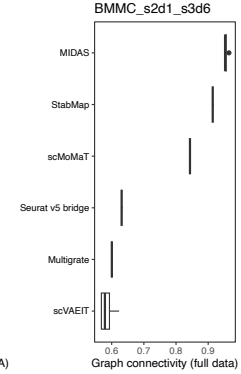

BMMC_s2d1_s3d6
sciPENN
Multigrate
scVAEIT
StabMap
MIDAS
totalVI
scMoMaT
Seurat v5 bridge
0.45
0.50
0.55
0.60
Isolated labels ASW level 1 (CITE+RNA)

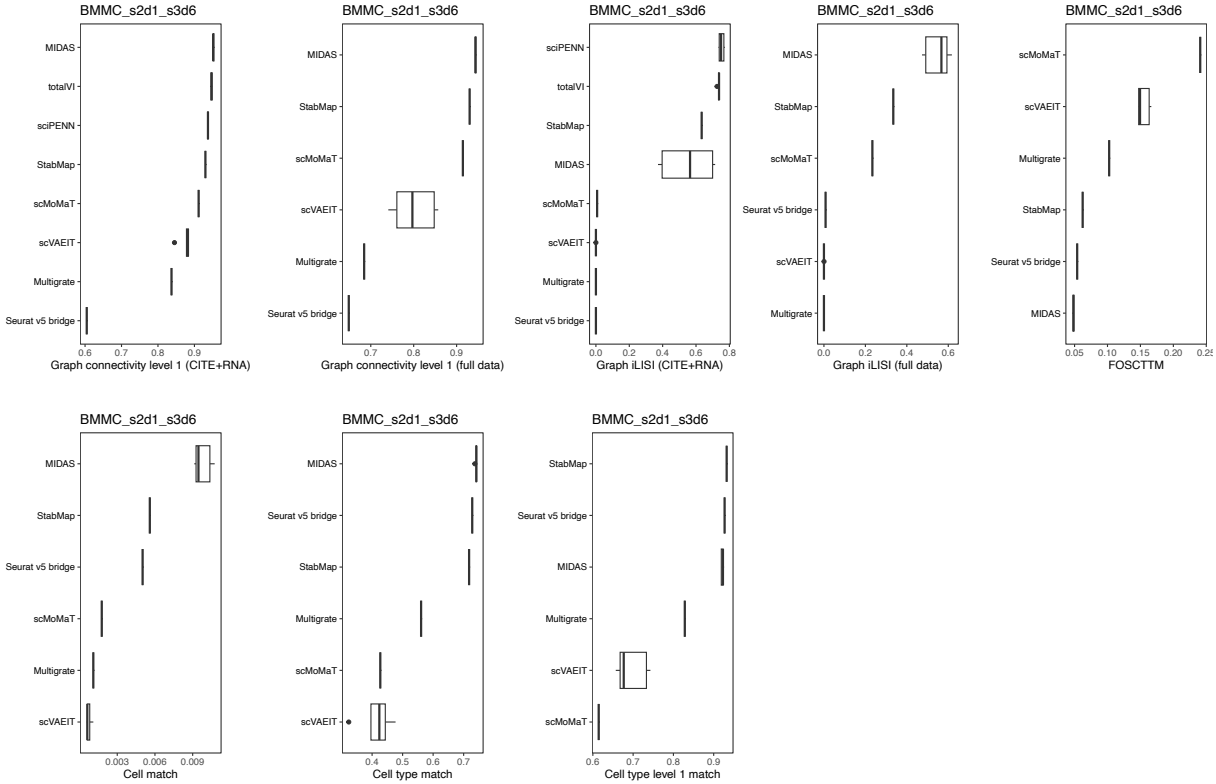

BMMC_s2d1_s3d6
Multigrate
StabMap
MIDAS
scVAEIT
scMoMaT
Seurat v5 bridge
0.50
0.52
0.54
0.56
0.58
isolated labels ASW level 1 (full data)


BMMC_s2d1_s3d6
MIDAS
totalVI
sciPENN
StabMap
scMoMaT
Multigrate
scVAEIT
Seurat v5 bridge
0.25
0.50
0.75
Batch ASW (CITE+RNA)

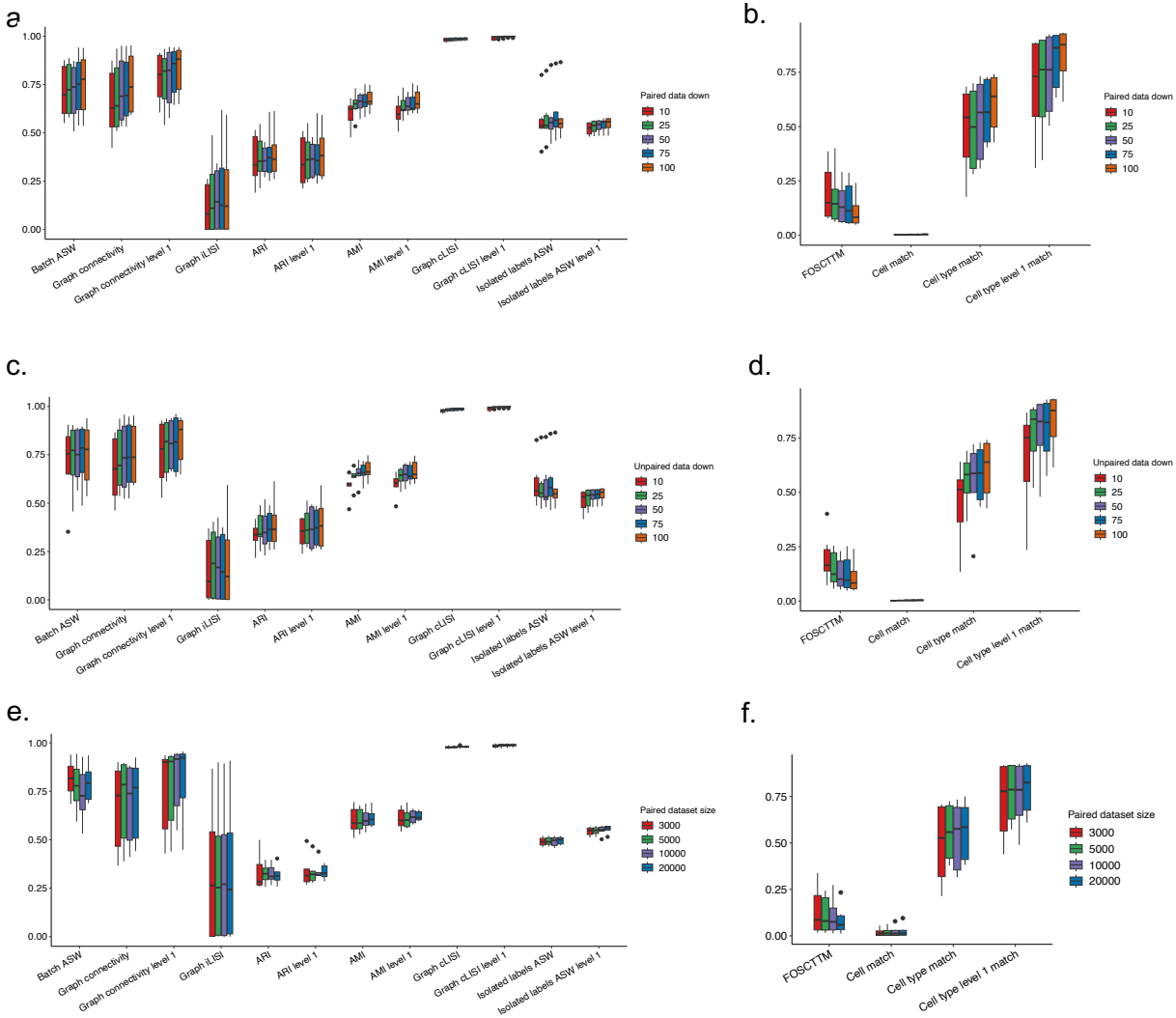

BMMC_s2d1_s3d6
MIDAS
scMoMaT
StabMap
Seurat v5 bridge
Multigrate
scVAEIT
0.5
0.6
0.7
0.8
0.9
Batch ASW (full data)


BMMC_s2d1_s3d6
MIDAS
totalVI
StabMap
sciPENN
scMoMaT
Multigrate
scVAEIT
Seurat v5 bridge
0.6
0.7
0.8
0.9
Graph connectivity (CITE+RNA)

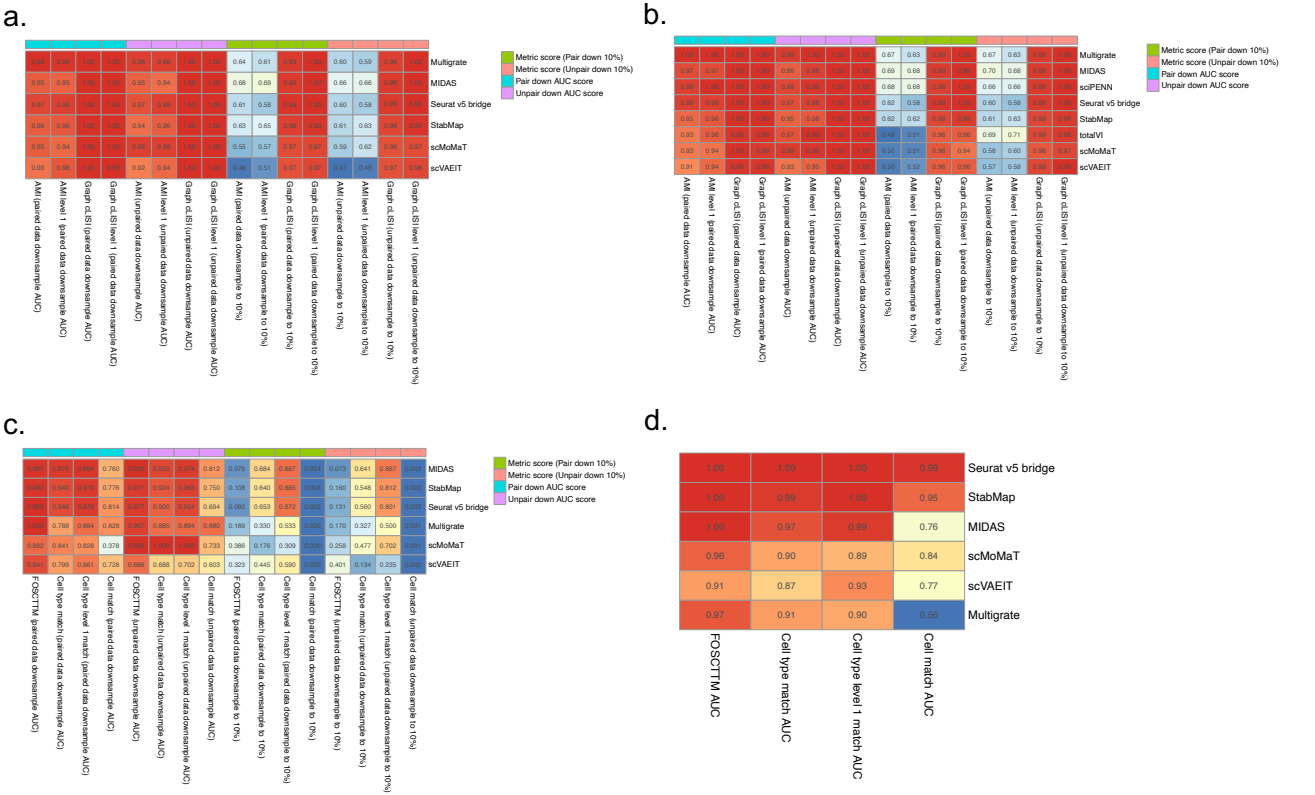

BMMC_s2d1_s3d6
MIDAS
StabMap
scMoMaT
Seurat v5 bridge
Multigrate
scVAEIT
0.6
0.7
0.8
0.9
Graph connectivity (full data)



Supplementary Fig. 20. Boxplot of the embedding accuracy and cell alignment accuracy metrics for unpaired scRNA and ADT mosaic integration algorithms in BMMC CITE-seq s2d1_s3d6 dataset, corresponding to Fig. 6b and Extended Data Fig. 8b. Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 $\times$ IQR (interquartile range).



Supplementary Fig. 21. Boxplot of the embedding accuracy and cell alignment accuracy metric values for all unpaired scRNA and ADT mosaic integration methods at (**a-b**) paired dataset downsampling levels, (**c-d**) unpaired dataset downsampling levels and (**e-f**) paired dataset sizes. The embedding accuracy and cell alignment accuracy metrics of different paired dataset size in (**e**) and (**f**) were calculated with unpaired cells only, Box plots show the median (centre line), the 25th and 75th percentiles (bounds of the box), and whiskers extend to 1.5 × IQR (interquartile range).



Supplementary Fig. 22. Details of the unpaired scRNA and ADT mosaic integration robustness metrics in Fig. 6c. (**a-b**). Detail robustness scores of embedding accuracy metrics of (**a**) full data and (**b**) CITE-seq and scRNA-seq cells in downsampling paired and unpaired multi-omic cells of BMMC CITE-seq dataset. (**c**) Detail robustness scores of cell alignment accuracy metrics cells in downsampling paired and unpaired multi-omic cells of BMMC s2d1_s3d6 dataset. (**d**) Detail AUCs of unpaired dataset cell alignment metrics in gradient paired dataset sizes of BMMC CITE-seq dataset.

Supplementary Notes：

Supplementary Note 1: Details of all benchmark methods

Here we outlined all details specific data preprocessing and parameter used in each Benchmarking single-cell multi-modal data integrations benchmark (SCMMIB) method. Other details of all benchmark methods are listed in **Supplementary Table 1.** The scripts of all benchmark integration methods are deposited in SCMMIB pipeline GitHub folder (https://github.com/bm2-lab/SCMMIB_pipeline/tree/main/benchmark_methods).

1. bindSC[1]

We evaluated bindSC for paired scRNA and ADT integration and unpaired scRNA and scATAC diagonal intergration tasks. For paired scRNA and ADT integration, we used default parameters in bindSC tutorial (https://htmlpreview.github.io/?https://github.com/KChen-lab/bindSC/blob/master/vignettes/CITE-seq/CITE_seq.html), and set the parameter *K* in BiCCA function as the smaller one in 15 and dimension of ADT data, as some datasets with too few ADT detected would run into error in dimension reduction step. We concatenated the output embedding of scRNA and ADT as joint embedding for bindSC. For unpaired scRNA and scATAC diagonal intergration, we took pre-computed gene activity matrix (Methods) and scRNA data as input, and used the recommended parameters from bindSC tutorial (https://htmlpreview.github.io/?https://github.com/KChen-lab/bindSC/blob/master/vignettes/A549/A549.html).

2. CiteFuse[2]

CiteFuse was evaluated for paired scRNA and ADT integration task. We used the default parameter in its' official tutorial (https://sydneybiox.github.io/CiteFuse/articles/CiteFuse.html) and generated the similarity network fusion (SNF) matrix as method graph output.

3. cobolt[3]

Cobolt was evaluated for both paired scRNA and scATAC integration task and unpaired scRNA and scATAC mosaic integration task. For both tasks, we used same parameters in the official tutorial (https://github.com/epurdom/cobolt/blob/master/docs/tutorial.ipynb).

4. DCCA[4]

DCCA was evaluated for paired scRNA and scATAC integration task. We followed all parameters and code used in the official tutorial (https://github.com/c

mzuo11/DCCA/wiki/Analysis-of-PBMC_3K-dataset-from-10X-Genomics-by-DCCA-model) and debugged the core functions "MVAE_cycleVAE.py" and "Cycle_model_missing.py", which forced to use two GPU device of certain names and raise error for device without exactly matched GPU hardware. We revised the script of GPU device as variable option in the revised version of source code and then reinstall the DCCA package.

5. DeepMAPS[5]

As reported in Methods section, we contacted with the author of DeepMAPS for GPU incompatibility in the GitHub issue (https://github.com/OSU-BMBL/deepmaps/issues/5). Due to its' incompatibility with our benchmark platform (NVIDIA 3090 and 1080Ti only), we did not test this method in our benchmark.

6. GCN-SC[6]

This method only provided some python files without any tutorial or introduction in GitHub (https://github.com/YuBinLab-QUST/GCN-SC). For this reason, we were not able to evaluate this method in accuracy and robustness sections.

7. GLUE[7]

GLUE was evaluated in unpaired scRNA and scATAC diagonal integration task. We ran this method following the official tutorial for version 0.2.0 (https://scglue.readthedocs.io/en/v0.2.1/), and took original scRNA and scATAC profiles as input.

8. Liger iNMF[8], Liger UiNMF[9], Liger online iNMF[10]

We evaluated three Liger functions in Liger package for unpaired scRNA and scATAC diagonal integration tasks. We followed the official tutorials of Liger iNMF (https://welch-lab.github.io/liger/articles/Integrating_scRNA_and_scATAC_data.html), Liger UiNMF (https://welch-lab.github.io/liger/articles/SNAREseq_walkthrough.html) and Liger online iNMF (https://welch-lab.github.io/liger/articles/online_iNMF_tutorial.html), and used same parameters in our benchmark.

9. MaxFuse[11]

MaxFuse was evaluated for both paired scRNA and ADT integration task and unpaired scRNA and scATAC diagonal integration task. We used the source code of MaxFuse in manuscript (https://github.com/shuxiaoc/maxfuse/tree/main/Archive) for both tasks. We followed parameters in the reproducibility code (https://github.com/shuxiaoc/maxfuse/blob/main/Archive/strong-link/10xpbmc/prep_pbmc.Rmd) for unpaired scRNA and scATAC diagonal integration task, and followi

ng the official tutorial (https://maxfuse.readthedocs.io/en/latest/citeseq_pbmc_evaluate.html) for paired scRNA and ADT integration task.

10. MOFA+[12], MEFISTO[13]

We evaluated two tools, MOFA+ and MEFISTO, in the MOFA package. MOFA+ was evaluated for paired scRNA and scATAC integration task and paired scRNA and ADT integration task. MEFISTO was evaluated for spatial CITE-seq integration task. We followed tutorials in muon package[14] (https://muon-tutorials.readthedocs.io/en/latest/) for both methods.

11. MIDAS[15]

MIDAS was evaluated for unpaired scRNA and scATAC mosaic integration task and unpaired scRNA and ADT mosaic integration task. As MIDAS uploaded the user executable code and official tutorial after AIP of our stage 1 manuscript (July 4th, 2024), we appended this algorithm after our stage 1 proposal with the permission from the editor. We followed the MIDAS official tutorial (https://scmidas.readthedocs.io/en/latest/) for both tasks.

12. Multigrate[16]

Multigrate was tested for paired scRNA and scATAC integration task, paired scRNA and ADT integration task, unpaired scRNA and scATAC mosaic integration task and unpaired scRNA and ADT mosaic integration task. We followed the official tutorials (https://multigrate.readthedocs.io/en/latest/tutorials.html) for all these tasks. We also evaluated the batch parameter by setting "*categorical_covariate_keys=['batch']*" in *setup_anndata* function.

13. MultiMAP[17]

MultiMAP was tested for unpaired scRNA and scATAC diagonal integration task and spatial CITE-seq integration task. For unpaired scRNA and scATAC diagonal integration task, we followed the official tutorial (https://nbviewer.org/github/Teichlab/MultiMAP/blob/master/examples/tutorial.ipynb). For spatial CITE-seq integration task, we followed the GitHub documentations for manifold alignment (https://github.com/Teichlab/MultiMAP) and aligned the spatial coordinates to CITE-seq manifold to generate joint reduction.

14. MultiVI[18], totalVI[19]

MultiVI and totalVI were two integration functions in scvi-tools[20]. We evaluated MultiVI for paired scRNA and scATAC integration task and unpaired scRNA and scATAC mosaic integration task. And we tested totalVI for paired scR

NA and ADT integration task and unpaired scRNA and ADT mosaic integratio n task. We ran the two methods following the tutorials of scvi-tools (https://do cs.scvi-tools.org/en/stable/index.html).

15. Pamona[21]

Pamona was evaluated for unpaired scRNA and scATAC diagonal integration t ask. We ran this method following the official tutorial (https://github.com/caokai 1073/Pamona/blob/master/example/Simulation1.ipynb) with default parameters.

16. SAILERX[22]

SAILERX was evaluated for paired scRNA and scATAC integration task. How ever, the input data format or demo was not available, and the author did not response to the unavailability of input data (https://github.com/uci-cbcl/SAILER X/issues/1). For this reason, we did not evaluate this method for accuracy and robustness metrics.

17. scAI[23]

scAI was evaluated for paired scRNA and scATAC integration task. We follow ed the parameter in official tutorial (https://htmlpreview.github.io/?https://github.c om/sqjin/scAI/blob/master/examples/walkthrough_A549dataset.html) for this metho d.

18. SCALEX[24]

SCALEX was evaluated for paired scRNA and ADT integration task and unpai red scRNA and scATAC diagonal integration task. We followed the official tut orial (https://scalex.readthedocs.io/en/latest/index.html) for both tasks.

19. scDEC[25]

scDEC was evaluated for paired scRNA and scATAC integration task. We foll owed the parameters in official tutorial for 10X PBMC dataset (https://github.c om/kimmo1019/scDEC/wiki/PBMC10k) for this method.

20. sciPENN[26]

sciPENN was evaluated for paired scRNA and ADT integration task and unpai red scRNA and ADT mosaic integration task. We followed the official tutorial in google drive (https://drive.google.com/drive/folders/1iY4s76UYNMFvF6v3XN4 JxD9gM77NIxoH) for both tasks.

21. scMCs[27]

scMCs was selected for paired scRNA and scATAC integration task. However, this method only provided the source code, and was not available for any tutorial or demo for user (https://www.sdu-idea.cn/codes.php?name=ScMCs).

22. scMDC[28]

scMDC was evaluated for paired scRNA and scATAC integration task and evaluated for paired scRNA and ADT integration task. We ran scMDC for both tasks with the default parameter introduced in the official GitHub folder (https://github.com/xianglin226/scMDC).

23. scMM[29]

scMM was evaluated for paired scRNA and scATAC integration task and evaluated for paired scRNA and ADT integration task. However, we encountered an error for exploding gradient in both tasks, which was also reported by other users (https://github.com/kodaim1115/scMM/issues/2), however, not responded by the author. For this reason, we did not evaluate this method for accuracy and robustness of both tasks.

24. scMoMaT[30]

scMoMaT was evaluated for unpaired scRNA and scATAC mosaic integration task and unpaired scRNA and ADT mosaic integration task. For both tasks, we followed the recommended parameter in demo of official GitHub (https://github.com/PeterZZQ/scMoMaT).

25. scMVAE[31]

scMVAE was evaluated for paired scRNA and scATAC integration task. The scMVAE should be appended with "__init__.py" for proper import in python script. Following a test code as demo (https://github.com/cmzuo11/scMVAE/blob/master/MVAE_test_Adbrain.py), we used same parameters and took GAM and scRNA as input of this method.

26. scMVP[32]

scMVP was evaluated for paired scRNA and scATAC integration task. We followed the parameters used in official tutorial (https://github.com/bm2-lab/scMVP/blob/master/demos/scMVP_tutorial.ipynb) for this method.

27. scVAEIT[33]

scVAEIT was evaluated for unpaired scRNA and scATAC mosaic integration task and unpaired scRNA and ADT mosaic integration task. For both tasks, we followed the tutorial provided by the author on Feb 22, 2024 in GitHub issue (https://github.com/jaydu1/scVAEIT/issues/6) for both tasks.

28. Seurat v3 CCA[34], Seurat v4 RPCA[34], Seurat v4 WNN[35], Seurat v5 bridge[36]

We evaluated four Seurat functions in Seurat package for our benchmark. Seurat v3 CCA and Seurat v4 RPCA were first released in version 3, and were evaluated for unpaired scRNA and scATAC diagonal integration task. However, we could only find RPCA integration function in Seurat version 4, and marked this method as Seurat v4 RPCA in our manuscript. Seurat v4 WNN was evaluated for paired scRNA and scATAC integration task and paired scRNA and ADT integration task. Seurat v5 bridge was evaluated for unpaired scRNA and scATAC mosaic integration task and unpaired scRNA and ADT mosaic integration task. We used same parameters as official tutorials for Seurat v3 CCA (https://satijalab.org/seurat/articles/integration_introduction.html), Seurat v4 RPCA (https://satijalab.org/seurat/articles/integration_rpca.html) and Seurat v4 WNN (https://satijalab.org/seurat/articles/weighted_nearest_neighbor_analysis). For Seurat v5 bridge, we used "*FindBridgeIntegrationAnchors*" function to replace "*FindBridgeTransferAnchors*" in official function document (https://satijalab.org/seurat/reference/findbridgeintegrationanchors), as we evaluated the integration accuracy rather than query accuracy in two mosaic integration benchmark tasks.

29. SIMBA[37]

SIMBA was evaluated for unpaired scRNA and scATAC diagonal integration task. We ran SIMBA for this task following the official tutorial (https://simba-bio.readthedocs.io/en/latest/multiome_10xpmbc10k_integration.html).

30. SpatialGlue[38]

SpatialGlue was evaluated of spatial CITE-seq integration task. As we used the datasets in SpatialGlue paper as our benchmark datasets, we ran this method with same scripts for same dataset in official tutorial (https://spatialglue-tutorials.readthedocs.io/en/latest/).

31. StabMap[39]

StabMap was evaluated for unpaired scRNA and scATAC mosaic integration task and unpaired scRNA and ADT mosaic integration task. We ran this method in unpaired scRNA and ADT mosaic integration task with recommended parameter in official tutorial (https://marionilab.github.io/StabMap/articles/stabMap_PBMC_Multiome.html). For unpaired scRNA and scATAC mosaic integration task, the filtering threshold in tutorial would lead to 0 features left for downsampling dataset, and the maximum features in StabMap is 1000 for all modalities, we set a fixed feature number of 600 for both scRNA and scATAC modalities.

32. unionCom[40]

unionCom was evaluated for unpaired scRNA and scATAC diagonal integration task. We followed parameters used in official tutorial (https://github.com/caokai1073/UnionCom/blob/master/Examples/Simulation1.ipynb) for this method.

33. uniPort[41]

uniPort was evaluated for paired scRNA and scATAC integration task and unp aired scRNA and scATAC diagonal integration task. We revised the parameter "chunk_size" in "batch_scale" function from default value of 20000 to 20001, as this method would raise a bug when the cell number in our scalability data set was divisible by the default "chunk_size" parameter. We also used lower le arning rate than the official tutorial (https://uniport.readthedocs.io/en/latest/), as t his method would encounter gradient explosion for datasets larger than 50,000 cells with default learning rate.

Supplementary Note 2: Summary of benchmark algorithms

For a comprehensive benchmarking analysis of multi-modal integration tools using different infrastructures, we collected a total of 40 multi-modal integration software packages encompassing 65 integration tasks. This collection included 17 paired scRNA and scATAC integration methods, 13 paired scRNA and ADT integration methods, 3 spatial multi-modal integration methods, 14 unpaired scRNA and scATAC diagonal integration methods, 8 CITE-seq mosaic integration methods and 8 multiome mosaic integration methods, whereas some of these tools were applicable for multiple multi-modal integrations. We executed using their default parameters as specified in public tutorials and methods detailed in their respective publications (details in Supplementary Note).

Among these benchmark algorithms, several algorithms were not executed for performance evaluations. Firstly, some tools were not available for our benchmark analysis due to the limited usability and a lack of source code maintenance. We were unable to execute scMCs[27] and GCN-SC[6] as we were unable to find any script or tutorial to guide us in using the algorithm or assist us with the source code. And SAILERX[22] did not provide the format of input data in the tutorial, and the hyperlink of example input data was not accessible. Additionally, We encountered an error when using scMM[29], an issue that had also been reported by us as well as other users on their GitHub page. However, ultimately, we received no response from the author. Secondly, some tools were restricted or not supported by our benchmark GPU platforms. DeepMAPS[5] was developed and tested on the NVIDIA A100 GPU, which, out of the box, does not support GeForce Ampere cards, like NVIDIA 3090 GPU or NVIDIA 1080Ti GPU used in our benchmark.

Supplementary Note 3: Details of all benchmark datasets

The characteristics of these datasets in the proposed benchmark study as follows:

BMMC datasets

The BMMC Multiome and CITE-seq datasets[42] contain nested batches of donors and experimental sites. The cell annotations were first generated for each modality and batch separately, and then manually annotated the cell subclusters inconsistent between the two modalities by evaluating which modality marker features more clearly described the specific cellular subpopulation[42]. The cell type annotations of BMMC datasets do not rely on joint embedding clustering or scRNA modality clustering, which largely ensures the independence of clustering results of all benchmark algorithms. For both Multiome and CITE-seq datasets, we grouped similar cell types and refine the annotations as broader cell types (Supplementary Table 3).

HSPC datasets

The HSPC Multiome and CITE-seq datasets[43] contain nested batches of HSPC cells collected from 4 samples and then cultured with 5 time points, which would be suitable for batch effect removal evaluation. The HSPC datasets annotated cells by scRNA cell clusters and their marker genes' gene expression using information from a previous study[44], which were not apply for evaluation with biological conservation metrics.

SHARE-seq dataset

The SHARE-seq mouse skin dataset[45] was annotated for scRNA and scATAC modalities separately, and then marked cell labels inconsistent between the two modalities as “mixed”. We used cells with consistent cell labels in two modalities as the ground truth for cell clustering, as these annotated cell labels would not favor the clustering results from joint embedding or scRNA embedding.

COVID-19 dataset

The COVID-19 PBMC CITE-seq dataset[46] is a large-scale dataset with 781,123 cells from 143 donors, which was available for batch removal and scalability evaluation. The COVID-19 dataset annotates cells by manually matching marker genes and surface protein expression to Leiden clustering results from the scRNA modality. For potential bias towards scRNA clustering-based results, we did not use the COVID-19 dataset for evaluation requiring ground truth cell labels.

10X PBMC dataset

The 10X PBMC dataset is merged from 10k PBMC Multiome RNA + ATAC dataset (https://www.10xgenomics.com/datasets/pbmc-from-a-healthy-donor-no-cell-sorting-10-k-1-standard-2-0-0) and 3k PBMC Multiome RNA + ATAC dataset (https://www.10xgenomics.com/datasets/pbmc-from-a-healthy-donor-no-cell-sorting-3-k-1-standard-2-0-0), including 15,021 cells in total. This dataset is available for evaluating batch removal between two datasets, as well as cell alignment accuracy evaluation.

10X mouse brain dataset

10X mouse brain dataset is a single cell Multiome RNA + ATAC dataset for brain coronal sections from one hemisphere of Alzheimer's disease mouse and wild type

mouse from 10X genomics website (https://www.10xgenomics.com/datasets/multiomic-integration-neuroscience-application-note-single-cell-multiome-rna-atac-alzheimers-disease-mouse-model-brain-coronal-sections-from-one-hemisphere-over-a-time-course-1-standard). We used two replicates for both wild type mouse and Alzheimer's disease mouse at time point of 2.5 months as our benchmark dataset. This dataset is available for batch removal evaluation and cell alignment accuracy evaluation.

Human WBC dataset

Human WBC dataset[34] is a human white blood cell (WBC) CITE-seq dataset with 24 samples from 8 donors and 3 time points, which was available for batch effect removal evaluation. This dataset was annotated by Seurat v4 package[34], which was not apply for evaluation with biological conservation metrics.

10X NSCLC dataset

10X NSCLC dataset is a CITE-seq dataset of non-small cell lung cancer (NSCLC) dissociated tumor cells from 10X Genomics website (https://www.10xgenomics.com/datasets/20k-mixture-of-nsclc-dtcs-from-7-donors-3-v3-1-with-intronic-reads-3-1-standard). This dataset includes 7 donors, and was available for batch removal evaluation.

10X kidney cancer dataset

10X kidney cancer dataset is a subset of dissociate tumor cells mixture dataset from 10X Genomics website (https://www.10xgenomics.com/datasets/40k-mixture-of-dissociated-tumor-cells-from-3-donors-stained-with-totalseqc-antibodies). We used two replicates of kidney cancer samples for batch removal evaluation.

Spatial multiomics datasets

We used three spatial CITE-seq datasets[47] for spatial multi-modal integration methods evaluation. Among the three spatial multi-omics datasets, we used the human lymph node dataset for biological conservation evaluation, as this dataset contains cell annotations by manually annotated tissue structures from histological images, which are irrelevant to joint embedding or scRNA embedding.

Scalability datasets

For scalability evaluation, we generated new simulation datasets (Supplementary Table 2) with a specific number of cells by either downsampling or upsampling from the BMMC Multiome and COVID-19 CITE-seq benchmark dataset. Since this is a scalability evaluation, the study would not be influenced by the proportion of cell type labels or cell batch information. Specifically, for simulation datasets smaller than the size of the original dataset, we randomly selected a specific number of cells from the original benchmark data as downsampling simulation data. For simulation datasets larger than the size of original data, we randomly sampled from the original datasets

for several times and merge these datasets as upsampling simulation datasets. To assess the robustness of the algorithms confronting data sparsity, we downsampled the total counts for each modality of each cell to 75%, 50%, 25% and 10% of the original counts, using the downsampleMatrix function in the scuttle package[48]. Further details regarding the benchmark datasets and associated tasks were listed in Supplementary Table 2.

Reference

1. Dou J*, et al.* Bi-order multimodal integration of single-cell data. *Genome Biology* **23**, 112 (2022).

2. Kim HJ, Lin Y, Geddes TA, Yang JYH, Yang P. CiteFuse enables multi-modal analysis of CITE-seq data. *Bioinformatics* **36**, 4137-4143 (2020).

3. Gong B, Zhou Y, Purdom E. Cobolt: integrative analysis of multimodal single-cell sequencing data. *Genome Biology* **22**, 351 (2021).

4. Zuo C, Dai H, Chen L. Deep cross-omics cycle attention model for joint analysis of single-cell multi-omics data. *Bioinformatics*, (2021).

5. Ma A*, et al.* Single-cell biological network inference using a heterogeneous graph transformer. *Nat Commun* **14**, 964 (2023).

6. Gao H, Zhang B, Liu L, Li S, Gao X, Yu B. A universal framework for single-cell multi-omics data integration with graph convolutional networks. *Brief Bioinform* **24**, (2023).

7. Cao Z-J, Gao G. Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. *Nature Biotechnology*, (2022).

8. Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, Macosko EZ. Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. *Cell* **177**, 1873-1887 e1817 (2019).

9. Kriebel AR, Welch JD. UINMF performs mosaic integration of single-cell multi-omic datasets using nonnegative matrix factorization. *Nature Communications* **13**, 780 (2022).

10. Gao C*, et al.* Iterative single-cell multi-omic integration using online learning. *Nat Biotechnol* **39**, 1000-1007 (2021).

11. Chen S*, et al.* Integration of spatial and single-cell data across modalities with weakly linked features. *Nat Biotechnol*, (2023).

12. Argelaguet R*, et al.* MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biology* **21**, 111 (2020).

13. Velten B*, et al.* Identifying temporal and spatial patterns of variation from multimodal data using MEFISTO. *Nat Methods* **19**, 179-186 (2022).

14. Bredikhin D, Kats I, Stegle O. MUON: multimodal omics analysis framework. *Genome Biol* **23**, 42 (2022).

15. He Z*, et al.* Mosaic integration and knowledge transfer of single-cell multimodal data with MIDAS. *Nat Biotechnol*, (2024).

16. Lotfollahi M, Litinetskaya A, Theis FJ. Multigrate: single-cell multi-omic data integration. *bioRxiv*, 2022.2003.2016.484643 (2022).

17. Jain MS*, et al.* MultiMAP: dimensionality reduction and integration of multimodal data. *Genome Biol* **22**, 346 (2021).

18. Ashuach T, Gabitto MI, Koodli RV, Saldi GA, Jordan MI, Yosef N. MultiVI: deep generative model for the integration of multimodal data. *Nat Methods*, (2023).

19. Gayoso A*, et al.* Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nature Methods* **18**, 272-282 (2021).

20. Gayoso A*, et al.* A Python library for probabilistic analysis of single-cell omics data. *Nat Biotechnol* **40**, 163-166 (2022).

21. Cao K, Hong Y, Wan L. Manifold alignment for heterogeneous single-cell multi-omics data integration using Pamona. *Bioinformatics* **38**, 211-219 (2021).

22. Cao Y*, et al.* Integrated analysis of multimodal single-cell data with structural similarity. *Nucleic Acids Res* **50**, e121 (2022).

23. Jin S, Zhang L, Nie Q. scAI: an unsupervised approach for the integrative analysis of parallel single-cell transcriptomic and epigenomic profiles. *Genome Biology* **21**, 25 (2020).

24. Xiong L, Tian K, Li Y, Zhang QC. Online single-cell data integration through projecting heterogeneous datasets into a common cell-embedding space. *bioRxiv*, 2021.2004.2006.438536 (2021).

25. Liu Q, Chen S, Jiang R, Wong WH. Simultaneous deep generative modelling and clustering of single-cell genomic data. *Nature Machine Intelligence* **3**, 536-544 (2021).

26. Lakkis J*, et al.* A multi-use deep learning method for CITE-seq and single-cell RNA-seq data integration with cell surface protein prediction and imputation. *Nat Mach Intell* **4**, 940-952 (2022).

27. Ren L, Wang J, Li Z, Li Q, Yu G. scMCs: a framework for single-cell multi-omics data integration and multiple clusterings. *Bioinformatics* **39**, (2023).

28. Lin X, Tian T, Wei Z, Hakonarson H. Clustering of single-cell multi-omics data with a multimodal deep learning method. *Nat Commun* **13**, 7705 (2022).

29. Minoura K, Abe K, Nam H, Nishikawa H, Shimamura T. A mixture-of-experts deep generative model for integrated analysis of single-cell multiomics data. *Cell Reports Methods* **1**, 100071 (2021).

30. Zhang Z*, et al.* scMoMaT jointly performs single cell mosaic integration and multi-modal bio-marker detection. *Nat Commun* **14**, 384 (2023).

31. Zuo C, Chen L. Deep-joint-learning analysis model of single cell transcriptome and open chromatin accessibility data. *Brief Bioinform* **22**, (2021).

32. Li G, *et al.* A deep generative model for multi-view profiling of single-cell RNA-seq and ATAC-seq data. *Genome Biology* **23**, 20 (2022).

33. Du JH, Cai Z, Roeder K. Robust probabilistic modeling for single-cell multimodal mosaic integration and imputation via scVAEIT. *Proc Natl Acad Sci U S A* **119**, e2214414119 (2022).

34. Stuart T, *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).

35. Hao Y, *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).

36. Hao Y, *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**, 293-304 (2024).

37. Chen H, Ryu J, Vinyard ME, Lerer A, Pinello L. SIMBA: single-cell embedding along with features. *Nat Methods*, (2023).

38. Long Y, *et al.* Deciphering spatial domains from spatial multi-omics with SpatialGlue. *Nature Methods* **21**, 1658-1667 (2024).

39. Ghazanfar S, Guibentif C, Marioni JC. Stabilized mosaic single-cell data integration using unshared features. *Nat Biotechnol* **42**, 284-292 (2024).

40. Cao K, Bai X, Hong Y, Wan L. Unsupervised topological alignment for single-cell multi-omics integration. *Bioinformatics* **36**, i48-i56 (2020).

41. Cao K, Gong Q, Hong Y, Wan L. A unified computational framework for single-cell data integration with optimal transport. *Nat Commun* **13**, 7419 (2022).

42. Luecken MD, *et al.* A sandbox for prediction and integration of DNA, RNA, and proteins in single cells. In: *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*) (2021).

43. Daniel Burkhardt RH, Malte Luecken, Andrew Benz, Peter Holderrieth, Jonathan Bloom, Christopher Lance, Ashley Chow. Open Problems - Multimodal Single-Cell Integration.). Kaggle (2022).

44. Velten L, *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nature Cell Biology* **19**, 271-281 (2017).

45. Ma S, *et al.* Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103-1116 e1120 (2020).

46. Stephenson E, *et al.* Single-cell multi-omics analysis of the immune response in COVID-19. *Nat Med* **27**, 904-916 (2021).

47. Ben-Chetrit N, *et al.* Integration of whole transcriptome spatial profiling with protein markers. *Nat Biotechnol* **41**, 788-793 (2023).

48. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179-1186 (2017).