

Table S1 Cross-modal generation identifies critical disease markers that would otherwise be missed. Top 5 differentially expressed proteins/-genes for ground-truth COVID-19 protein samples, cross-generated COVID-19 protein samples with perturbation, and ground-truth COVID-19 scRNA-seq samples, each compared to healthy controls. From the single-cell RNA-seq data alone, the cross-modal generation capability of scCross enables the identification of critical marker proteins that would otherwise remain undetected. Significant protein markers, such as CD38, were identified. Notably, CLEC12A, a top gene absent from the top protein biomarkers list identified with ground-truth single-cell protein data, was captured using our cross-modal generated protein data.

DE-Protein ground-truth	DE-Protein by scCross	DE-Gene ground-truth
CD38	CLEC12A	LY6E
PECAM1	CD38	IFI6
FCGR2A	PECAM1	TMSB10
LAMP1	FCGR2A	ISG15
CR1	CR1	IFITM3

Table S2 Detailed datasets information. Detailed information for datasets including Matched mouse cortex, Matched mouse atherosclerotic plaque immune cells, Matched mouse lymphonodus, Unmatched mouse cortex, Human cell atlas and COVID-19. Information like protocol, species, organ, cell number, access code/URL are included in this table.

Dataset	Protocol	Species	Organ	Cell number	Access code/URL
Matched mouse cortex [1, 2]	SNARE-seq	Mus musculus	Cortex	9,190	GSE126074
Matched mouse atherosclerotic plaque immune cells [3, 4]	10X Multiome	Mus musculus	Immune	5,984	GSE240753
Matched mouse lymphonodus [5, 6]	SNARE-seq	Mus musculus	Lymph	32,231	GSE140203
Unmatched mouse cortex scRNA-seq [7, 8]	Drop-seq	Mus musculus	Cortex	55,803	http://dropviz.org
Unmatched mouse cortex snmC-seq [9, 10]	snmC-seq	Mus musculus	Cortex	3,377	GSE97179
Unmatched mouse cortex scATAC-seq	10X ATAC	Mus musculus	Cortex	2,317	https://support.10xgenomics.com/single-cell-atac/datasets/1.1.0/atac.v1_adult_brain_fresh_5k
Human cell atlas scRNA-seq [11, 12]	sci-RNAseq3	Homo sapiens	Whole-organism	4,062,980	GSE156793
Human cell atlas scATAC-seq [13, 14]	sci-ATACseq3	Homo sapiens	Whole-organism	720,613	GSE149683
COVID-19 [15, 16]	CITE-seq	Homo sapiens	PBMC	647,366	https://covid19.cog.sanger.ac.uk/submissions/release1/haniffa21.processed.h5ad

References

- [1] Chen, S., Lake, B.B., Zhang, K.: High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nature biotechnology* **37**(12), 1452–1457 (2019)
- [2] Chen, S., Lake, B.B., Zhang, K.: High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. GEO. GSE12607 (2019)
- [3] Makhani, K., Yang, X., Dierick, F., Subramaniam, N., Gagnon, N., Ebrahimian, T., Wu, H., Ding, J., Mann, K.K.: Unveiling the impact of arsenic toxicity on immune cells in atherosclerotic plaques: Insights from single-cell multi-omics profiling. *bioRxiv*, 2023–11 (2023)
- [4] Makhani, K., Yang, X., Dierick, F., Subramaniam, N., Gagnon, N., Ebrahimian, T., Wu, H., Ding, J., Mann, K.K.: Unveiling the impact of arsenic toxicity on immune cells in atherosclerotic plaques: Insights from single-cell multi-omics profiling. GEO. GSE240753 (2023)
- [5] Ma, S., Zhang, B., LaFave, L.M., Earl, A.S., Chiang, Z., Hu, Y., Ding, J., Brack, A., Kartha, V.K., Tay, T., *et al.*: Chromatin potential identified by shared single-cell profiling of rna and chromatin. *Cell* **183**(4), 1103–1116 (2020)
- [6] Ma, S., Zhang, B., LaFave, L.M., Earl, A.S., Chiang, Z., Hu, Y., Ding, J., Brack, A., Kartha, V.K., Tay, T., *et al.*: Chromatin potential identified by shared single-cell profiling of rna and chromatin. GEO. GSE140203 (2020)
- [7] Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien, E., Baum, M., Bortolin, L., Wang, S., *et al.*: Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**(4), 1015–1030 (2018)
- [8] Saunders, A., Macosko, E.Z., Wysoker, A., Goldman, M., Krienen, F.M., de Rivera, H., Bien, E., Baum, M., Bortolin, L., Wang, S., *et al.*: Molecular diversity and specializations among the cells of the adult mouse brain. *Dropviz*. <http://dropviz.org> (2018)
- [9] Mulqueen, R.M., Pokholok, D., Norberg, S.J., Torkenczy, K.A., Fields, A.J., Sun, D., Sinnamon, J.R., Shendure, J., Trapnell, C., O’Roak, B.J., *et al.*: Highly scalable generation of dna methylation profiles in single cells. *Nature biotechnology* **36**(5), 428–431 (2018)
- [10] Mulqueen, R.M., Pokholok, D., Norberg, S.J., Torkenczy, K.A., Fields, A.J., Sun, D., Sinnamon, J.R., Shendure, J., Trapnell, C., O’Roak, B.J., *et*

- al.: Highly scalable generation of dna methylation profiles in single cells. GEO. GSE97179 (2018)
- [11] Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., *et al.*: A human cell atlas of fetal gene expression. *Science* **370**(6518), 7721 (2020)
 - [12] Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., *et al.*: A human cell atlas of fetal gene expression. GEO. GSE156793 (2020)
 - [13] Domcke, S., Hill, A.J., Daza, R.M., Cao, J., O'Day, D.R., Pliner, H.A., Aldinger, K.A., Pokholok, D., Zhang, F., Milbank, J.H., *et al.*: A human cell atlas of fetal chromatin accessibility. *Science* **370**(6518), 7612 (2020)
 - [14] Domcke, S., Hill, A.J., Daza, R.M., Cao, J., O'Day, D.R., Pliner, H.A., Aldinger, K.A., Pokholok, D., Zhang, F., Milbank, J.H., *et al.*: A human cell atlas of fetal chromatin accessibility. GEO. GSE149683 (2020)
 - [15] Stephenson, E., Reynolds, G., Botting, R.A., Calero-Nieto, F.J., Morgan, M.D., Tuong, Z.K., Bach, K., Sungnak, W., Worlock, K.B., Yoshida, M., *et al.*: Single-cell multi-omics analysis of the immune response in covid-19. *Nature medicine* **27**(5), 904–916 (2021)
 - [16] Stephenson, E., Reynolds, G., Botting, R.A., Calero-Nieto, F.J., Morgan, M.D., Tuong, Z.K., Bach, K., Sungnak, W., Worlock, K.B., Yoshida, M., *et al.*: Single-cell multi-omics analysis of the immune response in covid-19. Sanger. <https://covid19.cog.sanger.ac.uk/submissions/release1/haniffa21.processed.h5ad> (2021)