Supplemental References

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36:** 411–420. doi:10.1038/nbt.4096

Hunter JD. 2007. Matplotlib: a 2D graphics environment. *Comput Sci Eng* **9:** 90–95. doi:10.1109/MCSE.2007.55

Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. 2014. UpSet: visualization of intersecting sets. *IEEE Trans Vis Comput Graph* **20:** 1983–1992. doi:10.1109/TVCG.2014.2346248

McInnes L, Healy J, Melville J. 2018. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv:1802.03426 [stat.ML].

McKinney W. 2010. Data structures for statistical computing in Python. In *Proceedings of the Ninth Python in Science Conference*, Austin, TX (ed. van der Walt S, Millman J). doi:10.25080/Majora-92bf1922-00a

Oliphant TE. 2006. *Guide to numPy*. Trelgol Publishing, USA.

Oliphant TE. 2007. Python for scientific computing. *Comput Sci Eng* **9:** 10–20. doi:10.1109/MCSE.2007.58

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Duborg V, et al. 2011. Scikit-learn: machine learning in Python. *J Mach Learn Res* **12:** 2825–2830.

R Core Team. 2020. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

Waskom M, Botvinnik O, Hobson P, Cole JB, Halchenko Y, Hoyer S, Miles A, Augsperger T, Yarkoni T, Megies T, et al. 2014. Seaborn: v0.5.0 (November 2014). Zenodo. https://zenodo.org/record/4012680#.YNFQAm5Ok8Y

Wolf FA, Angerer P, Theis FJ. 2018. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19:** 15. doi:10.1186/s13059-017-1382-0