
HLab-PubMon

Release 0.1.0

Huang Lab @ HKU

Jun 29, 2023

MAIN

1	Journal list	3
2	Guidelines	5
2.1	Read in depth	5
2.2	Pub 2022-2023	7
2.3	Pub 2021-2022	11
2.4	Pub 2020-2021	20
2.5	Best3 in semester1-2020	43

HLab PubMon: Publication monitoring in Huang Lab @ HKU.

This repository aims to keep track of our reading habits on recent publications.

Codes: <https://github.com/StatBiomed/HLab-pubmon>

JOURNAL LIST

Topics are mainly about but not limited to computational biology, bioinformatics, systems biology, biostatistics, machine learning, artificial intelligence.

Here are some journals that we often find papers interesting to many of us:

- Nature: <https://www.nature.com>
- Science: <https://www.sciencemag.org>
- PNAS: <https://www.pnas.org>
- Genome Biology: <https://genomebiology.biomedcentral.com>
- Genome Research: <https://genome.cshlp.org>
- Bioinformatics: <https://academic.oup.com/bioinformatics>
- Plos comp bio: <https://journals.plos.org/ploscompbiol/>
- Nucleic Acids Research: <https://academic.oup.com/nar>
- Nature biotech: <https://www.nature.com/nbt/>
- Nature methods: <https://www.nature.com/nmeth/>
- Nature genetics: <https://www.nature.com/ng/>
- Nature communications: <https://www.nature.com/ncomms/>
- Conferences in ML, e.g., UAI, NeurIPS ...
- and more ...

GUIDELINES

The motivation is that everyone in the lab has a habit of reading, reasoning and sharing, which hopefully will keep us inspired and excited in science.

The guidelines here only aim to help us make better use of the limited time:

- Please make it within 10 min per person (up to 15min), including discussions.
- Introduce the background sufficiently, so everyone can understand what the question is.
- Don't need to go through the whole paper. Just mention 1 to 3 points that you want to share. It can also be questions that puzzle you.

Ideally, you will find this paper sharing is somewhat like a cake tasting party. So, bring the nice new cake you find and share why you like it, in an effective way.

2.1 Read in depth

2.1.1 2020-09-14

We read CHISEL paper for single-cell copy number variation calling [zaccaria2020characterizing] in detail today.

2.1.2 2020-10-11

We read variational auto-encoder paper [kingma2013auto] in detail today, with brief discussion on its single cell application scVI [lopez2018deep].

2.1.3 2021-05-17

We read three papers: RPCI [liu2021robust], Cell-ID: [cortal2021gene], scPred [alquicira2019scpred], all about linear projection of single-cell transcriptome to lower dimensional space with singular value decomposition (SVD), the implementation technique for Principal component analysis (PCA).

We also referred more details on PCA and SVD by reading the following book chapters:

- SVD quick reading: chapter 2.7-2.8 (4 pages): https://www.deeplearningbook.org/contents/linear_algebra.html
- SVD Longer reading: chapter 4.5: <https://mml-book.github.io/book/mml-book.pdf>
- PCA quick reading: chapter 12.1 <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>

2.1.4 2022-01-13

We read the Chapter 8 OptimizatiOn, PML book 1. <https://probml.github.io/pml-book/book1.html>

Rongting lead the reading of the foundation part, Mingze lead the discussion of the specific methods (SGD and the rest), and Chen C introduce the EM details.

2.1.5 2022-02-16

We read the Chapter 13-14 deep nueral networks(Introduction), PML book 1. <https://probml.github.io/pml-book/book1.html>

Xianjie, Qiaochen and Ruiyan lead the reading.

2.1.6 2022-03-02

We read the Chapter 14 VAE, PML book 1. <https://probml.github.io/pml-book/book1.html>

Fangxin and weizhong lead the reading.

2.1.7 References

1. Simone Zaccaria and Benjamin J Raphael. Characterizing allele-and haplotype-specific copy numbers in single cells with chisel. *Nature biotechnology*, pages 1–8, 2020. URL: <https://www.nature.com/articles/s41587-020-0661-6>.
2. Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv*, 2013. URL: <https://arxiv.org/abs/1312.6114>.
3. Romain Lopez, Jeffrey Regier, Michael B Cole, Michael I Jordan, and Nir Yosef. Deep generative modeling for single-cell transcriptomics. *Nature methods*, 15(12):1053–1058, 2018. URL: <https://www.nature.com/articles/s41592-018-0229-2>.
4. Yang Liu, Tao Wang, Bin Zhou, and Deyou Zheng. Robust integration of multiple single-cell RNA sequencing datasets using a single reference space. *Nature biotechnology*, pages 1–8, 2021. URL: <https://www.nature.com/articles/s41587-021-00859-x>.
5. Akira Cortal, Loredana Martignetti, Emmanuelle Six, and Antonio Rausell. Gene signature extraction and cell identity recognition at the single-cell level with Cell-ID. *Nature Biotechnology*, pages 1–8, 2021. URL: <https://www.nature.com/articles/s41587-021-00896-6>.
6. Jose Alquicira-Hernandez, Anuja Sathe, Hanlee P Ji, Quan Nguyen, and Joseph E Powell. scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome biology*, 20(1):1–17, 2019. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1862-5>.

2.2 Pub 2022-2023

2.2.1 2023-07-12

2.2.2 2023-06-21

1. Almut S Eisele, Marcel Tarbier, Alexia A Dormann, Vicent Pelechano, and David M Suter. Barcode-free prediction of cell lineages from scRNA-seq datasets. *bioRxiv*, pages 2022–09, 2022. URL: <https://www.biorxiv.org/content/10.1101/2022.09.20.508646v1>.

2.2.3 2023-05-04

1. Anna Arutyunyan, Kenny Roberts, Kevin Troulé, Frederick CK Wong, Megan A Sheridan, Ilia Kats, Luz Garcia-Alonso, Britta Velten, Regina Hoo, Elias R Ruiz-Morales, and others. Spatial multiomics map of trophoblast development in early pregnancy. *Nature*, 616(7955):143–151, 2023.
2. Isaac D Lutz, Shunzhi Wang, Christoffer Norn, Alexis Courbet, Andrew J Borst, Yan Ting Zhao, Annie Dosey, Longxing Cao, Jinwei Xu, Elizabeth M Leaf, and others. Top-down design of protein architectures with reinforcement learning. *Science*, 380(6642):266–273, 2023.
3. Steffen Schneider, Jin Hwa Lee, and Mackenzie Weygandt Mathis. Learnable latent embeddings for joint behavioural and neural analysis. *Nature*, pages 1–9, 2023. URL: <https://www.nature.com/articles/s41586-023-06031-6>.

2.2.4 2023-03-15

1. Gryte Satas, Simone Zaccaria, Geoffrey Mon, and Benjamin J Raphael. Scarlet: single-cell tumor phylogeny inference with copy-number constrained mutation losses. *Cell systems*, 10(4):323–332, 2020. URL: <https://pubmed.ncbi.nlm.nih.gov/31744515/>.
2. Dinithi Sumanaweera, Chenqu Suo, Daniele Muraro, Emma Dann, Krzysztof Polanski, Alexander S. Steemers, Jong-Eun Park, Bianca Dumitrascu, and Sarah A. Teichmann. Gene-level alignment of single cell trajectories informs the progression of in vitro t cell differentiation. *bioRxiv*, 2023. URL: <https://www.biorxiv.org/content/early/2023/03/10/2023.03.08.531713>, arXiv:<https://www.biorxiv.org/content/early/2023/03/10/2023.03.08.531713.full.pdf>, doi:10.1101/2023.03.08.531713.
3. Ya Cui, Frederick J Arnold, Fanglue Peng, Dan Wang, Jason Sheng Li, Sebastian Michels, Eric J Wagner, Albert R La Spada, and Wei Li. Alternative polyadenylation transcriptome-wide association study identifies APA-linked susceptibility genes in brain disorders. *Nature Communications*, 14(1):583, 2023. URL: <https://www.nature.com/articles/s41467-023-36311-8#Sec2>.
4. Anjun Ma, Xiaoying Wang, Jingxian Li, Cankun Wang, Tong Xiao, Yuntao Liu, Hao Cheng, Juexin Wang, Yang Li, Yuzhou Chang, and others. Single-cell biological network inference using a heterogeneous graph transformer. *Nature Communications*, 14(1):964, 2023.

2.2.5 2023-03-01

1. Brian L Hie, Kevin K Yang, and Peter S Kim. Evolutionary velocity with protein language models predicts evolutionary dynamics of diverse proteins. *Cell Systems*, 13(4):274–285, 2022.

2.2.6 2023-02-15

1. Mohammad Lotfollahi, Sergei Rybakov, Karin Hrovatin, Soroor Hediye-Zadeh, Carlos Talavera-López, Alexander V Misharin, and Fabian J Theis. Biologically informed deep learning to query gene programs in single-cell atlases. *Nature Cell Biology*, pages 1–14, 2023. URL: <https://www.nature.com/articles/s41556-022-01072-x>.
2. Cheng-kai Shiao, Lina Lu, Rachel Kieser, Kazutaka Fukumura, Timothy Pan, Hsiao-Yun Lin, Jie Yang, Eric L Tong, GaHyun Lee, Yuanqing Yan, and others. Delineating genotypes and phenotypes of individual cells from long-read single cell transcriptomes. *bioRxiv*, pages 2023–01, 2023. URL: <https://www.biorxiv.org/content/10.1101/2023.01.24.525264v2>.
3. Anna Lyubetskaya, Brian Rabe, Andrew Fisher, Anne Lewin, Isaac Neuhaus, Constance Brett, Todd Brett, Ethel Pereira, Ryan Golhar, Sami Kebede, and others. Assessment of spatial transcriptomics for oncology discovery. *Cell Reports Methods*, 2(11):100340, 2022.

2.2.7 2023-02-01

1. Ziqi Zhang, Haoran Sun, Ragunathan Mariappan, Xi Chen, Xinyu Chen, Mika S Jain, Mirjana Efremova, Sarah A Teichmann, Vaibhav Rajan, and Xiuwei Zhang. Scmomat jointly performs single cell mosaic integration and multi-modal bio-marker detection. *Nature Communications*, 14(1):384, 2023. URL: <https://www.nature.com/articles/s41467-023-36066-2>.
2. Fulong Yu, Liam D Cato, Chen Weng, L Alexander Liggett, Soyoung Jeon, Keren Xu, Charleston WK Chiang, Joseph L Wiemels, Jonathan S Weissman, Adam J de Smith, and others. Variant to function mapping at single-cell resolution through network propagation. *Nature Biotechnology*, pages 1–10, 2022. URL: <https://www.nature.com/articles/s41587-022-01341-y>.
3. Raquel García-Pérez, Jose Miguel Ramirez, Aida Ripoll-Cladellas, Ruben Chazarra-Gil, Winona Oliveros, Oleksandra Soldatkina, Mattia Bosio, Paul Joris Rognon, Salvador Capella-Gutierrez, Miquel Calvo, and others. The landscape of expression and alternative splicing variation across human traits. *Cell Genomics*, 2023. URL: <https://www.sciencedirect.com/science/article/pii/S2666979X22002075>.
4. Alexander P Wu, Rohit Singh, and Bonnie Berger. Granger causal inference on dags identifies genomic loci regulating transcription. In *International Conference on Learning Representations*. 2022.
5. Junjie Huang, Yanchao Xu, Yunfan Xue, Yue Huang, Xu Li, Xiaohui Chen, Yao Xu, Dongxiang Zhang, Peng Zhang, Junbo Zhao, and others. Identification of potent antimicrobial peptides via a machine-learning pipeline that mines the entire space of peptide sequences. *Nature Biomedical Engineering*, pages 1–14, 2023.

2.2.8 2023-01-11

1. Julia Joung, Sai Ma, Tristan Tay, Kathryn R. Geiger-Schuller, Paul C. Kirchgatterer, Vanessa K. Verdine, Baolin Guo, Mario A. Arias-Garcia, William E. Allen, Ankita Singh, Olena Kuksenko, Omar O. Abudayyeh, Jonathan S. Gootenberg, Zhanyan Fu, Rhiannon K. Macrae, Jason D. Buenrostro, Aviv Regev, and Feng Zhang. A transcription factor atlas of directed differentiation. *Cell*, 186(1):209–229.e26, jan 2023. URL: <https://doi.org/10.1016%2Fj.cell.2022.11.026>, doi:10.1016/j.cell.2022.11.026.
2. Kaelan J Brennan, Melanie Weilert, Sabrina Krueger, Anusri Pampari, Hsiao-Yun Liu, Ally WH Yang, Timothy R Hughes, Christine A Rushlow, Anshul Kundaje, and Julia Zeitlinger. Chromatin accessibility is a two-tier process regulated by transcription factor pioneering and enhancer activation. *bioRxiv*, 2022.
3. Dongze He, Charlotte Sonesson, and Rob Patro. Understanding and evaluating ambiguity in single-cell and single-nucleus rna-sequencing. *bioRxiv*, 2023. URL: <https://www.biorxiv.org/content/early/2023/01/04/2023.01.04.522742>, arXiv:<https://www.biorxiv.org/content/early/2023/01/04/2023.01.04.522742.full.pdf>, doi:10.1101/2023.01.04.522742.

2.2.9 2022-12-07

2.2.10 2022-11-23

1. Edwin Fong, Chris Holmes, and Stephen G Walker. Martingale posterior distributions. *arXiv preprint arXiv:2103.15671*, 2021.
2. Dylan M Cable, Evan Murray, Vignesh Shanmugam, Simon Zhang, Luli S Zou, Michael Diao, Haiqi Chen, Evan Z Macosko, Rafael A Irizarry, and Fei Chen. Cell type-specific inference of differential expression in spatial transcriptomics. *Nature methods*, 19(9):1076–1087, 2022. URL: <https://www.nature.com/articles/s41592-022-01575-3>.
3. Lei Xiong, Kang Tian, Yuzhe Li, and Qiangfeng Cliff Zhang. Online single-cell data integration through projecting heterogeneous datasets into a common cell-embedding space. *bioRxiv*, 2021. URL: <https://www.nature.com/articles/s41467-022-33758-z>.

2.2.11 2022-11-09

1. Yuxuan Hu, Jiazhen Rong, Runzhi Xie, Yafei Xu, Jacqueline Peng, Lin Gao, and Kai Tan. Learning predictive models of tissue cellular neighborhoods from cell phenotypes with graph pooling. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/11/06/2022.11.06.515344>, arXiv:<https://www.biorxiv.org/content/early/2022/11/06/2022.11.06.515344.full.pdf>, doi:10.1101/2022.11.06.515344.
2. Qian Qin, Eli Bingham, Gioele La Manno, David M Langenau, and Luca Pinello. Pyro-velocity: probabilistic rna velocity inference from single-cell data. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/10.1101/2022.09.12.507691v2>.
3. Qing Zhou, Guannan Kang, Peiyong Jiang, Rong Qiao, WK Jacky Lam, Stephanie CY Yu, Mary-Jane L Ma, Lu Ji, Suk Hang Cheng, Wanxia Gai, and others. Epigenetic analysis of cell-free dna by fragmentomic profiling. *Proceedings of the National Academy of Sciences*, 119(44):e2209852119, 2022.

2.2.12 2022-10-26

1. Gherman Novakovsky, Nick Dexter, Maxwell W Libbrecht, Wyeth W Wasserman, and Sara Mostafavi. Obtaining genetics insights from deep learning via explainable artificial intelligence. *Nature Reviews Genetics*, pages 1–13, 2022.
2. Zhaoyang Liu, Dongqing Sun, and Chenfei Wang. Evaluation of cell-cell interaction methods by integrating single-cell rna sequencing data with spatial information. *Genome biology*, 23(1):1–38, 2022.
3. Alexander J Tarashansky, Yuan Xue, Pengyang Li, Stephen R Quake, and Bo Wang. Self-assembling manifolds in single-cell rna sequencing data. *Elife*, 8:e48994, 2019.
4. Yuqiong Hu, Zhenhuan Jiang, Kexuan Chen, Zhangxian Zhou, Xin Zhou, Yan Wang, Jingwei Yang, Bo Zhang, Lu Wen, and Fuchou Tang. Scnanoatac-seq: a long-read single-cell atac sequencing method to detect chromatin accessibility and genetic variants simultaneously within an individual cell. *Cell Research*, pages 1–4, 2022. URL: <https://www.nature.com/articles/s41422-022-00730-x>.

2.2.13 2022-10-12

1. Kun Wang, Liangzhen Hou, Zhaolian Lu, Xin Wang, Zhike Zi, Weiwei Zhai, Xionglei He, Christina Curtis, Da Zhou, and Zheng Hu. Cell division history encodes directional information of fate transitions. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/10.1101/2022.10.06.511094v2>.
2. Wei Wei, Katherine R Schon, Greg Elgar, Andrea Orioli, Melanie Tanguy, Adam Giess, Marc Tischkowitz, Mark J Caulfield, and Patrick F Chinnery. Nuclear-embedded mitochondrial dna sequences in 66,083 human genomes. *Nature*, pages 1–10, 2022.

2.2.14 2022-09-28

1. Grace Hui Ting Yeo, Sachit D Saksena, and David K Gifford. Generative modeling of single-cell time series with prescient enables prediction of cell trajectories with interventions. *Nature communications*, 12(1):1–12, 2021. URL: <https://www.nature.com/articles/s41467-021-23518-w#Sec10>.
2. Teng Gao, Ruslan Soldatov, Hirak Sarkar, Adam Kurkiewicz, Evan Biederstedt, Po-Ru Loh, and Peter V Kharchenko. Haplotype-aware analysis of somatic copy number variations from single-cell transcriptomes. *Nature Biotechnology*, pages 1–10, 2022. URL: <https://www.nature.com/articles/s41587-022-01468-y>.

2.2.15 2022-09-14

1. Qi Liu, Chih-Yuan Hsu, and Yu Shyr. Scalable and model-free detection of spatial patterns and colocalization. *Genome Research*, 2022. URL: <https://genome.cshlp.org/content/early/2022/09/09/gr.276851.122.abstract>.
2. Chao Zhang, Renchao Chen, and Yi Zhang. Accurate inference of genome-wide spatial expression with ispatial. *Science Advances*, 2022. URL: <https://www.science.org/doi/10.1126/sciadv.abq0990>.
3. Han Yuan and David R Kelley. Scbasset: sequence-based modeling of single-cell atac-seq using convolutional neural networks. *Nature Methods*, pages 1–9, 2022. URL: <https://www.nature.com/articles/s41592-022-01562-8>.
4. Francisco X Galdos, Sidra Xu, William R Goodyer, Lauren Duan, Yuhsin V Huang, Soah Lee, Han Zhu, Carissa Lee, Nicholas Wei, Daniel Lee, and others. Devcellpy is a machine learning-enabled pipeline for automated annotation of complex multilayered single-cell transcriptomic data. *Nature communications*, 13(1):1–20, 2022. URL: <https://www.nature.com/articles/s41467-022-33045-x>.

2.3 Pub 2021-2022

2.3.1 2022-07-20

1. Zhiyi Qin, Peter Stoilov, Xuegong Zhang, and Yi Xing. Seastar: systematic evaluation of alternative transcription start sites in rna. *Nucleic Acids Research*, 46(8):e45–e45, 2018. URL: <https://academic.oup.com/nar/article/46/8/e45/4931254#116446609>.
2. Yunlong Ma, Fei Qiu, Chunyu Deng, Jingjing Li, Yukuan Huang, Zeyi Wu, Yijun Zhou, Yaru Zhang, Yichun Xiong, Yinghao Yao, and others. Integrating single-cell sequencing data with gwas summary statistics reveals cd16+ monocytes and memory cd8+ t cells involved in severe covid-19. *Genome medicine*, 14(1):1–21, 2022. URL: <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-022-01021-1#Sec17>.
3. Haoyu Chen, Marta Pelizzola, and Andreas Futschik. Haplotype based testing for a better understanding of the selective architecture. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/07/18/2022.07.18.500395>, arXiv:<https://www.biorxiv.org/content/early/2022/07/18/2022.07.18.500395.full.pdf>, doi:10.1101/2022.07.18.500395.
4. Hong Seo Lim and Peng Qiu. Quantifying cell-type-specific differences of single-cell datasets using umap and shap. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/07/18/2022.07.15.500285>, arXiv:<https://www.biorxiv.org/content/early/2022/07/18/2022.07.15.500285.full.pdf>, doi:10.1101/2022.07.15.500285.
5. Xiaojie Qiu, Yan Zhang, Jorge D Martin-Rufino, Chen Weng, Shayan Hosseinzadeh, Dian Yang, Angela N Pogson, Marco Y Hein, Kyung Hoi Joseph Min, Li Wang, and others. Mapping transcriptomic vector fields of single cells. *Cell*, 185(4):690–711, 2022. URL: <https://www.sciencedirect.com/science/article/pii/S0092867421015774>.
6. Cynthia A Kalita and Alexander Gusev. Decaf: a novel method to identify cell-type specific regulatory variants and their role in cancer risk. *Genome Biology*, 23(1):1–22, 2022. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-022-02708-9>.

2.3.2 2022-06-22

1. Atul Deshpande, Melanie Loth, Dimitrios N. Sidiropoulos, Shuming Zhang, Long Yuan, Alexander Bell, Qingfeng Zhu, Won Jin Ho, Cesar Santa-Maria, Daniele Gilkes, Stephen R. Williams, Cedric R. Uyttingco, Jennifer Chew, Andrej Hartnett, Zachary W. Bent, Alexander V. Favorov, Aleksander S. Popel, Mark Yarchoan, Lei Zheng, Elizabeth M. Jaffee, Robert Anders, Ludmila Danilova, Genevieve Stein-Otextquoteright Brien, Luciane T. Kagohara, and Elana J. Fertig. Uncovering the spatial landscape of molecular interactions within the tumor microenvironment through latent spaces. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/06/02/2022.06.02.490672>, arXiv:<https://www.biorxiv.org/content/early/2022/06/02/2022.06.02.490672.full.pdf>, doi:10.1101/2022.06.02.490672.
2. Shaolong Cao, Jennifer R Wang, Shuangxi Ji, Peng Yang, Yaoyi Dai, Shuai Guo, Matthew D Montierth, John Paul Shen, Xiao Zhao, Jingxiao Chen, and others. Estimation of tumor cell total mrna expression in 15 cancer types predicts disease progression. *Nature Biotechnology*, pages 1–10, 2022. URL: <https://www.nature.com/articles/s41587-022-01342-x>.
3. Maxwell Sherman, others, and Bonnie Berger. Genome-wide mapping of somatic mutation rates uncovers drivers of cancer. *Nature Biotechnology*, pages 1–10, 2022. URL: <https://www.nature.com/articles/s41587-022-01353-8>.
4. Fulong Yu, Liam D Cato, Chen Weng, L Alexander Liggett, Soyoung Jeon, Keren Xu, Charleston WK Chiang, Joseph L Wiemels, Jonathan S Weissman, Adam J de Smith, and others. Variant to function mapping at single-

cell resolution through network propagation. *Nature Biotechnology*, pages 1–10, 2022. URL: <https://www.nature.com/articles/s41587-022-01341-y>.

5. Pascal Notin, José Miguel Hernández-Lobato, and Yarin Gal. Improving black-box optimization in vae latent space using decoder uncertainty. *Advances in Neural Information Processing Systems*, 34:802–814, 2021.
6. Christopher D Steele, Ammal Abbasi, SM Islam, Amy L Bowes, Azhar Khandekar, Kerstin Haase, Shadi Hames-Fathi, Dolapo Ajayi, Annelien Verfaillie, Pawan Dhama, and others. Signatures of copy number alterations in human cancer. *Nature*, pages 1–8, 2022. URL: <https://www.nature.com/articles/s41586-022-04738-6>.
7. Taro Leo Saito, Shin-ichi Hashimoto, Sam Guoping Gu, J Jason Morton, Michael Stadler, Thomas Blumenthal, Andrew Fire, and Shinichi Morishita. The transcription start site landscape of *c. elegans*. *Genome research*, 23(8):1348–1361, 2013. URL: <https://genome.cshlp.org/content/23/8/1348.short>.

2.3.3 2022-06-08

1. Brendan F Miller, Feiyang Huang, Lyla Atta, Arpan Sahoo, and Jean Fan. Reference-free cell type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data. *Nature communications*, 13(1):1–13, 2022.
2. Peiyao Zhao, Jiaqiang Zhu, Ying Ma, and Xiang Zhou. Modeling zero inflation is not necessary for spatial transcriptomics. *Genome Biology*, 23(1):1–19, 2022.
3. Dian Yang, Matthew G Jones, Santiago Naranjo, William M Rideout III, Kyung Hoi Joseph Min, Raymond Ho, Wei Wu, Joseph M Replogle, Jennifer L Page, Jeffrey J Quinn, and others. Lineage tracing reveals the phylogenetics, plasticity, and paths of tumor evolution. *Cell*, 185(11):1905–1923, 2022.
4. Anja Nusser, Jeremy B Swann, Brigitte Krauth, Dagmar Diekhoff, Lesly Calderon, Christiane Happe, Dominic Grün, Thomas Boehm, and others. Developmental dynamics of two bipotent thymic epithelial progenitor types. *Nature*, pages 1–7, 2022.

2.3.4 2022-05-18

1. Kasper Karlsson, Moritz Przybilla, Hang Xu, Eran Kotler, Kremena Karagyzova, Alexandra Sockell, Katherine Liu, Amanda Mah, Yuan-Hung Lo, Bingxin Lu, Kathleen E. Houlihan, Aziz Khan, Zhicheng Ma, Carlos J. Suarez, Christopher P. Barnes, Calvin J. Kuo, and Christina Curtis. Experimental evolution in tp53 deficient gastric organoids recapitulates tumorigenesis. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/04/10/2022.04.09.487529>, doi:10.1101/2022.04.09.487529.
2. Rebecca Elyanow, Ron Zeira, Max Land, and Benjamin J Raphael. Starch: copy number and clone inference from spatial transcriptomics data. *Physical Biology*, 18(3):035001, 2021.
3. Andrew Erickson, Emelie Berglund, Mengxiao He, Maja Marklund, Reza Mirzazadeh, Niklas Schultz, Ludvig Bergenstråhle, Linda Kvastad, Alma Andersson, Joseph Bergenstråhle, and others. The spatial landscape of clonal somatic mutations in benign and malignant tissue. *bioRxiv*, 2021.
4. Eli N Weinstein, Alan N Amin, Will S Grathwohl, Daniel Kassler, Jean Disset, and Debora Marks. Optimal design of stochastic dna synthesis protocols based on generative sequence models. In *International Conference on Artificial Intelligence and Statistics*, 7450–7482. PMLR, 2022. URL: <https://www.biorxiv.org/content/10.1101/2021.10.28.466307v1>.
5. Aparna Nathan, Samira Asgari, Kazuyoshi Ishigaki, Cristian Valencia, Tiffany Amariuta, Yang Luo, Jessica I Beynor, Yuriy Baglaenko, Sara Suliman, Alkes L Price, and others. Single-cell eqtl models reveal dynamic t cell state dependence of disease loci. *Nature*, 606(7912):120–128, 2022. URL: <https://www.nature.com/articles/s41586-022-04713-1>.

6. Zhi-Jie Cao and Ge Gao. Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. *Nature Biotechnology*, pages 1–9, 2022. URL: <https://www.nature.com/articles/s41587-022-01284-4>.

2.3.5 2022-05-04

1. Carlos F Buen Abad Najar, Prakruthi Burra, Nir Yosef, and Liana F Lareau. Identifying cell-state associated alternative splicing events and their co-regulation. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.07.23.453605v2>.
2. Mislav Balunovic, Anian Ruoss, and Martin Vechev. Fair normalizing flows. In *International Conference on Learning Representations*. 2022. URL: <https://openreview.net/forum?id=BrFIKuxrZE>.

2.3.6 2022-04-13

1. Tianyu Zhu, Jacklyn Liu, Stephan Beck, Sun Pan, David Capper, Matt Lechner, Chrissie Thirlwell, Charles E Breeze, and Andrew E Teschendorff. A pan-tissue dna methylation atlas enables in silico decomposition of human tissue methylomes at cell-type resolution. *Nature methods*, 19(3):296–306, 2022. URL: <https://www.nature.com/articles/s41592-022-01412-7>.
2. Andrew E Teschendorff, Tianyu Zhu, Charles E Breeze, and Stephan Beck. Episcor: cell type deconvolution of bulk tissue dna methylomes from single-cell rna-seq data. *Genome biology*, 21(1):1–33, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02126-9>.
3. Longxing Cao, Brian Coventry, Inna Goreschnik, Buwei Huang, Joon Sung Park, Kevin M Jude, Iva Marković, Rameshwar U Kadam, Koen HG Verschueren, Kenneth Verstraete, and others. Design of protein binding proteins from target structure alone. *Nature*, pages 1–1, 2022. URL: <https://www.nature.com/articles/s41586-022-04654-9>.
4. Sergey Nurk, Sergey Koren, Arang Rhie, Mikko Rautiainen, Andrey V Bzikadze, Alla Mikheenko, Mitchell R Vollger, Nicolas Altomose, Lev Uralsky, Ariel Gershman, and others. The complete sequence of a human genome. *Science*, 376(6588):44–53, 2022. URL: <https://www.science.org/doi/10.1126/science.abj6987>.
5. Seyhan Yazar, Jose Alquicira-Hernandez, Kristof Wing, Anne Senabouth, M Grace Gordon, Stacey Andersen, Qinyi Lu, Antonia Rowson, Thomas RP Taylor, Linda Clarke, and others. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science*, 376(6589):eabf3041, 2022. URL: <https://www.science.org/doi/10.1126/science.abf3041>.
6. Kasper Karlsson, Moritz Przybilla, Hang Xu, Eran Kotler, Kremena Karagyzova, Alexandra Sockell, Katherine Liu, Amanda Mah, Yuan-Hung Lo, Bingxin Lu, Kathleen E. Houlahan, Aziz Khan, Zhicheng Ma, Carlos J. Suarez, Christopher P. Barnes, Calvin J. Kuo, and Christina Curtis. Experimental evolution in tp53 deficient gastric organoids recapitulates tumorigenesis. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/04/10/2022.04.09.487529>, doi:10.1101/2022.04.09.487529.
7. Joannella Morales, Shashikant Pujar, Jane E Loveland, Alex Astashyn, Ruth Bennett, Andrew Berry, Eric Cox, Claire Davidson, Olga Ermolaeva, Catherine M Farrell, and others. A joint ncbi and embl-ebi transcript set for clinical genomics and research. *Nature*, pages 1–6, 2022. URL: <https://www.nature.com/articles/s41586-022-04558-8>.
8. Zhong Wang, Alexandra G Chivu, Lauren A Choate, Edward J Rice, Donald C Miller, Tinyi Chu, Shao-Pei Chou, Nicole B Kingsley, Jessica L Petersen, Carrie J Finno, and others. Prediction of histone post-translational modification patterns based on nascent transcription data. *Nature Genetics*, pages 1–11, 2022.

2.3.7 2022-03-30

1. Yang Xu and Rachel Patton McCord. Costa: unsupervised convolutional neural network learning for spatial transcriptomics analysis. *BMC bioinformatics*, 22(1):1–26, 2021.
2. Caibin Sheng, Rui Lopes, Gang Li, Sven Schuierer, Annick Waladt, Rachel Cuttat, Slavica Dimitrieva, Audrey Kauffmann, Eric Durand, Giorgio G. Galli, Guglielmo Roma, and Antoine de Weck. Probabilistic machine learning ensures accurate ambient denoising in droplet-based single-cell omics. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/03/24/2022.01.14.476312>, arXiv:<https://www.biorxiv.org/content/early/2022/03/24/2022.01.14.476312.full.pdf>, doi:10.1101/2022.01.14.476312.
3. Sayan Ghosal, Qiang Chen, Giulio Pergola, Aaron L Goldman, William Ulrich, Daniel R Weinberger, and Archana Venkataraman. A biologically interpretable graph convolutional network to link genetic risk pathways and imaging phenotypes of disease. In *International Conference on Learning Representations*. 2022. URL: <https://openreview.net/forum?id=Lwr8We4MIxn>.
4. Deniz Demircioğlu, Engin Cukuroglu, Martin Kindermans, Tannistha Nandi, Claudia Calabrese, Nuno A Fonseca, André Kahles, Kjong-Van Lehmann, Oliver Stegle, Alvis Brazma, and others. A pan-cancer transcriptome analysis reveals pervasive regulation through alternative promoters. *Cell*, 178(6):1465–1477, 2019.
5. Lu Pan, Huy Q Dinh, Yudi Pawitan, and Trung Nghia Vu. Isoform-level quantification for single-cell rna sequencing. *Bioinformatics*, 38(5):1287–1294, 2022.
6. Zexian Zeng, Yawei Li, Yiming Li, and Yuan Luo. Statistical and machine learning methods for spatially resolved transcriptomics data analysis. *Genome Biology*, 23(1):1–23, 2022. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-022-02653-7#Sec7>.
7. Felipe A. Simão, Robert M Waterhouse, Panagiotis Ioannidis, Evgenia V Kriventseva, and Evgeny M Zdobnov. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19):3210–3212, 2015. arXiv:[arXiv:1410.3554v1](https://arxiv.org/abs/1410.3554v1), doi:10.1093/bioinformatics/btv351.

2.3.8 2022-03-09

1. Teng Gao, Ruslan Soldatov, Hira Sarkar, Adam Kurkiewicz, Evan Biederstedt, Po-Ru Loh, and Peter Kharchenko. Haplotype-enhanced inference of somatic copy number profiles from single-cell transcriptomes. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/02/10/2022.02.07.479314>, arXiv:<https://www.biorxiv.org/content/early/2022/02/10/2022.02.07.479314.full.pdf>, doi:10.1101/2022.02.07.479314.
2. Marc J Williams, Tyler Funnell, Ciara H O’Flanagan, Andrew McPherson, Sohrab Salehi, Ignacio Vázquez-García, Farhia Kabeer, Hakwoo Lee, Tehmina Masud, Peter Eirew, Damian Yap, Beixi Wang, Jazmine Brimhall, Justina Biele, Jerome Ting, Sean Beatty, Daniel Lai, Jenifer Pham, Diljot Grewal, Douglas Abrams, Eliyahu Havasov, Samantha Leung, Viktoria Bojilova, Adam C Weiner, Nicole Rusk, Florian Uhlitz, Nicholas Ceglia, IMAXT consortium, Samuel Aparicio, and Sohrab P. Shah. Evolutionary tracking of cancer haplotypes at single-cell resolution. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/06/06/2021.06.04.447031>, arXiv:<https://www.biorxiv.org/content/early/2021/06/06/2021.06.04.447031.full.pdf>, doi:10.1101/2021.06.04.447031.
3. Biswajyoti Sahu, Tuomo Hartonen, Päivi Pihlajamaa, Bei Wei, Kashyap Dave, Fangjie Zhu, Eevi Kaasinen, Katja Lidschreiber, Michael Lidschreiber, Carsten O Daub, and others. Sequence determinants of human gene regulatory elements. *Nature Genetics*, pages 1–12, 2022. URL: <https://www.nature.com/articles/s41588-021-01009-4>.
4. Julia Eve Olivieri, Roozbeh Dehghannasiri, and Julia Salzman. The spliz generalizes ‘percent spliced in’ to reveal regulated splicing at single-cell resolution. *Nature Methods*, pages 1–4, 2022. URL: <https://www.nature.com/articles/s41592-022-01400-x>.

5. Gonzalo Benegas, Jonathan Fischer, and Yun S Song. Robust and annotation-free analysis of alternative splicing across diverse cell types in mice. *eLife*, 11:e73520, 2022. URL: <https://elifesciences.org/articles/73520>.
6. Tyler E Miller, Caleb A Lareau, Julia A Verga, Erica AK DePasquale, Vincent Liu, Daniel Ssozi, Katalin Sandor, Yajie Yin, Leif S Ludwig, Chadi A El Farran, and others. Mitochondrial variant enrichment from high-throughput single-cell rna sequencing resolves clonal populations. *Nature Biotechnology*, pages 1–5, 2022. URL: <https://www.nature.com/articles/s41587-022-01210-8>.

2.3.9 2022-02-23

1. Shou-Wen Wang, Michael J Herriges, Kilian Hurley, Darrell N Kotton, and Allon M Klein. Cospar identifies early cell fate biases from single-cell transcriptomic and lineage information. *Nature Biotechnology*, pages 1–9, 2022. URL: <https://www.nature.com/articles/s41587-022-01209-1>.
2. Ran Su, YiXuan Huang, De-gan Zhang, Guobao Xiao, and Leyi Wei. Srdfm: siamese response deep factorization machine to improve anti-cancer drug recommendation. *Briefings in Bioinformatics*, 01 2022. bbab534. URL: <https://doi.org/10.1093/bib/bbab534>, arXiv:<https://academic.oup.com/bib/advance-article-pdf/doi/10.1093/bib/bbab534/42236165/bbab534.pdf>, doi:10.1093/bib/bbab534.

2.3.10 2022-02-09

1. Morten Rasmussen, Mitsu Reddy, Rory Nolan, Joan Camunas-Soler, Arkady Khodursky, Nikolai M Scheller, David E Cantonwine, Line Engelbrechtsen, Jia Dai Mi, Arup Dutta, and others. Rna profiles reveal signatures of future health and disease in pregnancy. *Nature*, pages 1–6, 2022. URL: <https://www.nature.com/articles/s41586-021-04249-w>.
2. Alexei Baevski, Wei-Ning Hsu, Qiantong Xu, Arun Babu, Jiatao Gu, and Michael Auli. Data2vec: a general framework for self-supervised learning in speech, vision and language. *Meta AI*, 2022. URL: <https://ai.facebook.com/research/data2vec-a-general-framework-for-self-supervised-learning-in-speech-vision-and-language>.
3. Jingsi Ming, Jia Zhao, and Can Yang. Scpi: a scalable framework for probabilistic inference in single-cell rna-sequencing data analysis. *Statistics in Biosciences*, pages 1–24, 2022. URL: <https://link.springer.com/article/10.1007/s12561-022-09335-9>.
4. Michael Eisenstein. Seven technologies to watch in 2022. *Nature*, 601(7894):658–661, 2022. URL: <https://www.nature.com/articles/d41586-022-00163-x>.
5. Sevahn K Vorperian, Mira N Moufarrej, Stephen R Quake, Tabula Sapiens Consortium, and others. Cell types of origin in the cell free transcriptome in human health and disease. *bioRxiv*, 2021. URL: <https://www.nature.com/articles/s41587-021-01188-9>.
6. Logan G Wright, Tatsuhiko Onodera, Martin M Stein, Tianyu Wang, Darren T Schachter, Zoey Hu, and Peter L McMahon. Deep physical neural networks trained with backpropagation. *Nature*, 601(7894):549–555, 2022. URL: <https://www.nature.com/articles/s41586-021-04223-6>.

2.3.11 2022-01-19

1. Vitalii Kleshchevnikov, Artem Shmatko, Emma Dann, Alexander Aivazidis, Hamish W. King, Tong Li, Rasa Elmentaite, Artem Lomakin, Veronika Kedlian, Adam Gayoso, Mika Sarkin Jain, Jun Sung Park, Lauma Ramona, Elizabeth Tuck, Anna Arutyunyan, Roser Vento-Tormo, Moritz Gerstung, Louisa James, Oliver Stegle, and Omer Ali Bayraktar. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nature Biotechnology*, Jan 2022. URL: <https://doi.org/10.1038/s41587-021-01139-4>, doi:10.1038/s41587-021-01139-4.
2. Britta Velten, Jana M. Braunger, Ricard Argelaguet, Damien Arnol, Jakob Wirbel, Danila Bredikhin, Georg Zeller, and Oliver Stegle. Identifying temporal and spatial patterns of variation from multimodal data using

- mefisto. *Nature Methods*, Jan 2022. URL: <https://doi.org/10.1038/s41592-021-01343-9>, doi:10.1038/s41592-021-01343-9.
3. Marius Lange, Volker Bergen, Michal Klein, Manu Setty, Bernhard Reuter, Mostafa Bakhti, Heiko Lickert, Meshal Ansari, Janine Schniering, Herbert B. Schiller, Dana Pe'er, and Fabian J. Theis. Cellrank for directed single-cell fate mapping. *Nature Methods*, Jan 2022. URL: <https://doi.org/10.1038/s41592-021-01346-6>, doi:10.1038/s41592-021-01346-6.
 4. Jialin Qu and Yuehua Cui. Gene set analysis with graph-embedded kernel association test. *Bioinformatics*, 12 2021. btab851. URL: <https://doi.org/10.1093/bioinformatics/btab851>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btab851/42003921/btab851.pdf>, doi:10.1093/bioinformatics/btab851.
 5. Xuejian Li, Shiqiang Ma, Jin Liu, Jijun Tang, and Fei Guo. Inferring gene regulatory network via fusing gene expression image and RNA-seq data. *Bioinformatics*, 01 2022. btac008. URL: <https://doi.org/10.1093/bioinformatics/btac008>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btac008/42091226/btac008.pdf>, doi:10.1093/bioinformatics/btac008.
 6. Ilia Kats, Roser Vento-Tormo, and Oliver Stegle. Spatialde2: fast and localized variance component analysis of spatial transcriptomics. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/11/11/2021.10.27.466045>, arXiv:<https://www.biorxiv.org/content/early/2021/11/11/2021.10.27.466045.full.pdf>, doi:10.1101/2021.10.27.466045.
 7. J Monroe, Thanvi Srikant, Pablo Carbonell-Bejerano, Claude Becker, Mariele Lensink, Moises Exposito-Alonso, Marie Klein, Julia Hildebrandt, Manuela Neumann, Daniel Kliebenstein, and others. Mutation bias reflects natural selection in arabidopsis thaliana. *Nature*, pages 1–5, 2022. URL: <https://www.nature.com/articles/s41586-021-04269-6>.
 8. Chen Li, Maria Virgilio, Kathleen Collins, and Joshua D Welch. Single-cell multi-omic velocity infers dynamic and decoupled gene regulation. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.12.13.472472v1>.
 9. Wei Xu, Weilong Yang, Yunlong Zhang, Yawen Chen, Qian Zhang, Xuefei Wang, Kun Song, Wenfei Jin, and Xi Chen. Issaac-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. *bioRxiv*, 2022. URL: <https://www.biorxiv.org/content/early/2022/01/17/2022.01.16.476488>, doi:10.1101/2022.01.16.476488.
 10. Gaoyang Li, Shaliu Fu, Shuguang Wang, Chenyu Zhu, Bin Duan, Chen Tang, Xiaohan Chen, Guohui Chuai, Ping Wang, and Qi Liu. A deep generative model for multi-view profiling of single-cell rna-seq and atac-seq data. *Genome Biology*, 23(1):1–23, 2022.
 11. Liangtao Zheng, Shishang Qin, Wen Si, Anqiang Wang, Baocai Xing, Ranran Gao, Xianwen Ren, Li Wang, Xiaojiang Wu, Ji Zhang, and others. Pan-cancer single-cell landscape of tumor-infiltrating t cells. *Science*, 374(6574):abe6474, 2021.

2.3.12 2021-12-10

1. Antonio De Falco, Francesca P Caruso, Xiao Dong Su, Antonio Iavarone, and Michele Ceccarelli. A fast variational algorithm to detect the clonal copy number substructure of tumors from single-cell data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.11.20.469390v1>.
2. Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, and others. Advancing mathematics by guiding human intuition with ai. *Nature*, 600(7887):70–74, 2021.
3. Minking Pang, Kenong Su, and Mingyao Li. Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/11/28/2021.11>.

28.470212, arXiv:<https://www.biorxiv.org/content/early/2021/11/28/2021.11.28.470212.full.pdf>, doi:10.1101/2021.11.28.470212.

4. Allen W. Lynch, Christina V. Theodoris, Henry Long, Myles Brown, X. Shirley Liu, and Clifford A. Meyer. MIRA: joint regulatory modeling of multimodal expression and chromatin accessibility in single cells. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.12.06.471401v1?rss=1>, doi:10.1101/2021.12.06.471401.
5. Ludvig Bergenstr hle, Bryan He, Joseph Bergenstr hle, Xes s Abalo, Reza Mirzazadeh, Kim Thrane, Andrew L Ji, Alma Andersson, Ludvig Larsson, Nathalie Stakenborg, and others. Super-resolved spatial transcriptomics by deep data fusion. *Nature biotechnology*, pages 1–4, 2021. URL: <https://www.nature.com/articles/s41587-021-01075-3>.
6. Youjin Hu, Jiawei Zhong, Yuhua Xiao, Zheng Xing, Katherine Sheu, Shuxin Fan, Qin An, Yuanhui Qiu, Yingfeng Zheng, Xialin Liu, and others. Single-cell rna cap and tail sequencing (scrcat-seq) reveals subtype-specific isoforms differing in transcript demarcation. *Nature communications*, 11(1):1–11, 2020. URL: <https://www.nature.com/articles/s41467-020-18976-7>.

2.3.13 2021-11-26

1. Jian Zhou and Olga G Troyanskaya. An analytical framework for interpretable and generalizable single-cell data analysis. *Nature Methods*, pages 1–5, 2021. URL: <https://www.nature.com/articles/s41592-021-01286-1>.
2. Jing Gong, Kui Xu, Ziyuan Ma, Zhi John Lu, and Qiangfeng Cliff Zhang. A deep learning method for recovering missing signals in transcriptome-wide rna structure profiles from probing experiments. *Nature Machine Intelligence*, 3(11):995–1006, 2021.
3. Claudio Lorenzi, Sylvain Barriere, Katharina Arnold, Reini F Luco, Andrew J Oldfield, and William Ritchie. Irfinder-s: a comprehensive suite to discover and explore intron retention. *Genome biology*, 22(1):1–13, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02515-8>.
4. Luyi Tian, Jafar S Jabbari, Rachel Thijssen, Quentin Gouil, Shanika L Amarasinghe, Oliver Voogd, Hasaru Kariyawasam, Mei RM Du, Jakob Schuster, Changqing Wang, and others. Comprehensive characterization of single-cell full-length isoforms in human and mouse with long-read sequencing. *Genome biology*, 22(1):1–24, 2021.
5. Duanchen Sun, Xiangnan Guan, Amy E Moran, Ling-Yun Wu, David Z Qian, Pepper Schedin, Mu-Shui Dai, Alexey V Danilov, Joshi J Alumkal, Andrew C Adey, and others. Identifying phenotype-associated subpopulations by integrating bulk and single-cell sequencing data. *Nature Biotechnology*, pages 1–12, 2021. URL: <https://www.nature.com/articles/s41587-021-01091-3>.
6. Maren Buettner, Johannes Ostner, Christian L Mueller, Fabian J Theis, and Benjamin Schubert. Scoda is a bayesian model for compositional single-cell data analysis. *Nature communications*, 12(1):1–10, 2021. URL: <https://www.nature.com/articles/s41467-021-27150-6>.

2.3.14 2021-11-12

1. Luke Zappia and Fabian J Theis. Over 1000 tools reveal trends in the single-cell rna-seq analysis landscape. *bioRxiv*, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02519-4>.
2. PF Ferreira, J Kuipers, and N Beerenwinkel. Mapping single-cell transcriptomes to copy number evolutionary trees. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.11.04.467244v1.full.pdf+html>.
3. Viktor Petukhov, Rosalind J Xu, Ruslan A Soldatov, Paolo Cadinu, Konstantin Khodosevich, Jeffrey R Moffitt, and Peter V Kharchenko. Cell segmentation in imaging-based spatial transcriptomics. *Nature Biotechnology*, pages 1–10, 2021. URL: <https://www.nature.com/articles/s41587-021-01044-w>.

4. Tommaso Biancalani, Gabriele Scalia, Lorenzo Buffoni, Raghav Avasthi, Ziqing Lu, Aman Sanger, Neriman Tokcan, Charles R Vanderburg, Åsa Segerstolpe, Meng Zhang, and others. Deep learning and alignment of spatially resolved single-cell transcriptomes with tangram. *Nature Methods*, pages 1–11, 2021. URL: <https://www-nature-com.eproxy.lib.hku.hk/articles/s41592-021-01264-7>.
5. Jonathan Moody, Tsukasa Kouno, Akari Suzuki, Youtaro Shibayama, Chikashi Terao, Jen-Chien Chang, Fernando López-Redondo, Chi Wai Yip, Jessica Severin, Hiroyuki Suetsugu, and others. Profiling of transcribed cis-regulatory elements in single cells. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.04.04.438388v2.full>.

2.3.15 2021-10-29

1. Ran Cheng and Peter K Jackson. Identifying cancer drivers. *Science*, 374(6563):38–39, 2021. URL: <https://www.science.org/doi/epdf/10.1126/science.abl9080>.
2. J. Frazer, P. Notin, M. Dias, A. Gomez, J. Min, K. Brock, Y. Gal, and D. Marks. Disease variant prediction with deep generative models of evolutionary data. *Nature*, pages 1–10, 2021. URL: <https://www.nature.com/articles/s41586-021-04043-8>.
3. Sabrina Rashid, Sohrab Shah, Ziv Bar-Joseph, and Ravi Pandya. Dhaka: variational autoencoder for unmasking tumor heterogeneity from single cell genomic data. *Bioinformatics*, 37(11):1535–1543, 02 2019. URL: <https://doi.org/10.1093/bioinformatics/btz095>, arXiv:<https://academic.oup.com/bioinformatics/article-pdf/37/11/1535/38923762/btz095.pdf>, doi:10.1093/bioinformatics/btz095.
4. Jian Hu, Xiangjie Li, Kyle Coleman, Amelia Schroeder, David J Irwin, Edward B Lee, Russell T Shinohara, and Mingyao Li. Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *bioRxiv*, 2020.
5. Ruochen Jiang, Tianyi Sun, Dongyuan Song, and Jingyi Jessica Li. Statistics or biology: the zero-inflation controversy about scrna-seq data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/09/21/2020.12.28.424633>, arXiv:<https://www.biorxiv.org/content/early/2021/09/21/2020.12.28.424633.full.pdf>, doi:10.1101/2020.12.28.424633.
6. Xian Adiconis, Adam L Haber, Sean K Simmons, Ami Levy Moonshine, Zhe Ji, Michele A Busby, Xi Shi, Justin Jacques, Madeline A Lancaster, Jen Q Pan, and others. Comprehensive comparative analysis of 5-end rna-sequencing methods. *Nature methods*, 15:505–511, 2018. URL: <https://www.nature.com/articles/s41592-018-0014-2>.
7. Jingjing Xu, Hao Zhou, Chun Gan, Zaixiang Zheng, and Lei Li. Vocabulary learning via optimal transport for neural machine translation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 7361–7373. Online, August 2021. Association for Computational Linguistics. URL: <https://aclanthology.org/2021.acl-long.571>, doi:10.18653/v1/2021.acl-long.571.

2.3.16 2021-10-15

1. Gamze Gürsoy, Nancy Lu, Sarah Wagner, and Mark Gerstein. Recovering genotypes and phenotypes using allele-specific genes. *Genome biology*, 22(1):1–9, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02477-x>.
2. Johan Winnubst and Silvia Arber. A census of cell types in the brain’s motor cortex. 2021. URL: <https://www.nature.com/articles/d41586-021-02493-8>.
3. A Booeshaghi, Zizhen Yao, Cindy van Velthoven, Kimberly Smith, Bosiljka Tasic, Hongkui Zeng, and Lior Pachter. Isoform cell-type specificity in the mouse primary motor cortex. *Nature*, 598(7879):195–199, 2021.

4. Lukas Breitwieser, Ahmad Hesam, Jean de Montigny, Vasileios Vavourakis, Alexandros Iosif, Jack Jennings, Marcus Kaiser, Marco Manca, Alberto Di Meglio, Zaid Al-Ars, Fons Rademakers, Onur Mutlu, and Roman Bauer. BioDynaMo: a modular platform for high-performance agent-based simulation. *Bioinformatics*, 09 2021. btab649. URL: <https://doi.org/10.1093/bioinformatics/btab649>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btab649/40395475/btab649.pdf>, doi:10.1093/bioinformatics/btab649.
5. Samuel Morabito, Emily Miyoshi, Neethu Michael, Saba Shahin, Alessandra Cadete Martini, Elizabeth Head, Justine Silva, Kelsey Leavy, Mari Perez-Rosendahl, and Vivek Swarup. Single-nucleus chromatin accessibility and transcriptomic characterization of alzheimer's disease. *Nature Genetics*, 53(8):1143–1155, 2021. URL: <https://www-nature-com.eproxy.lib.hku.hk/articles/s41588-021-00894-z>.
6. Michael VanInsberghe, Jeroen van den Berg, Amanda Andersson-Rolf, Hans Clevers, and Alexander van Oude-naarden. Single-cell ribo-seq reveals cell cycle-dependent translational pausing. *Nature*, pages 1–5, 2021. URL: <https://www.nature.com/articles/s41586-021-03887-4>.

2.3.17 2021-09-17

1. Kamila Naxerova. Mutation fingerprints encode cellular histories. 2021. URL: <https://www.nature.com/articles/d41586-021-02269-0>.
2. Natsuhiko Kumasaka, Raghd Rostom, Ni Huang, Krzysztof Polanski, Kerstin Meyer, Sharad Patel, Rachel Boyd, Celine Gomez, Sam Barnett, Nikolaos Panousis, and others. Mapping interindividual dynamics of innate immune response at single-cell resolution. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.09.01.457774v1>.
3. Anna SE Cuomo, Tobias Heinen, Danai Vagiaki, Danilo Horta, John Marioni, and Oliver Stegle. Cellregmap: a statistical framework for mapping context-specific regulatory variants using scrna-seq. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.09.01.458524v1>.
4. Wei Wei, Daniel J Gaffney, and Patrick F Chinnery. Cell reprogramming shapes the mitochondrial dna landscape. *Nature Communications*, 12(1):1–15, 2021.
5. Akdes Serin Harmanci, Arif O Harmanci, Xiaobo Zhou, Benjamin Deneen, Ganesh Rao, Tiemo Klisch, and Akash Patel. Scregulocity: detection of local rna velocity patterns in embeddings of single cell rna-seq data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/06/02/2021.06.01.446674>, arXiv:<https://www.biorxiv.org/content/early/2021/06/02/2021.06.01.446674.full.pdf>, doi:10.1101/2021.06.01.446674.
6. Ismael Lemhadri, Feng Ruan, and Rob Tibshirani. Lassonet: neural networks with feature sparsity. In Arindam Banerjee and Kenji Fukumizu, editors, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of Proceedings of Machine Learning Research, 10–18. PMLR, 13–15 Apr 2021. URL: <https://proceedings.mlr.press/v130/lemhadri21a.html>.
7. Marissa Connor, Gregory Canal, and Christopher Rozell. Variational autoencoder with learned latent structure. In Arindam Banerjee and Kenji Fukumizu, editors, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of Proceedings of Machine Learning Research, 2359–2367. PMLR, 13–15 Apr 2021. URL: <https://proceedings.mlr.press/v130/connor21a.html>.
8. Wenjing Ma, Kenong Su, and Hao Wu. Evaluation of some aspects in supervised cell type identification for single-cell rna-seq: classifier, feature selection, and reference construction. *Genome Biology*, 22(1):1–23, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02480-2>.
9. Zachary J. DeBruine, Karsten Melcher, and Timothy J. Triche. Fast and robust non-negative matrix factorization for single-cell experiments. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/09/01/2021.09.01.458620>, arXiv:<https://www.biorxiv.org/content/early/2021/09/01/2021.09.01.458620.full.pdf>, doi:10.1101/2021.09.01.458620.

10. Mohammad Lotfollahi, Mohsen Naghipourfar, Malte D Luecken, Matin Khajavi, Maren Büttner, Marco Wagenstetter, Žiga Avsec, Adam Gayoso, Nir Yosef, Marta Interlandi, and others. Mapping single-cell data to reference atlases by transfer learning. *Nature Biotechnology*, pages 1–10, 2021. URL: <https://www.nature.com/articles/s41587-021-01001-7>.

2.4 Pub 2020-2021

2.4.1 2021-07-26

1. Ploy N Pratanwanich, Fei Yao, Ying Chen, Casslynn WQ Koh, Yuk Kei Wan, Christopher Hendra, Polly Poon, Yeek Teck Goh, Phoebe ML Yap, Jing Yuan Chooi, and others. Identification of differential rna modifications from nanopore direct rna sequencing with xpore. *Nature Biotechnology*, pages 1–9, 2021. URL: <https://www.nature.com/articles/s41587-021-00949-w>.
2. Nathanael Andrews, Jason T Serviss, Natalie Geyer, Agneta B Andersson, Ewa Dzwonkowska, Iva Šutevski, Rosan Heijboer, Ninib Baryawno, Marco Gerling, and Martin Enge. An unsupervised method for physical cell interaction profiling of complex tissues. *Nature Methods*, pages 1–9, 2021. URL: <https://www.nature.com/articles/s41592-021-01196-2>.
3. Leah L Weber, Palash Sashittal, and Mohammed El-Kebir. Doubletd: detecting doublets in single-cell dna sequencing data. *Bioinformatics*, 37(Supplement_1):i214–i221, 2021. URL: <https://doi.org/10.1093/bioinformatics/btab266>.
4. Ariya Shajii, Ibrahim Numanagić, Alexander T Leighton, Haley Greenyer, Saman Amarasinghe, and Bonnie Berger. A python-based programming language for high-performance computational genomics. *Nature Biotechnology*, pages 1–2, 2021. URL: <https://www.nature.com/articles/s41587-021-00985-6>.
5. David M. Kurtz, Joanne Soo, Lyron Co Ting Keh, Stefan Alig, Jacob J. Chabon, Brian J. Sworder, Andre Schultz, Michael C. Jin, Florian Scherer, Andrea Garofalo, Charles W. Macaulay, Emily G. Hamilton, Binbin Chen, Mari Olsen, Joseph G. Schroers-Martin, Alexander F. M. Craig, Everett J. Moding, Mohammad S. Esfahani, Chih Long Liu, Ulrich Dührsen, Andreas Hüttmann, René-Olivier Casasnovas, Jason R. Westin, Mark Roschewski, Wyndham H. Wilson, Gianluca Gaidano, Davide Rossi, Maximilian Diehn, and Ash A. Alizadeh. Enhanced detection of minimal residual disease by targeted sequencing of phased variants in circulating tumor dna. *Nature Biotechnology*, Jul 2021. URL: <https://doi.org/10.1038/s41587-021-00981-w>, doi:10.1038/s41587-021-00981-w.
6. Ricard Argelaguet, Anna SE Cuomo, Oliver Stegle, and John C Marioni. Computational principles and challenges in single-cell data integration. *Nature Biotechnology*, pages 1–14, 2021. URL: <https://www.nature.com/articles/s41587-021-00895-7>.
7. Xiaoyang Jing and Jinbo Xu. Fast and effective protein model refinement using deep graph neural networks. *Nature Computational Science*, pages 1–8, 2021. URL: <https://www.nature.com/articles/s43588-021-00098-9>.
8. Naozumi Hiranuma, Hahnbeom Park, Minkyung Baek, Ivan Anishchenko, Justas Dauparas, and David Baker. Improved protein structure refinement guided by deep learning based accuracy estimation. *Nature communications*, 12(1):1–11, 2021. URL: <https://www.nature.com/articles/s41467-021-21511-x>.
9. Linhua Wang and Zhandong Liu. Missing-value imputation and in-silico region detection for spatially resolved transcriptomics. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/05/17/2021.05.14.443446>, arXiv:<https://www.biorxiv.org/content/early/2021/05/17/2021.05.14.443446.full.pdf>, doi:10.1101/2021.05.14.443446.
10. Andrew Erickson, Emelie Berglund, Mengxiao He, Maja Marklund, Reza Mirzazadeh, Niklas Schultz, Ludvig Bergenstråhle, Linda Kvastad, Alma Andersson, Joseph Bergenstråhle, Ludvig Larsson, Alia Shamikh, Elisa Basmaci, Teresita Diaz De Ståhl, Timothy Rajakumar, Kim Thrane, Andrew L Ji, Paul A Khavari, Firaz Tarish, Anna Tanoglidi, Jonas Maaskola, Richard Colling, Tuomas Mirtti, Freddie C Hamdy, Dan J Woodcock, Thomas

Helleday, Ian G. Mills, Alastair D Lamb, and Joakim Lundeberg. The spatial landscape of clonal somatic mutations in benign and malignant tissue. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/07/12/2021.07.12.452018>, arXiv:<https://www.biorxiv.org/content/early/2021/07/12/2021.07.12.452018.full.pdf>, doi:10.1101/2021.07.12.452018.

11. Saket Choudhary and Rahul Satija. Comparison and evaluation of statistical error models for scRNA-seq. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.07.07.451498v1>.

2.4.2 2021-07-12

1. Yasushi Okochi, Shunta Sakaguchi, Ken Nakae, Takefumi Kondo, and Honda Naoki. Model-based prediction of spatial gene expression via generative linear mapping. *Nature Communications*, 12(1):1–13, 2021. URL: <https://www.nature.com/articles/s41467-021-24014-x>.
2. Brendan Francis Miller, Lyla Atta, Arpan Sahoo, Feiyang Huang, and Jean Fan. Reference-free cell-type deconvolution of pixel-resolution spatially resolved transcriptomics data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.06.15.448381v1>.
3. Hananeh Aliee and Fabian J Theis. AutoGeneS: Automatic gene selection using multi-objective optimization for RNA-seq deconvolution. *Cell Systems*, 2021. URL: <https://www.sciencedirect.com/science/article/pii/S2405471221001927>.
4. Lily Zheng, Laura Wood, Rachel Karchin, and Robert Scharpf. Estimation of cancer cell fractions and clone trees from multi-region sequencing of tumors. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.06.12.448194v1>.
5. Peter Carbonetto, Abhishek Sarkar, Zihao Wang, and Matthew Stephens. Non-negative matrix factorization algorithms greatly improve topic model fits. *arXiv preprint arXiv:2105.13440*, 2021. URL: <https://arxiv.org/abs/2105.13440>.
6. Sohrab Salehi, Farhia Kabeer, Nicholas Ceglia, Mirela Andronescu, Marc J Williams, Kieran R Campbell, Tehmina Masud, Beixi Wang, Justina Biele, Jazmine Brimhall, and others. Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature*, pages 1–6, 2021. URL: <https://www.nature.com/articles/s41586-021-03648-3>.
7. Livius Penter, Satyen H Gohil, Caleb Lareau, Leif S Ludwig, Erin M Parry, Teddy Huang, Shuqiang Li, Wandu Zhang, Dimitri Livitz, Ignaty Leshchiner, and others. Longitudinal single-cell dynamics of chromatin accessibility and mitochondrial mutations in chronic lymphocytic leukemia mirror disease history. *Cancer Discovery*, 2021. URL: <https://cancerdiscovery.aacrjournals.org/content/early/2021/06/10/2159-8290.CD-21-0276>.
8. Junyan Lu, Ester Cannizzaro, Fabienne Meier-Abt, Sebastian Scheinost, Peter-Martin Bruch, Holly AR Giles, Almut Lütge, Jennifer Hülle, Lena Wagner, Brian Giacomelli, and others. Multi-omics reveals clinically relevant proliferative drive associated with mTOR-myc-oxphos activity in chronic lymphocytic leukemia. *Nature Cancer*, pages 1–12, 2021. URL: <https://www.nature.com/articles/s43018-021-00216-6>.
9. Chenchen Zhu, Jingyan Wu, Han Sun, Francesca Briganti, Benjamin Meder, Wu Wei, and Lars M Steinmetz. Single-molecule, full-length transcript isoform sequencing reveals disease mutation-associated rna isoforms in cardiomyocytes. *Nature Communications*, 2021. URL: <https://www.nature.com/articles/s41467-021-24484-z>.
10. Chong Chu, Rebeca Borges-Monroy, Vinayak V Viswanadham, Soohyun Lee, Heng Li, Eunjung Alice Lee, and Peter J Park. Comprehensive identification of transposable element insertions using multiple sequencing technologies. *Nature Communications*, 12(1):1–12, 2021.
11. Peter V Kharchenko. The triumphs and limitations of computational methods for scRNA-seq. *Nature Methods*, pages 1–10, 2021.
12. Carter Allen, Yuzhou Chang, Brian Neelon, Won Chang, Hang J. Kim, Zihai Li, Qin Ma, and Dongjun Chung. A bayesian multivariate mixture model for spatial transcriptomics data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/06/24/2021>.

- 06.23.449615, arXiv:<https://www.biorxiv.org/content/early/2021/06/24/2021.06.23.449615.full.pdf>, doi:10.1101/2021.06.23.449615.
13. Jun Zhao, Ariel Jaffe, Henry Li, Ofir Lindenbaum, Esen Sefik, Ruaidhrí Jackson, Xiuyuan Cheng, Richard A Flavell, and Yuval Kluger. Detection of differentially abundant cell subpopulations in scRNA-seq data. *Proceedings of the National Academy of Sciences*, 2021. URL: <https://www.pnas.org/content/118/22/e2100293118>.
 14. Shang-Hua Gao, Qi Han, Duo Li, Pai Peng, Ming-Ming Cheng, and Pai Peng. Representative batch normalization with feature calibration. In *CVPR*. 2021.
 15. Yue You, Luyi Tian, Shian Su, Xueyi Dong, Jafar S Jabbari, Peter F Hickey, and Matthew E Ritchie. Benchmarking umi-based single cell RNA-seq preprocessing workflows. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/06/17/2021.06.17.448895>, arXiv:<https://www.biorxiv.org/content/early/2021/06/17/2021.06.17.448895.full.pdf>, doi:10.1101/2021.06.17.448895.
 16. Dongze He, Mohsen Zakeri, Hirak Sarkar, Charlotte Sonesson, Avi Srivastava, and Rob Patro. Alevin-fry unlocks rapid, accurate, and memory-frugal quantification of single-cell RNA-seq data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/07/01/2021.06.29.450377>, arXiv:<https://www.biorxiv.org/content/early/2021/07/01/2021.06.29.450377.full.pdf>, doi:10.1101/2021.06.29.450377.
 17. Alexander S Garruss, Katherine M Collins, and George M Church. Deep representation learning improves prediction of laci-mediated transcriptional repression. *Proceedings of the National Academy of Sciences*, 2021. URL: <https://www.pnas.org/content/118/27/e2022838118>.

2.4.3 2021-05-31

1. Stefanie Warnat-Herresthal, Hartmut Schultze, Krishna Prasad Lingadahalli Shastry, Sathyanarayanan Manamohan, Saikat Mukherjee, Vishesh Garg, Ravi Sarveswara, Kristian Haendler, Peter Pickkers, N Ahmad Aziz, and others. Swarm Learning as a privacy-preserving machine learning approach for disease classification. *Nature*, 2021. URL: <https://www.nature.com/articles/s41586-021-03583-3>.
2. Ming Y Lu, Tiffany Y Chen, Drew FK Williamson, Melissa Zhao, Maha Shady, Jana Lipkova, and Faisal Mahmood. AI-based pathology predicts origins for cancers of unknown primary. *Nature*, pages 1–5, 2021. URL: <https://www.nature.com/articles/s41586-021-03512-4>.
3. Michael Spencer Chapman, Anna Maria Ranzoni, Brynnele Myers, Nicholas Williams, Tim HH Coorens, Emily Mitchell, Timothy Butler, Kevin J Dawson, Yvette Hooks, Luiza Moore, and others. Lineage tracing of human development through somatic mutations. *Nature*, pages 1–6, 2021. URL: <https://www.nature.com/articles/s41586-021-03548-6>.
4. Paolo Marangio, Ka Ying Toby Law, Guido Sanguinetti, and Sander Granneman. DiffBUM-HMM: a robust statistical modeling approach for detecting RNA flexibility changes in high-throughput structure probing data. *Genome Biology*, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02379-y>.
5. Abhishek Sarkar and Matthew Stephens. Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis. *Nature Genetics*, May 2021. URL: <https://doi.org/10.1038/s41588-021-00873-4>, doi:10.1038/s41588-021-00873-4.
6. Tianyi Sun, Dongyuan Song, Wei Vivian Li, and Jingyi Jessica Li. Scdesign2: a transparent simulator that generates high-fidelity single-cell gene expression count data with gene correlations captured. *Genome Biology*, 22(1):1–37, 2021.
7. Carolyn Shasha, Yuan Tian, Florian Mair, Helen E.R. Miller, and Raphael Gottardo. Superscan: supervised single-cell annotation. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/05/22/2021.05.20.445014>, arXiv:<https://www.biorxiv.org/content/early/2021/05/22/2021.05.20.445014.full.pdf>, doi:10.1101/2021.05.20.445014.

8. Massimo Andreatta, Jesus Corria-Osorio, Sören Müller, Rafael Cubas, George Coukos, and Santiago J. Carmona. Interpretation of T cell states from single-cell transcriptomics data using reference atlases. *Nature Communications*, 12(1):2965, May 2021. URL: <https://doi.org/10.1038/s41467-021-23324-4>, doi:10.1038/s41467-021-23324-4.
9. Jonathan R. Belyeu, Murad Chowdhury, Joseph Brown, Brent S. Pedersen, Michael J. Cormier, Aaron R. Quinlan, and Ryan M. Layer. Samplot: a platform for structural variant visual validation and automated filtering. *Genome Biology*, 22(1):161, May 2021. URL: <https://doi.org/10.1186/s13059-021-02380-5>, doi:10.1186/s13059-021-02380-5.
10. Lucille Lopez-Delisle and Jean-Baptiste Delisle. Baredsc: bayesian approach to retrieve expression distribution of single-cell. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/05/26/2021.05.26.445740.1>, arXiv:<https://www.biorxiv.org/content/early/2021/05/26/2021.05.26.445740.1.full.pdf>, doi:10.1101/2021.05.26.445740.
11. Seojin Bang, Pengtao Xie, Heewook Lee, Wei Wu, and Eric Xing. Explaining a black-box using deep variational information bottleneck approach. In *AAAI 2021*. AAAI, 2021. URL: <https://arxiv.org/pdf/1902.06918.pdf>.
12. Magda Markowska, Tomasz Cakała, Błażej Miasojedow, Dilafruz Juraeva, Johanna Mazur, Edith Ross, Eike Staub, and Ewa Szczurek. Conet: copy number event tree model of evolutionary tumor history for single-cell data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.04.23.441204v1>.
13. Chi-Yun Wu, Billy T Lau, Heon Seok Kim, Anuja Sathe, Susan M Grimes, Hanlee P Ji, and Nancy R Zhang. Integrative single-cell analysis of allele-specific copy number alterations and chromatin accessibility in cancer. *Nature Biotechnology*, pages 1–11, 2021. URL: <https://www.nature.com/articles/s41587-021-00911-w>.

2.4.4 2021-05-03

1. Jérémie Breda, Mihaela Zavolan, and Erik van Nimwegen. Bayesian inference of gene expression states from single-cell RNA-seq data. *Nature biotech*, 2021. URL: <https://www.nature.com/articles/s41587-021-00875-x>.
2. Akira Cortal, Loredana Martignetti, Emmanuelle Six, and Antonio Rausell. Cell-ID: gene signature extraction and cell identity recognition at individual cell level. *Nature biotech*, 2020. URL: <https://www.nature.com/articles/s41587-021-00896-6>.
3. Chao Gao, Jialin Liu, April R Kriebel, Sebastian Preissl, Chongyuan Luo, Rosa Castanon, Justin Sandoval, Angeline Rivkin, Joseph R Nery, Margarita M Behrens, and others. Iterative single-cell multi-omic integration using online learning. *Nature Biotechnology*, pages 1–8, 2021. URL: <https://www.nature.com/articles/s41587-021-00867-x>.
4. Osvaldo D Rivera, Michael J Mallory, Mathieu Quesnel-Vallières, Rakesh Chatrikhi, David C Schultz, Martin Carroll, Yoseph Barash, Sara Cherry, and Kristen W Lynch. Alternative splicing redefines landscape of commonly mutated genes in acute myeloid leukemia. *Proceedings of the National Academy of Sciences*, 2021. URL: <https://www.pnas.org/content/118/15/e2014967118>.
5. Chantierint-Andreas Kapourani, Ricard Argelaguet, Guido Sanguinetti, and Catalina A Vallejos. scMET: Bayesian modeling of DNA methylation heterogeneity at single-cell resolution. *Genome Biology*, 22(1):1–21, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02329-8>.
6. Akira Cortal, Loredana Martignetti, Emmanuelle Six, and Antonio Rausell. Cell-ID: gene signature extraction and cell identity recognition at individual cell level. *Nature biotech*, 2020. URL: <https://www.nature.com/articles/s41587-021-00896-6>.
7. Youjin Lee, Derek Bogdanoff, Yutong Wang, George C Hartoularos, Jonathan M Woo, Cody T Mowery, Hunter M Nisonoff, David S Lee, Yang Sun, James Lee, and others. Xyzeq: spatially resolved single-cell rna sequencing reveals expression heterogeneity in the tumor microenvironment. *Science Advances*, 7(17):eabg4755, 2021. URL: <https://advances.sciencemag.org/content/7/17/eabg4755?rss=1>.

8. Iain C Clark, Cristina Gutiérrez-Vázquez, Michael A Wheeler, Zhaorong Li, Veit Rothhammer, Mathias Linerbauer, Liliana M Sanmarco, Lydia Guo, Manon Blain, Stephanie EJ Zandee, and others. Barcoded viral tracing of single-cell interactions in central nervous system inflammation. *Science*, 2021. URL: <https://science.sciencemag.org/content/372/6540/eabf1230?rss=1>.
9. Alexander N. Gorelick, Minsoo Kim, Walid K. Chatila, Konnor La, A. Ari Hakimi, Michael F. Berger, Barry S. Taylor, Payam A. Gammage, and Ed Reznik. Respiratory complex and tissue lineage drive recurrent mutations in tumour mtDNA. *Nature Metabolism*, 3(4):558–570, April 2021. URL: <https://doi.org/10.1038/s42255-021-00378-8>, doi:10.1038/s42255-021-00378-8.
10. Tal Ashuach, Daniel A Reidenbach, Adam Gayoso, and Nir Yosef. Peakvi: a deep generative model for single cell chromatin accessibility analysis. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/04/30/2021.04.29.442020>, arXiv:<https://www.biorxiv.org/content/early/2021/04/30/2021.04.29.442020.full.pdf>, doi:10.1101/2021.04.29.442020.
11. Wendell Jones, Binsheng Gong, Natalia Novorodovskaya, Dan Li, Rebecca Kusko, Todd A. Richmond, Donald J. Johann, Halil Bisgin, Sayed Mohammad Ebrahim Sahraeian, Pierre R. Bushel, and et al. A verified genomic reference sample for assessing performance of cancer panels detecting small variants of low allele frequency. *Genome Biology*, Apr 2021. URL: <https://doi.org/10.1186/s13059-021-02316-z>.
12. Magda Markowska, Tomasz Cakała, Błażej Miasojedow, Dilafruz Juraeva, Johanna Mazur, Edith Ross, Eike Staub, and Ewa Szczurek. Conet: copy number event tree model of evolutionary tumor history for single-cell data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.04.23.441204v1.full>.
13. M Gordian Adam, Georg Beyer, Nicole Christiansen, Beate Kamlage, Christian Pilarsky, Marius Distler, Tim Falbusch, Ansgar Chromik, Fritz Klein, Marcus Bahra, and others. Identification and validation of a multivariable prediction model based on blood plasma and serum metabolomics for the distinction of chronic pancreatitis subjects from non-pancreas disease control subjects. *Gut*, 2021. URL: <https://gut.bmj.com/content/early/2021/02/18/gutjnl-2020-320723.abstract>.

2.4.5 2021-04-12

1. Qiao Liu, Jiaze Xu, Rui Jiang, and Wing Hung Wong. Density estimation using deep generative neural networks. *Proceedings of the National Academy of Sciences*, 2021. URL: <https://www.pnas.org/content/118/15/e2101344118>, arXiv:<https://www.pnas.org/content/118/15/e2101344118.full.pdf>, doi:10.1073/pnas.2101344118.
2. Daniel P Cooke, David C Wedge, and Gerton Lunter. A unified haplotype-based method for accurate and comprehensive variant calling. *Nature Biotechnology*, pages 1–8, 2021. URL: <https://www.nature.com/articles/s41587-021-00861-3>.
3. Surojit Biswas, Grigory Khimulya, Ethan C Alley, Kevin M Esvelt, and George M Church. Low-n protein engineering with data-efficient deep learning. *BioRxiv*, 2020. URL: <https://www.nature.com/articles/s41592-021-01100-y>.
4. Matteo Borella, Graziano Martello, Davide Risso, and Chiara Romualdi. PsiNorm: a scalable normalization for single-cell RNA-seq data. *BioRxiv*, 2021. URL: <https://doi.org/10.1101/2021.04.07.438822>, doi:10.1101/2021.04.07.438822.
5. Alexey Kolesnikov, Sidharth Goel, Maria Nattestad, Taedong Yun, Gunjan Baid, Howard Yang, Cory Y McLean, Pi-Chuan Chang, and Andrew Carroll. Deeptrio: variant calling in families using deep learning. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/04/06/2021.04.05.438434>, arXiv:<https://www.biorxiv.org/content/early/2021/04/06/2021.04.05.438434.full.pdf>, doi:10.1101/2021.04.05.438434.
6. Michael E Nelson, Simone G Riva, and Ann Cvejic. Smash: a scalable, general marker gene identification framework for single-cell rna sequencing and spatial transcriptomics. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/04/09/2021.04>.

08.438978, arXiv:<https://www.biorxiv.org/content/early/2021/04/09/2021.04.08.438978.full.pdf>, doi:10.1101/2021.04.08.438978.

7. Jonas A. Sibbesen, Jordan M. Eizenga, Adam M. Novak, Jouni Sirén, Xian Chang, Erik Garrison, and Benedict Paten. Haplotype-aware pantranscriptome analyses using spliced pangenome graphs. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/03/28/2021.03.26.437240>, arXiv:<https://www.biorxiv.org/content/early/2021/03/28/2021.03.26.437240.full.pdf>, doi:10.1101/2021.03.26.437240.
8. Zhiyuan Hu, Ahmed Ashour Ahmed, and Christopher Yau. An interpretable meta-clustering framework for single-cell rna-seq data integration and evaluation. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.03.29.437525v1.full>.
9. Zhe Wang, Shiyi Yang, Yusuke Koga, Sean E Corbett, W Evan Johnson, Masanao Yajima, and Joshua D Campbell. Celda: a bayesian model to perform bi-clustering of genes into modules and cells into subpopulations using single-cell rna-seq data. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.16.373274v2>.
10. Pablo Moreno, Ni Huang, Jonathan R Manning, Suhaib Mohammed, Andrey Solovyev, Krzysztof Polanski, Wendi Bacon, Ruben Chazarra, Carlos Talavera-López, Maria A Doyle, and others. User-friendly, scalable tools and workflows for single-cell rna-seq analysis. *Nature Methods*, pages 1–2, 2021. URL: <https://www.nature.com/articles/s41592-021-01102-w>.

2.4.6 2021-03-29

1. Wanze Chen, Orane Guillaume-Gentil, Riccardo Dainese, Pernille Yde Rainer, Magda Zachara, Christoph G. Gäbelein, Julia A. Vorholt, and Bart Deplancke. Genome-wide molecular recording using live-seq. *bioRxiv*, mar 2021. URL: <https://doi.org/10.1101/2021.03.24.436752>, doi:10.1101/2021.03.24.436752.
2. Juexin Wang, Anjun Ma, Yuzhou Chang, Jianting Gong, Yuexu Jiang, Ren Qi, Cankun Wang, Hongjun Fu, Qin Ma, and Dong Xu. scGNN is a novel graph neural network framework for single-cell RNA-seq analyses. *Nature Communications*, mar 2021. URL: <https://doi.org/10.1038/s41467-021-22197-x>, doi:10.1038/s41467-021-22197-x.
3. Christoffer Norn, Basile IM Wicky, David Juergens, Sirui Liu, David Kim, Doug Tischer, Brian Koepnick, Ivan Anishchenko, David Baker, and Sergey Ovchinnikov. Protein sequence design by conformational landscape optimization. *Proceedings of the National Academy of Sciences*, 2021. URL: <https://www.pnas.org/content/118/11/e2017228118>.
4. Yang Liu, Tao Wang, bin Zhou, and Deyou Zheng. Robust integration of multiple single-cell rna sequencing datasets using a single reference space. *Nature biotechnology*, 2021. URL: <https://www.nature.com/articles/s41587-021-00859-x>.
5. Karthik A Jagadeesh, Kushal K Dey, Daniel T Montoro, Steven Gazal, Jesse M Engreitz, Ramnik J Xavier, Alkes L Price, and Aviv Regev. Identifying disease-critical cell types and cellular processes across the human body by integration of single-cell profiles and human genetics. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.03.19.436212v1>.
6. He Zhao, Piyush Rai, Lan Du, Wray Buntine, Dinh Phung, and Mingyuan Zhou. Variational autoencoders for sparse and overdispersed discrete data. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, 1684–1694. PMLR, 26–28 Aug 2020. URL: <http://proceedings.mlr.press/v108/zhao20c.html>.
7. Suoqin Jin, Christian F Guerrero-Juarez, Lihua Zhang, Ivan Chang, Raul Ramos, Chen-Hsiang Kuan, Peggy Myung, Maksim V Plikus, and Qing Nie. Inference and analysis of cell-cell communication using cellchat. *Nature communications*, 12(1):1–20, 2021. URL: <https://www.nature.com/articles/s41467-021-21246-9>.
8. Daniel N. Baker, Nathan Dyjack, Vladimir Braverman, Stephanie C. Hicks, and Ben Langmead. Minicore: fast scRNA-seq clustering with various distances.

- bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/03/25/2021.03.24.436859>, arXiv:<https://www.biorxiv.org/content/early/2021/03/25/2021.03.24.436859.full.pdf>, doi:10.1101/2021.03.24.436859.
9. Tatsuhiko Naito, Ken Suzuki, Jun Hirata, Yoichiro Kamatani, Koichi Matsuda, Tatsushi Toda, and Yukinori Okada. A deep learning method for hla imputation and trans-ethnic mhc fine-mapping of type 1 diabetes. *Nature Communications*, 12(1):1–14, 2021.
 10. Galadriel Briere, Elodie Darbo, Patricia Thebault, and Raluca Uricaru. Consensus clustering applied to multi-omic disease subtyping. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/03/25/2020.10.19.345389>, arXiv:<https://www.biorxiv.org/content/early/2021/03/25/2020.10.19.345389.full.pdf>, doi:10.1101/2020.10.19.345389.
 11. Attila Gabor, Marco Tognetti, Alice Driessen, Jovan Tanevski, Baosen Guo, Wencai Cao, He Shen, Thomas Yu, Verena Chung, Bernd Bodenmiller, and Julio Saez-Rodriguez. Cell-to-cell and type-to-type heterogeneity of signaling networks: insights from the crowd. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/03/23/2021.03.23.436603>, arXiv:<https://www.biorxiv.org/content/early/2021/03/23/2021.03.23.436603.full.pdf>, doi:10.1101/2021.03.23.436603.
 12. Yonghan Yu, Lingxi Chen, Xinyao Miao, and Shuai Cheng Li. Spechap: a diploid phasing algorithm based on spectral graph theory. *bioRxiv*, pages 870972, 2021. URL: <https://www.biorxiv.org/content/10.1101/870972v2.full>.
 13. Fang Wang, Qihan Wang, Vakul Mohanty, Shaoheng Liang, Jinzhuang Dou, Jincheng Han, Darlan Conterno Minussi, Ruli Gao, Li Ding, Nicholas Navin, and Ken Chen. Medalt: single-cell copy number lineage tracing enabling gene discovery. *Genome Biology*, 22(1):70, Feb 2021. URL: <https://doi.org/10.1186/s13059-021-02291-5>, doi:10.1186/s13059-021-02291-5.
 14. Roy Rabbie, Doreen Lau, Richard M. White, and David J. Adams. Unraveling the cartography of the cancer ecosystem. *Genome Biology*, 22(1):87, Mar 2021. URL: <https://doi.org/10.1186/s13059-021-02310-5>, doi:10.1186/s13059-021-02310-5.
 15. Yanxiang Deng, Di Zhang, Yang Liu, Graham Su, Archibald Enniful, Zhiliang Bai, and Rong Fan. Spatial epigenome sequencing at tissue scale and cellular level. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.03.11.434985v1.full>.

2.4.7 2021-03-15

1. Marc Jan Bonder, Craig Smail, Michael J Gloudemans, Laure Frésard, David Jakubosky, Matteo D’Antonio, Xin Li, Nicole M Ferraro, Ivan Carcamo-Orive, Bogdan Mirauta, and others. Identification of rare and common regulatory variants in pluripotent cells using population-scale transcriptomics. *Nature genetics*, 53(3):313–321, 2021. URL: <https://www.nature.com/articles/s41588-021-00800-7>.
2. Julie Jerber, Daniel D Seaton, Anna SE Cuomo, Natsuhiko Kumasaka, James Haldane, Juliette Steer, Minal Patel, Daniel Pearce, Malin Andersson, Marc Jan Bonder, and others. Population-scale single-cell rna-seq profiling across dopaminergic neuron differentiation. *Nature genetics*, pages 1–9, 2021. URL: <https://www.nature.com/articles/s41588-021-00801-6>.
3. John-William Sidhom, H Benjamin Larman, Drew M Pardoll, and Alexander S Baras. Deeptcr: a deep learning framework for revealing structural concepts within tcr repertoire. *Nature communications*, 2021. URL: <https://www.nature.com/articles/s41467-021-21879-w>.
4. Pawel F. Przytycki and Katherine S. Pollard. Cellwalker integrates single-cell and bulk data to resolve regulatory elements across cell types in complex tissues. *Genome Biology*, 2021. URL: <https://doi.org/10.1186/s13059-021-02279-1>, doi:10.1186/s13059-021-02279-1.
5. Mor Nitzan and Michael P. Brenner. Revealing lineage-related signals in single-cell gene expression using random matrix theory. *Proceedings of the National Academy of Sciences*, 2021. URL: <https://www>.

- pnas.org/content/118/11/e1913931118, arXiv:<https://www.pnas.org/content/118/11/e1913931118.full.pdf>, doi:10.1073/pnas.1913931118.
6. Steffen Albrecht, Maximilian Sprang, Miguel A. Andrade-Navarro, and Jean-Fred Fontaine. Seqqscorer: automated quality control of next-generation sequencing data using machine learning. *Genome Biology*, 2021. URL: <https://doi.org/10.1186/s13059-021-02294-2>, doi:10.1186/s13059-021-02294-2.
 7. Eddie Park, Yan Jiang, Lili Hao, Jingyi Hui, and Yi Xing. Genetic variation and microrna targeting of a-to-i rna editing fine tune human tissue transcriptomes. *Genome Biology*, 22(1):77, Mar 2021. URL: <https://doi.org/10.1186/s13059-021-02287-1>, doi:10.1186/s13059-021-02287-1.
 8. Changcai Huang, Guangyu Li, Jiayu Wu, Junbo Liang, and Xiaoyue Wang. Identification of pathogenic variants in cancer genes using base editing screens with editing efficiency correction. *Genome Biology*, 22(1):80, Mar 2021. URL: <https://doi.org/10.1186/s13059-021-02305-2>, doi:10.1186/s13059-021-02305-2.
 9. Bianca Dumitrascu, Soledad Villar, Dustin G Mixon, and Barbara E Engelhardt. Optimal marker gene selection for cell type discrimination in single cell analyses. *Nature communications*, 12(1):1–8, 2021. URL: <https://www.nature.com/articles/s41467-021-21453-4>.
 10. Gennady Gorin, Valentine Svensson, and Lior Pachter. Protein velocity and acceleration from single-cell multi-omics experiments. *Genome biology*, 21(1):1–6, 2020.
 11. Alexander Binder, Michael Bockmayr, Miriam Hägele, Stephan Wienert, Daniel Heim, Katharina Hellweg, Masaru Ishii, Albrecht Stenzinger, Andreas Hocke, Carsten Denkert, and others. Morphological and molecular breast cancer profiling through explainable machine learning. *Nature Machine Intelligence*, pages 1–12, 2021.
 12. Lars Velten, Benjamin A Story, Pablo Hernández-Malmierca, Simon Raffel, Daniel R Leonce, Jennifer Milbank, Malte Paulsen, Aykut Demir, Chelsea Szu-Tu, Robert Frömel, and others. Identification of leukemic and pre-leukemic stem cells by clonal tracking from single-cell transcriptomics. *Nature communications*, 12(1):1–13, 2021. URL: <https://www.nature.com/articles/s41467-021-21650-1>.
 13. Daniel B Burkhardt, Jay S Stanley, Alexander Tong, Ana Luisa Perdigoto, Scott A Gigante, Kevan C Herold, Guy Wolf, Antonio J Giraldez, David van Dijk, and Smita Krishnaswamy. Quantifying the effect of experimental perturbations at single-cell resolution. *Nature Biotechnology*, pages 1–11, 2021.
 14. Sheng Wang, Emily R Flynn, and Russ B Altman. Gaussian embedding for large-scale gene set analysis. *Nature machine intelligence*, 2(7):387–395, 2020. URL: <https://1drv.ms/u/s!Asst2n-g8HI883ob82p2QH00vDTB>.
 15. Gryte Satas, Simone Zaccaria, Mohammed El-Kebir, and Benjamin J Raphael. Deciphering the elusive cancer cell fraction in tumor heterogeneity and evolution. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.02.27.429196v1.full>.
 16. Francisca Rojas Ringeling, Stefan Canzar, and others. Linear-time cluster ensembles of large-scale single-cell rna-seq and multimodal data. *Genome research*, pages gr–267906, 2021.
 17. Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre H Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Daniel Guo, Mohammad Gheshlaghi Azar, and others. Bootstrap your own latent: a new approach to self-supervised learning. *arXiv preprint arXiv:2006.07733*, 2020. URL: <https://arxiv.org/abs/2006.07733>.
 18. Xinlei Chen and Kaiming He. Exploring simple siamese representation learning. *arXiv preprint arXiv:2011.10566*, 2020. URL: <https://arxiv.org/abs/2011.10566>.

2.4.8 2021-03-01

1. Adam Gayoso, Zoë Steier, Romain Lopez, Jeffrey Regier, Kristopher L Nazor, Aaron Streets, and Nir Yosef. Joint probabilistic modeling of single-cell multi-omic data with totalvi. *Nature Methods*, pages 1–11, 2021. URL: <https://www.nature.com/articles/s41592-020-01050-x>.
2. Benjamin L Emert, Christopher J Cote, Eduardo A Torre, Ian P Dardani, Connie L Jiang, Naveen Jain, Sydney M Shaffer, and Arjun Raj. Variability within rare cell states enables multiple paths toward drug resistance. *Nature Biotechnology*, pages 1–12, 2021. URL: <https://www.nature.com/articles/s41587-021-00837-3>.
3. Dylan M Cable, Evan Murray, Luli S Zou, Aleksandrina Goeva, Evan Z Macosko, Fei Chen, and Rafael A Irizarry. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nature Biotechnology*, pages 1–10, 2021. URL: <https://www.nature.com/articles/s41587-021-00830-w>.
4. Fang Wang, Qihan Wang, Vakul Mohanty, Shaoheng Liang, Jinzhuang Dou, Jincheng Han, Darlan Conterno Minussi, Ruli Gao, Li Ding, Nicholas Navin, and others. Medalt: single-cell copy number lineage tracing enabling gene discovery. *Genome Biology*, 22(1):1–22, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02291-5>.
5. Hongyu Guo and Jun Li. Scsorter: assigning cells to known cell types according to marker genes. *Genome Biology*, 22(1):1–18, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02281-7>.
6. Dongfang Wang, Siyu Hou, Lei Zhang, Xiliang Wang, Baolin Liu, and Zemin Zhang. Imap: integration of multiple single-cell datasets by adversarial paired transfer networks. *Genome biology*, 22(1):1–24, 2021. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-021-02280-8>.
7. Stephan Schmeing and Mark D. Robinson. Reseq simulates realistic illumina high-throughput sequencing data. *Genome Biology*, 2021. doi:10.1186/s13059-021-02265-7.
8. Stephanie C. Hicks, Ruoxi Liu, Yuwei Ni, Elizabeth Purdom, and Davide Risso. Mbkmeans: fast clustering for single cell data using mini-batch k-means. *PLOS Computational Biology*, 17(1):1–18, 01 2021. URL: <https://doi.org/10.1371/journal.pcbi.1008625>, doi:10.1371/journal.pcbi.1008625.
9. Jacob M. Graving and Iain D. Couzin. Vae-sne: a deep generative model for simultaneous dimensionality reduction and clustering. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/early/2020/07/17/2020.07.17.207993>, arXiv:<https://www.biorxiv.org/content/early/2020/07/17/2020.07.17.207993.full.pdf>, doi:10.1101/2020.07.17.207993.
10. Tim Sainburg, Leland McInnes, and Timothy Q Gentner. Parametric umap: learning embeddings with deep neural networks for representation and semi-supervised learning. 2020. arXiv:2009.12981.
11. Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. Dynamic graph cnn for learning on point clouds. 2019. arXiv:1801.07829.
12. Anees Kazi, Luca Cosmo, Nassir Navab, and Michael Bronstein. Differentiable graph module (dgm) for graph convolutional networks. 2020. arXiv:2002.04999.
13. Luca Cosmo, Anees Kazi, Seyed-Ahmad Ahmadi, Nassir Navab, and Michael Bronstein. Latent patient network learning for automatic diagnosis. *Lecture Notes in Computer Science*, pages 643–653, 2020. URL: http://dx.doi.org/10.1007/978-3-030-59713-9_62, doi:10.1007/978-3-030-59713-9_62.
14. Žiga Avsec, Melanie Weilert, Avanti Shrikumar, Sabrina Krueger, Amr Alexandari, Khyati Dalal, Robin Fropf, Charles McAnany, Julien Gagneur, Anshul Kundaje, and others. Base-resolution models of transcription-factor binding reveal soft motif syntax. *Nature Genetics*, pages 1–13, 2021.
15. Hai Yang, Rui Chen, Dongdong Li, and Zhe Wang. Subtype-GAN: a deep learning approach for integrative cancer subtyping of multi-omics data. *Bioinformatics*, 02 2021. btab109. URL: <https://doi.org/10.1093/bioinformatics/btab109>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btab109/36411133/btab109.pdf>, doi:10.1093/bioinformatics/btab109.

16. Hongyu Guo and Jun Li. Scsorter: assigning cells to known cell types according to marker genes. *Genome Biology*, 22(1):69, Feb 2021. URL: <https://doi.org/10.1186/s13059-021-02281-7>, doi:10.1186/s13059-021-02281-7.
17. Bobby Ranjan, Florian Schmidt, Wenjie Sun, Jinyu Park, Mohammad Amin Honardoost, Joanna Tan, Nir-mala Arul Rayan, and Shyam Prabhakar. Scconsensus: combining supervised and unsupervised clustering for cell type identification in single-cell rna sequencing data. *bioRxiv*, 2020.
18. Akshat Gupta and others. Blind descent: a prequel to gradient descent. *arXiv preprint arXiv:2006.11505*, 2020.
19. Chitong Rao, Katharine Z Coyte, Wayne Bainter, Raif S Geha, Camilia R Martin, and Seth Rakoff-Nahoum. Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature*, pages 1–6, 2021. URL: <https://www.nature.com/articles/s41586-021-03241-8>.
20. Diana Mahdessian, Anthony J Cesnik, Christian Gnann, Frida Danielsson, Lovisa Stenström, Muhammad Arif, Cheng Zhang, Trang Le, Fredric Johansson, Rutger Shutten, and others. Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature*, 590(7847):649–654, 2021. URL: <https://www.nature.com/articles/s41586-021-03232-9>.
21. Jeffrey M. Granja, M. Ryan Corces, Sarah E. Pierce, S. Tansu Bagdatli, Hani Choudhry, Howard Y. Chang, and William J. Greenleaf. Archr is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature Genetics*, Feb 2021. URL: <https://doi.org/10.1038/s41588-021-00790-6>, doi:10.1038/s41588-021-00790-6.

2.4.9 2021-02-01

1. Brian Hie, Ellen D Zhong, Bonnie Berger, and Bryan Bryson. Learning the language of viral evolution and escape. *Science*, 371(6526):284–288, 2021. URL: <https://science.sciencemag.org/content/371/6526/284>.
2. Ashwin Narayan, Bonnie Berger, and Hyunghoon Cho. Assessing single-cell transcriptomic variability through density-preserving data visualization. *Nature Biotechnology*, pages 1–10, 2021. URL: <https://www.nature.com/articles/s41587-020-00801-7>.
3. Lyla Atta and Jean Fan. Veloviz: rna-velocity informed 2d embeddings for visualizing cellular trajectories. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.01.28.425293v1>.
4. Karren Dai Yang, Anastasiya Belyaeva, Saradha Venkatachalapathy, Karthik Damodaran, Abigail Katcoff, Adityanarayanan Radhakrishnan, GV Shivashankar, and Caroline Uhler. Multi-domain translation between single-cell imaging and sequencing data using autoencoders. *Nature Communications*, 12(1):1–10, 2021. URL: <https://www.nature.com/articles/s41467-020-20249-2>.
5. Brian Hie, Ellen D Zhong, Bonnie Berger, and Bryan Bryson. Learning the language of viral evolution and escape. *Science*, 371(6526):284–288, 2021. URL: <https://science.sciencemag.org/content/371/6526/284.abstract>.
6. Anna S.E. Cuomo, Giordano Alvari, Christina B. Azodi, Davis J. McCarthy, and Marc Jan Bonder. Optimising expression quantitative trait locus mapping workflows for single-cell studies. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/21/2021.01.20.427401>, arXiv:<https://www.biorxiv.org/content/early/2021/01/21/2021.01.20.427401.full.pdf>, doi:10.1101/2021.01.20.427401.
7. Yang Xu, Priyojit Das, and Rachel Patton McCord. Smile: mutual information learning for integration of single cell omics data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/29/2021.01.28.428619>, arXiv:<https://www.biorxiv.org/content/early/2021/01/29/2021.01.28.428619.full.pdf>, doi:10.1101/2021.01.28.428619.
8. Jian Yan, Yunjiang Qiu, André M Ribeiro Dos Santos, Yimeng Yin, Yang E Li, Nick Vinckier, Naoki Nariai, Paola Benaglio, Anugraha Raman, Xiaoyu Li, and others. Systematic analysis of binding of transcription factors to noncoding variants. *Nature*, pages 1–5, 2021. URL: <https://www.nature.com/articles/s41586-021-03211-0>.

9. An Zheng, Michael Lamkin, Hanqing Zhao, Cynthia Wu, Hao Su, and Melissa Gymrek. Deep neural networks identify sequence context features predictive of transcription factor binding. *Nature Machine Intelligence*, pages 1–9, 2021. URL: <https://www.biorxiv.org/content/10.1101/2020.02.26.965343v1.full>.
10. Oliver Hilsenbeck, Michael Schwarzfischer, Dirk Loeffler, Sotiris Dimopoulos, Simon Hastreiter, Carsten Marr, Fabian J Theis, and Timm Schroeder. Faster: a user-friendly tool for ultrafast and robust cell segmentation in large-scale microscopy. *Bioinformatics*, 33(13):2020–2028, 2017.
11. Charlotte Soneson, Avi Srivastava, Rob Patro, and Michael B Stadler. Preprocessing choices affect rna velocity results for droplet scrna-seq data. *PLOS Computational Biology*, 17(1):e1008585, 2021.
12. K. Charmpi, T. Guo, Q. Zhong, and et al. Convergent network effects along the axis of gene expression during prostate cancer progression. *BMC Genome Biology*, pages 21, 2021. URL: <https://doi.org/10.1186/s13059-020-02188-9>.
13. Vu V H Pham, Lin Liu, Cameron P Bracken, Gregory J Goodall, Jiuyong Li, and Thuc D Le. DriverGroup: a novel method for identifying driver gene groups. *Bioinformatics*, 36(Supplement_2):i583–i591, 12 2020. URL: <https://doi.org/10.1093/bioinformatics/btaa797>, arXiv:https://academic.oup.com/bioinformatics/article-pdf/36/Supplement_2/i583/35336898/btaa797.pdf, doi:10.1093/bioinformatics/btaa797.
14. Stefano Mangiola, Ramyar Molania, Ruining Dong, Maria A. Doyle, and Anthony T. Papenfuss. Tidybulk: an r tidy framework for modular transcriptomic data analysis. *Genome Biology*, 22(1):42, Jan 2021. URL: <https://doi.org/10.1186/s13059-020-02233-7>, doi:10.1186/s13059-020-02233-7.
15. Nam D Nguyen, Jiawei Huang, and Daifeng Wang. Deepmanreg: a deep manifold-regularized learning model for improving phenotype prediction from multi-modal data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/30/2021.01.28.428715>, arXiv:<https://www.biorxiv.org/content/early/2021/01/30/2021.01.28.428715.full.pdf>, doi:10.1101/2021.01.28.428715.
16. Nuha BinTayyash, Sokratia Georgaka, ST John, Sumon Ahmed, Alexis Boukouvalas, James Hensman, and Magnus Rattray. Non-parametric modelling of temporal and spatial counts data from rna-seq experiments. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/30/2020.07.29.227207>, arXiv:<https://www.biorxiv.org/content/early/2021/01/30/2020.07.29.227207.full.pdf>, doi:10.1101/2020.07.29.227207.
17. Zhanlin Chen, Jing Zhang, Jason Liu1, Zixuan Zhang, Jiangqi Zhu, Donghoon Lee, Min Xu, and Mark Gerstein. SCAN-ATAC-Sim: a scalable and efficient method for simulating single-cell ATAC-seq data from bulk-tissue experiments. *Bioinformatics*, 01 2021. btaa1039. URL: <https://doi.org/10.1093/bioinformatics/btaa1039>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa1039/35927179/btaa1039.pdf>, doi:10.1093/bioinformatics/btaa1039.
18. Jialin Liu, Chao Gao, Joshua Sodicoff, Velina Kozareva, Evan Z Macosko, and Joshua D Welch. Jointly defining cell types from multiple single-cell datasets using liger. *Nature Protocols*, 15(11):3632–3662, 2020. URL: <https://www.nature.com/articles/s41596-020-0391-8>.
19. Joshua D Welch, Velina Kozareva, Ashley Ferreira, Charles Vanderburg, Carly Martin, and Evan Z Macosko. Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell*, 177(7):1873–1887, 2019. URL: <https://www.sciencedirect.com/science/article/pii/S0092867419305045>.
20. Zi Yang and George Michailidis. A non-negative matrix factorization method for detecting modules in heterogeneous omics multi-modal data. *Bioinformatics*, 32(1):1–8, 2016. URL: <https://doi.org/10.1093/bioinformatics/btv544>.

2.4.10 2021-01-18

1. Rens van de Schoot, Sarah Depaoli, Ruth King, Bianca Kramer, Kaspar Märtens, Mahlet G. Tadesse, Marina Vannucci, Andrew Gelman, Duco Veen, Joukje Willemsen, and Christopher Yau. Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1):1–26, 2021. URL: <https://www.nature.com/articles/s43586-020-00001-2>, doi:10.1038/s43586-020-00001-2.
2. Ze Zhang, Danyi Xiong, Xinlei Wang, Hongyu Liu, and Tao Wang. Mapping the functional landscape of t cell receptor repertoires by single-t cell transcriptomics. *Nature Methods*, 18(1):92–99, 2021. URL: <https://www.nature.com/articles/s41592-020-01020-3>.
3. Ronen Sadeh, Israa Sharkia, Gavriel Fialkoff, Ayelet Rahat, Jenia Gutin, Alon Chappleboim, Mor Nitzan, Ilana Fox-Fisher, Daniel Neiman, Guy Meler, and others. Chip-seq of plasma cell-free nucleosomes identifies gene expression programs of the cells of origin. *Nature Biotechnology*, pages 1–13, 2021. URL: <https://www.nature.com/articles/s41587-020-00775-6>.
4. Ruiping Wang, Minghao Dang, Kazuto Harada, Guangchun Han, Fang Wang, Melissa Pool Pizzi, Meina Zhao, Ghia Tatlonghari, Shaojun Zhang, Dapeng Hao, and others. Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nature Medicine*, pages 1–11, 2021. URL: <https://www.nature.com/articles/s41591-020-1125-8>.
5. David Fawcner-Corbett, Agne Antanaviciute, Kaushal Parikh, Marta Jagielowicz, Ana Sousa Gerós, Tarun Gupta, Neil Ashley, Doran Khamis, Darren Fowler, Edward Morrissey, and others. Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell*, 2021. URL: [https://www.cell.com/cell/fulltext/S0092-8674\(20\)31686-X](https://www.cell.com/cell/fulltext/S0092-8674(20)31686-X).
6. Jieun Hani Kim, Kevin YX Wang, Carissa Chen, Yingxin Lin, Patrick PL Tam, Dave M Lin, Jean Yee Hwa Yang, and Pengyi Yang. Cepo uncovers cell identity through differential stability. *bioRxiv*, pages 2021–01, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.01.10.426138v1>.
7. Yingxin Lin, Tung-Yu Wu, Sheng Wan, Jean YH Yang, YX Rachel Wang, and Wing H Wong. Scjoint: transfer learning for data integration of single-cell rna-seq and atac-seq. *bioRxiv*, pages 2020–12, 2021. URL: <https://www.biorxiv.org/content/10.1101/2020.12.31.424916v1>.
8. Luke Zappia and Alicia Oshlack. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *GigaScience*, 07 2018. giy083. URL: <https://doi.org/10.1093/gigascience/giy083>, arXiv:<https://academic.oup.com/gigascience/article-pdf/7/7/giy083/25205178/giy083.pdf>, doi:10.1093/gigascience/giy083.
9. Yinghao Cao, Xiaoyue Wang, and Gongxin Peng. Scsa: a cell type annotation tool for single-cell rna-seq data. *Frontiers in Genetics*, 11:490, 2020. URL: <https://www.frontiersin.org/article/10.3389/fgene.2020.00490>, doi:10.3389/fgene.2020.00490.
10. Sagi Abelson, Andy G. X. Zeng, Ido Nofech-Mozes, Ting Ting Wang, Stanley W. K. Ng, Mark D. Minden, Trevor J. Pugh, Philip Awadalla, Liran I. Shlush, Tracy Murphy, Steven M. Chan, John E. Dick, and Scott V. Bratman. Integration of intra-sample contextual error modeling for improved detection of somatic mutations from deep sequencing. *Science Advances*, 2020. URL: <https://advances.sciencemag.org/content/6/50/eabe3722>, arXiv:<https://advances.sciencemag.org/content/6/50/eabe3722.full.pdf>, doi:10.1126/sciadv.abe3722.
11. Yuqi Sheng, Ying Jiang, Yang Yang, Xiangmei Li, Jiayue Qiu, Jiashuo Wu, Liang Cheng, and Junwei Han. CNA2Subpathway: identification of dysregulated subpathway driven by copy number alterations in cancer. *Briefings in Bioinformatics*, 01 2021. bbaa413. URL: <https://doi.org/10.1093/bib/bbaa413>, arXiv:<https://academic.oup.com/bib/advance-article-pdf/doi/10.1093/bib/bbaa413/35590234/bbaa413.pdf>, doi:10.1093/bib/bbaa413.
12. Xian F. Mallory, Mohammadamin Edrisi, Nicholas Navin, and Luay Nakhleh. Methods for copy number aberration detection from single-cell DNA-sequencing data. *Genome Biology*, 21(1):208, August 2020. URL: <https://doi.org/10.1186/s13059-020-02119-8> (visited on 2021-01-18), doi:10.1186/s13059-020-02119-8.

13. Nae-Chyun Chen, Alexey Kolesnikov, Sidharth Goel, Taedong Yun, Pi-Chuan Chang, and Andrew Carroll. Improving variant calling using population data and deep learning. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/07/2021.01.06.425550>, arXiv:<https://www.biorxiv.org/content/early/2021/01/07/2021.01.06.425550.full.pdf>, doi:10.1101/2021.01.06.425550.
14. Bas van Opheusden, Luigi Acerbi, and Wei Ji Ma. Unbiased and efficient log-likelihood estimation with inverse binomial sampling. *PLOS Computational Biology*, 16(12):1–32, 12 2020. URL: <https://doi.org/10.1371/journal.pcbi.1008483>, doi:10.1371/journal.pcbi.1008483.
15. Matthew G Durrant and Ami S Bhatt. Automated prediction and annotation of small open reading frames in microbial genomes. *Cell Host & Microbe*, 2020.
16. Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41592-020-01018-x>.
17. Ali Mahdipour-Shirayeh, Natalie Erdmann, Chungyee Leung-Hagesteijn, and Rodger E Tiedemann. Scicnv: high-throughput paired profiling of transcriptomes and dna copy number variations at single cell resolution. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.02.10.942607v2.full>.
18. Stefan Canzar and others. A generalization of t-sne and umap to single-cell multimodal omics. *bioRxiv*, pages 2021–01, 2021. URL: <https://www.biorxiv.org/content/10.1101/2021.01.10.426098v1.full>.
19. Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: transformers for image recognition at scale. In *International Conference on Learning Representations*. 2021. URL: <https://openreview.net/forum?id=YicbFdNTTy>.
20. Yifan Zhao, Huiyu Cai, Zuobai Zhang, Jian Tang, and Yue Li. Learning interpretable cellular and gene signature embeddings from single-cell transcriptomic data. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/15/2021.01.13.426593>, arXiv:<https://www.biorxiv.org/content/early/2021/01/15/2021.01.13.426593.full.pdf>, doi:10.1101/2021.01.13.426593.
21. Ziqi Zhang and Xiuwei Zhang. Velosim: simulating single cell gene-expression and rna velocity. *bioRxiv*, 2021. URL: <https://www.biorxiv.org/content/early/2021/01/13/2021.01.11.426277>, arXiv:<https://www.biorxiv.org/content/early/2021/01/13/2021.01.11.426277.full.pdf>, doi:10.1101/2021.01.11.426277.
22. Hyo Young Choi, Heejoon Jo, Xiaobei Zhao, Katherine A. Hoadley, Scott Newman, Jeremiah Holt, Michele C. Hayward, Michael I. Love, J. S. Marron, and D. Neil Hayes. SCISSOR: a framework for identifying structural changes in RNA transcripts. *Nature Communications*, January 2021. URL: <https://doi.org/10.1038/s41467-020-20593-3>, doi:10.1038/s41467-020-20593-3.

2.4.11 2020-12-14

1. Shilpa Garg, Arkarachai Fungtammasan, Andrew Carroll, Mike Chou, Anthony Schmitt, Xiang Zhou, Stephen Mac, Paul Peluso, Emily Hatas, Jay Ghurye, and others. Chromosome-scale, haplotype-resolved assembly of human genomes. *Nature Biotechnology*, pages 1–4, 2020. URL: <https://www.nature.com/articles/s41587-020-0711-0>.
2. Robert R Stickels, Evan Murray, Pawan Kumar, Jilong Li, Jamie L Marshall, Daniela J Di Bella, Paola Arlotta, Evan Z Macosko, and Fei Chen. Highly sensitive spatial transcriptomics at near-cellular resolution with slide-seq2. *Nature Biotechnology*, 2020. URL: <https://www.nature.com/articles/s41587-020-0739-1>.
3. Jonathan R Bowles, Caroline Hoppe, Hilary L Ashe, and Magnus Rattray. Scalable inference of transcriptional kinetic parameters from ms2 time series data. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.12.04.412049v1>.

4. Benjamin Chidester, Tianming Zhou, and Jian Ma. Spicemix: integrative single-cell spatial modeling for inferring cell identity. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.29.383067v1>.
5. Duncan K. Ralph and Frederick A. Matsen. Using b cell receptor lineage structures to predict affinity. *for Public Library of Science (PLOS)*, 2020. doi:10.1101/2020.04.24.060194.
6. Teresa M.r. Noviello, Michele Ceccarelli, and Luigi Cerulo. Deep learning predicts non-coding rna functions from only raw sequence data. *for Public Library of Science (PLOS)*, 2020. doi:10.1101/2020.05.27.118778.
7. Simon M. Hoser. Intronic trnas of mitochondrial origin regulate constitutive and alternative splicing. *BMC Genome Biology*, 2020. doi:10.26226/morressier.5ebd45acffea6f735881aeb6.
8. Shuai He, Lin-He Wang, Yang Liu, Yi-Qi Li, Hai-Tian Chen, Jing-Hong Xu, Wan Peng, Guo-Wang Lin, Pan-Pan Wei, Bo Li, and et al. Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs. *BMC Genome Biology*, 2020. doi:10.1101/2020.03.18.996975.
9. Quang-Huy Nguyen and Duc-Hau Le. Improving existing analysis pipeline to identify and analyze cancer driver genes using multi-omics data. *Scientific Reports*, 10(1):20521, Nov 2020. URL: <https://doi.org/10.1038/s41598-020-77318-1>, doi:10.1038/s41598-020-77318-1.
10. Stefano Avanzini, David M. Kurtz, Jacob J. Chabon, Everett J. Moding, Sharon Seiko Hori, Sanjiv Sam Gambhir, Ash A. Alizadeh, Maximilian Diehn, and Johannes G. Reiter. A mathematical model of ctdna shedding predicts tumor detection size. *Science Advances*, 2020. URL: <https://advances.sciencemag.org/content/6/50/eabc4308>, arXiv:<https://advances.sciencemag.org/content/6/50/eabc4308.full.pdf>, doi:10.1126/sciadv.abc4308.
11. Tallulah S Andrews, Vladimir Yu Kiselev, Davis McCarthy, and Martin Hemberg. Tutorial: guidelines for the computational analysis of single-cell rna sequencing data. *Nature Protocols*, 2020. URL: <https://www.nature.com/articles/s41596-020-00409-w>.
12. Lifei Wang, Rui Nie, Zeyang Yu, Ruyue Xin, Caihong Zheng, Zhang Zhang, Jiang Zhang, and Jun Cai. An interpretable deep-learning architecture of capsule networks for identifying cell-type gene expression programs from single-cell rna-sequencing data. *Nature Machine Intelligence*, pages 1–11, 2020. URL: <https://www.nature.com/articles/s42256-020-00244-4>.
13. Jan Lause, Philipp Berens, and Dmitry Kobak. Analytic pearson residuals for normalization of single-cell rna-seq umi data. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/early/2020/12/02/2020.12.01.405886>, arXiv:<https://www.biorxiv.org/content/early/2020/12/02/2020.12.01.405886.full.pdf>, doi:10.1101/2020.12.01.405886.
14. Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. Nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, Dec 2020. URL: <https://doi.org/10.1038/s41592-020-01008-z>, doi:10.1038/s41592-020-01008-z.
15. Anonymous. Capturing label characteristics in VAEs. In *Submitted to International Conference on Learning Representations*. 2021. under review. URL: <https://openreview.net/forum?id=wQRISUZ5V7B>.
16. Anonymous. Sparse gaussian process variational autoencoders. In *Submitted to International Conference on Learning Representations*. 2021. under review. URL: <https://openreview.net/forum?id=czv8Ac3Kg7L>.
17. Xin Jin, Zelalem Demere, Karthik Nair, Ahmed Ali, Gino B. Ferraro, Ted Natoli, Amy Deik, Lia Petronio, Andrew A. Tang, Cong Zhu, Li Wang, Danny Rosenberg, Vamsi Mangena, Jennifer Roth, Kwanghun Chung, Rakesh K. Jain, Clary B. Clish, Matthew G. Vander Heiden, and Todd R. Golub. A metastasis map of human cancer cell lines. *Nature*, 588(7837):331–336, Dec 2020. URL: <https://doi.org/10.1038/s41586-020-2969-2>, doi:10.1038/s41586-020-2969-2.

2.4.12 2020-11-30

1. Ameen A Salahudeen, Shannon S Choi, Arjun Rustagi, Junjie Zhu, Vincent van Unen, M Sean, Ryan A Flynn, Mar Margalef-Català, António JM Santos, Jihang Ju, and others. Progenitor identification and sars-cov-2 infection in human distal lung organoids. *Nature*, pages 1–9, 2020. URL: <https://www.nature.com/articles/s41586-020-3014-1>.
2. Jacob Schreiber, Ritambhara Singh, Jeffrey Bilmes, and William Stafford Noble. A pitfall for machine learning methods aiming to predict across cell types. *Genome biology*, 21(1):1–6, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02177-y>.
3. Emma Dann, Neil C Henderson, Sarah A Teichmann, Michael D Morgan, and John C Marioni. Milo: differential abundance testing on single-cell data using k-nn graphs. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.23.393769v1>.
4. Tim Lohoff, Shila Ghazanfar, Alsu Missarova, Noushin Koulana, Nico Pierson, Jonathan A Griffiths, Evan S Bardot, Chee-Huat Linus Eng, Richard CV Tyser, Ricard Argelaguet, and others. Highly multiplexed spatially resolved gene expression profiling of mouse organogenesis. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.20.391896v1>.
5. Vitalii Kleshchevnikov, Artem Shmatko, Emma Dann, Alexander Aivazidis, Hamish W King, Tong Li, Artem Lomakin, Veronika Kedlian, Mika Sarkin Jain, Jun Sung Park, and others. Comprehensive mapping of tissue cell architecture via integrated single cell and spatial transcriptomics. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.15.378125v1>.
6. Jacob Hepkema, Nicholas Keone Lee, Benjamin J. Stewart, Siwat Ruangroengkulrith, Varodom Charoen-sawan, Menna R. Clatworthy, and Martin Hemberg. Predicting the impact of sequence motifs on gene regulation using single-cell data. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/early/2020/11/27/2020.11.26.400218>, arXiv:<https://www.biorxiv.org/content/early/2020/11/27/2020.11.26.400218.full.pdf>, doi:10.1101/2020.11.26.400218.
7. Apostolia M. Tsimberidou, Elena Fountzilas, Leonidas Bleris, and Razelle Kurzrock. Transcriptomics and solid tumors: the next frontier in precision cancer medicine. *Seminars in Cancer Biology*, 2020. URL: <http://www.sciencedirect.com/science/article/pii/S1044579X20301966>, doi:<https://doi.org/10.1016/j.semcancer.2020.09.007>.
8. Tatiana A. Gurbich and Valery Vladimirovich Ilinsky. Classifycnv: a tool for clinical annotation of copy-number variants. *Scientific Reports*, 10(1):20375, Nov 2020. URL: <https://doi.org/10.1038/s41598-020-76425-3>, doi:10.1038/s41598-020-76425-3.
9. Xuanhua P. Xie, Dan R. Laks, Daochun Sun, Asaf Poran, Ashley M. Laughney, Zilai Wang, Jessica Sam, German Belenguer, Isabel Fariñas, Olivier Elemento, Xiuping Zhou, and Luis F. Parada. High-resolution mouse subventricular zone stem-cell niche transcriptome reveals features of lineage, anatomy, and aging. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/11/20/2014389117>, arXiv:<https://www.pnas.org/content/early/2020/11/20/2014389117.full.pdf>, doi:10.1073/pnas.2014389117.
10. someone. Identification of a novel cancer stem cell subpopulation that promotes progression of human fatal renal cell carcinoma by single-cell rna-seq analysis. *International journal of biological sciences*, 16(16):3149–3162, 2020.
11. Sushant Kumar, Arif Harmanci, Jagath Vytheeswaran, and Mark B. Gerstein. Svfx: a machine learning framework to quantify the pathogenicity of structural variants. *Genome Biology*, 21(1):1–274, 2020.
12. Andrew E Teschendorff, Alok K Maity, Xue Hu, Chen Weiyan, and Matthias Lechner. Ultra-fast scalable estimation of single-cell differentiation potency from scRNA-Seq data. *Bioinformatics*, 11 2020. btaa987. URL: <https://doi.org/10.1093/bioinformatics/btaa987>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa987/34557516/btaa987.pdf>, doi:10.1093/bioinformatics/btaa987.
13. Anonymous. Fully unsupervised diversity denoising with convolutional variational autoencoders. In *Submitted to International Conference on Learning Representations*. 2021. under review. URL: <https://openreview.net/>

forum?id=agHLCOBM5jP.

14. Pacôme Prompsy, Pia Kirchmeier, Justine Marsolier, Marc Deloger, Nicolas Servant, and Céline Vallot. Interactive analysis of single-cell epigenomic landscapes with chromscape. *Nature Communications*, 11(1):1–9, 2020. URL: <https://www.nature.com/articles/s41467-020-19542-x>.
15. Jae Won Yun, Lixing Yang, Hye-Young Park, Chang-Woo Lee, Hongui Cha, Hyun-Tae Shin, Ka-Won Noh, Yoon-La Choi, Woong-Yang Park, and Peter J Park. Dysregulation of cancer genes by recurrent intergenic fusions. *Genome biology*, 21(1):1–20, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02076-2>.
16. Noam Bar, Tal Korem, Omer Weissbrod, David Zeevi, Daphna Rothschild, Sigal Leviatan, Noa Kosower, Maya Lotan-Pompan, Adina Weinberger, Caroline I Le Roy, and others. A reference map of potential determinants for the human serum metabolome. *Nature*, pages 1–6, 2020.
17. Junha Cha and Insuk Lee. Single-cell network biology for resolving cellular heterogeneity in human diseases. *Experimental & Molecular Medicine*, Nov 2020. URL: <https://doi.org/10.1038/s12276-020-00528-0>, doi:10.1038/s12276-020-00528-0.
18. Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A Wichmann. Shortcut learning in deep neural networks. *arXiv preprint arXiv:2004.07780*, 2020.
19. Anthony Culos, Amy S Tsai, Natalie Stanley, Martin Becker, Mohammad S Ghaemi, David R McIlwain, Ramin Fallahzadeh, Athena Tanada, Huda Nassar, Camilo Espinosa, and others. Integration of mechanistic immunological knowledge into a machine learning pipeline improves predictions. *Nature Machine Intelligence*, pages 1–10, 2020.

2.4.13 2020-11-16

1. Alexia Giannoula, Emilio Centeno, Miguel-Angel Mayer, Ferran Sanz, and Laura I Furlong. A system-level analysis of patient disease trajectories based on clinical, phenotypic and molecular similarities. *Bioinformatics*, 11 2020. btaa964. URL: <https://doi.org/10.1093/bioinformatics/btaa964>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa964/34297863/btaa964.pdf>, doi:10.1093/bioinformatics/btaa964.
2. Santiago Codesido, Mohamed Hanafi, Yoric Gagnebin, Víctor González-Ruiz, Serge Rudaz, and Julien Boccard. Network principal component analysis: a versatile tool for the investigation of multigroup and multiblock datasets. *Bioinformatics*, 11 2020. btaa954. URL: <https://doi.org/10.1093/bioinformatics/btaa954>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa954/34178616/btaa954.pdf>, doi:10.1093/bioinformatics/btaa954.
3. Yang Liu, Mingyu Yang, Yanxiang Deng, Graham Su, Cindy Guo, Di Zhang, Dongjoo Kim, Zhiliang Bai, Yang Xiao, Stephanie Halene, and Rong Fan. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell*, 2020. URL: [https://www.cell.com/cell/fulltext/S0092-8674\(20\)31390-8](https://www.cell.com/cell/fulltext/S0092-8674(20)31390-8).
4. Britta Velten, Jana Muriel Braunger, Damien Arnol, Ricard Argelaguet, and Oliver Stegle. Identifying temporal and spatial patterns of variation from multi-modal data using mefisto. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.03.366674v1>.
5. Lukas M Weber, Ariel A Hippen, Peter F Hickey, Kristofer C Berrett, Jason Gertz, Jennifer Anne Doherty, Casey S Greene, and Stephanie C Hicks. Genetic demultiplexing of pooled single-cell rna-sequencing samples in cancer facilitates effective experimental design. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.11.06.371963v1>.
6. Erick Armingol, Adam Officer, Olivier Harismendy, and Nathan E Lewis. Deciphering cell-cell interactions and communication from gene expression. *Nature Reviews Genetics*, pages 1–18, 2020. URL: <https://www.nature.com/articles/s41576-020-00292-x>.

7. Karsten Borgwardt, Elisabetta Ghisu, Felipe Llinares-López, Leslie O’Bray, and Bastian Rieck. Graph kernels: state-of-the-art and future challenges. *arXiv preprint arXiv:2011.03854*, 2020. URL: <https://arxiv.org/abs/2011.03854>.
8. someone. V(dd)j recombination is an important and evolutionarily conserved mechanism for generating antibodies with unusually long cdr3s. *Genome research*, 30(11):1547–1558, 2020.
9. someone. Rna timestamps identify the age of single molecules in rna sequencing. *Nature biotechnology*, 2020.
10. someone. Svfx: a machine learning framework to quantify the pathogenicity of structural variants. *Genome Biology*, 21(1):274, Nov 2020. URL: <https://doi.org/10.1186/s13059-020-02178-x>.
11. someone. Multiomics profiling of primary lung cancers and distant metastases reveals immunosuppression as a common characteristic of tumor cells with metastatic plasticity. *Genome Biology*, 21(1):271, Nov 2020. URL: <https://doi.org/10.1186/s13059-020-02175-0>.
12. Lifei Wang, Rui Nie, Zeyang Yu, Ruyue Xin, Caihong Zheng, Zhang Zhang, Jiang Zhang, and Jun Cai. An interpretable deep-learning architecture of capsule networks for identifying cell-type gene expression programs from single-cell rna-sequencing data. *Nature Machine Intelligence*, pages 1–11, 2020. URL: <https://www.pnas.org/content/early/2020/10/29/2005990117>.
13. Ingrid C Romero, Shu Kong, Charless C Fowlkes, Carlos Jaramillo, Michael A Urban, Francisca Obob-Ikuenobe, Carlos D’Apolito, and Surangi W Punyasena. Improving the taxonomy of fossil pollen using convolutional neural networks and superresolution microscopy. *Proceedings of the National Academy of Sciences*, 117(45):28496–28505, 2020. URL: <https://www.pnas.org/content/117/45/28496>.
14. Junyue Cao, Diana R. O’textquoteright Day, Hannah A. Pliner, Paul D. Kingsley, Mei Deng, Riza M. Daza, Michael A. Zager, Kimberly A. Aldinger, Ronnie Blecher-Gonen, Fan Zhang, Malte Spielmann, James Palis, Dan Doherty, Frank J. Steemers, Ian A. Glass, Cole Trapnell, and Jay Shendure. A human cell atlas of fetal gene expression. *Science*, 2020. URL: <https://science.sciencemag.org/content/370/6518/eaba7721>, [arXiv:https://science.sciencemag.org/content/370/6518/eaba7721.full.pdf](https://science.sciencemag.org/content/370/6518/eaba7721.full.pdf), doi:10.1126/science.aba7721.
15. Silvia Domcke, Andrew J. Hill, Riza M. Daza, Junyue Cao, Diana R. O’textquoteright Day, Hannah A. Pliner, Kimberly A. Aldinger, Dmitry Pokholok, Fan Zhang, Jennifer H. Milbank, Michael A. Zager, Ian A. Glass, Frank J. Steemers, Dan Doherty, Cole Trapnell, Darren A. Cusanovich, and Jay Shendure. A human cell atlas of fetal chromatin accessibility. *Science*, 2020. URL: <https://science.sciencemag.org/content/370/6518/eaba7612>, [arXiv:https://science.sciencemag.org/content/370/6518/eaba7612.full.pdf](https://science.sciencemag.org/content/370/6518/eaba7612.full.pdf), doi:10.1126/science.aba7612.
16. Junil Kim, Simon T. Jakobsen, Kedar N Natarajan, and Kyoung-Jae Won. TENET: gene network reconstruction using transfer entropy reveals key regulatory factors from single cell transcriptomic data. *Nucleic Acids Research*, 11 2020. gkaa1014. URL: <https://doi.org/10.1093/nar/gkaa1014>, [arXiv:https://academic.oup.com/nar/advance-article-pdf/doi/10.1093/nar/gkaa1014/34237557/gkaa1014.pdf](https://academic.oup.com/nar/advance-article-pdf/doi/10.1093/nar/gkaa1014/34237557/gkaa1014.pdf), doi:10.1093/nar/gkaa1014.

2.4.14 2020-11-02

1. Gabriela S Kinker, Alissa C Greenwald, Rotem Tal, Zhanna Orlova, Michael S Cuoco, James M McFarland, Allison Warren, Christopher Rodman, Jennifer A Roth, Samantha A Bender, and others. Pan-cancer single cell rna-seq uncovers recurring programs of cellular heterogeneity. *Nature Genetics*, pages 807552, 2019.
2. Paula Nieto, Marc Elosua-Bayes, Juan L. Trincado, Domenica Marchese, Ramon Massoni-Badosa, Maria Salvany, Ana Henriques, Elisabetta Mereu, Catia Moutinho, Sara Ruiz, Patricia Lorden, Vanessa T. Chin, Dominik Kaczorowski, Chia-Ling Chan, Richard Gallagher, Angela Chou, Ester Planas-Rigol, Carlota Rubio-Perez, Ivo Gut, Josep M. Piulats, Joan Seoane, Joseph E. Powell, Eduard Batlle, and Holger Heyn. A single-cell tumor immune atlas for precision oncology. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/early/2020/10/26/2020.10.26.354829>, [arXiv:https://www.biorxiv.org/content/early/2020/10/26/2020.10.26.354829.full.pdf](https://www.biorxiv.org/content/early/2020/10/26/2020.10.26.354829.full.pdf), doi:10.1101/2020.10.26.354829.

3. Linde A Miles, Robert L Bowman, Tiffany R Merlinsky, Isabelle S Csete, Aik Ooi, Robert Durruthy-Durruthy, Michael Bowman, Christopher Famulare, Minal A Patel, Pedro Mendez, and others. Single-cell mutation analysis of clonal evolution in myeloid malignancies. *Nature*, 2020. URL: <https://www.nature.com/articles/s41586-020-2864-x>.
4. Kiyomi Morita, Feng Wang, Katharina Jahn, Jack Kuipers, Yuanqing Yan, Jairo Matthews, Latasha Little, Curtis Gumbs, Shujuan Chen, Jianhua Zhang, and others. Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. *Nature communications*, 2020. URL: <https://www.nature.com/articles/s41467-020-19119-8>.
5. C.R. Nowosad, L. Mesin, T.B.R. Castro, and others. Tunable dynamics of b cell selection in gut germinal centres. *Nature*, 2020. URL: <https://www.nature.com/articles/s41586-020-2865-9>.
6. Marius Lange, Volker Bergen, Michal Klein, Manu Setty, Bernhard Reuter, Mostafa Bakhti, Heiko Lickert, Meshal Ansari, Janine Schniering, Herbert B Schiller, and others. Cellrank for directed single-cell fate mapping. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.10.19.345983v1>.
7. Edward Zhao, Matthew R Stone, Xing Ren, Thomas Pulliam, Paul Nghiem, Jason H Bielas, and Raphael Gottardo. Bayesspace enables the robust characterization of spatial gene expression architecture in tissue sections at increased resolution. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.09.04.283812v1>.
8. Viktor Petukhov, Konstantin Khodosevich, Ruslan A Soldatov, and Peter V Kharchenko. Bayesian segmentation of spatially resolved transcriptomics data. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.10.05.326777v1>.
9. Kentaro Kawata, Hiroyasu Wakida, Toshimichi Yamada, Kenzui Taniue, Han Han, Masahide Seki, Yutaka Suzuki, and Nobuyoshi Akimitsu. Metabolic labeling of rna using multiple ribonucleoside analogs enables the simultaneous evaluation of rna synthesis and degradation rates. *Genome research*, 30(10):gr.264408.120–1491, 2020.
10. Tracey W. Chan, Ting Fu, Jae Hoon Bahn, Hyun-Ik Jun, Jae-Hyung Lee, Giovanni Quinones-Valdez, Chonghui Cheng, and Xinshu Xiao. Rna editing in cancer impacts mrna abundance in immune response pathways. *Genome Biology*, 2020. doi:10.1186/s13059-020-02171-4.
11. Shengqing Stan Gu, Xiaoqing Wang, Xihao Hu, Peng Jiang, Ziyi Li, Nicole Traugh, Xia Bu, Qin Tang, Chenfei Wang, Zexian Zeng, and et al. Clonal tracing reveals diverse patterns of response to immune checkpoint blockade. *Genome Biology*, 2020. doi:10.1186/s13059-020-02166-1.
12. Krzysztof Choromanski, Valerii Likhoshesterov, David Dohan, Xingyou Song, Andreea Gane, Tamas Sarlos, Peter Hawkins, Jared Davis, Afroz Mohiuddin, Lukasz Kaiser, and others. Rethinking attention with performers. *arXiv preprint arXiv:2009.14794*, 2020. URL: <https://arxiv.org/abs/2009.14794>.
13. Thin Nguyen, Hang Le, and Svetha Venkatesh. Graphdta: prediction of drug–target binding affinity using graph convolutional networks. *Bioinformatics*, 2019. URL: <https://www.biorxiv.org/content/early/2019/06/27/684662>, arXiv:<https://www.biorxiv.org/content/early/2019/06/27/684662.full.pdf>, doi:10.1101/684662.
14. Masaki Asada, Makoto Miwa, and Yutaka Sasaki. Using Drug Descriptions and Molecular Structures for Drug-Drug Interaction Extraction from Literature. *Bioinformatics*, 10 2020. btaa907. URL: <https://doi.org/10.1093/bioinformatics/btaa907>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa907/34012017/btaa907.pdf>, doi:10.1093/bioinformatics/btaa907.

2.4.15 2020-10-19

1. Jialin Liu, Chao Gao, Joshua Sodicoff, Velina Kozareva, Evan Z Macosko, and Joshua D Welch. Jointly defining cell types from multiple single-cell datasets using liger. *Nature Protocols*, pages 1–31, 2020. URL: <https://doi.org/10.1038/s41596-020-0391-8>.
2. Shiyi Yang, Sean E Corbett, Yusuke Koga, Zhe Wang, W Evan Johnson, Masanao Yajima, and Joshua D Campbell. Decontamination of ambient rna in single-cell rna-seq with decontx. *Genome biology*, 21(1):1–15, 2020. URL: <https://doi.org/10.1186/s13059-020-1950-6>.
3. Y. Zhang, P. Qu, Y. Ji, and et al. A system hierarchy for brain-inspired computing. *Nature*, 2020. URL: <https://www.nature.com/articles/s41586-020-2782-y>.
4. M. Song, M. Pebworth, X. Yang, and et al. Cell-type-specific 3d epigenomes in the developing human cortex. *Nature*, 2020. URL: <https://doi.org/10.1038/s41586-020-2825-4>.
5. Kalliopi Skamaki, Stephane Emond, Matthieu Chodorge, John Andrews, D Gareth Rees, Daniel Cannon, Bojana Popovic, Andrew Buchanan, Ralph Minter, and Florian Hollfelder. In vitro evolution of antibody affinity via insertional mutagenesis scanning of an entire antibody variable region. *PNAS*, 2020. URL: <https://doi.org/10.1073/pnas.2002954117>.
6. Claudio Lorenzi, Sylvain Barriere, Laureline Villemain, Jean-Philippevand Dejardin Bretones, Alban Mancheron, and William Ritchie. Imoka: k-mer based software to anahttps://github.com/statbiomed/hlab-pubmon/blob/master/bibs/m2020-10-19.biblyze large collections of sequencing data. *Genome Biology*, 21(1):261, Oct 2020. URL: <https://doi.org/10.1186/s13059-020-02165-2>, doi:10.1186/s13059-020-02165-2.
7. Shengqing Stan Gu, Xiaoqing Wang, Xihao Hu, Peng Jiang, Ziyi Li, Nicole Traugh, Xia Bu, Qin Tang, Chenfei Wang, Zexian Zeng, Jingxin Fu, Cliff Meyer, Yi Zhang, Paloma Cejas, Klothilda Lim, Jin Wang, Wubing Zhang, Collin Tokheim, Avinash Das Sahu, Xiaofang Xing, Benjamin Kroger, Zhangyi Ouyang, Henry Long, Gordon J. Freeman, Myles Brown, and X. Shirley Liu. Clonal tracing reveals diverse patterns of response to immune checkpoint blockade. *Genome Biology*, 21(1):263, Oct 2020. URL: <https://doi.org/10.1186/s13059-020-02166-1>, doi:10.1186/s13059-020-02166-1.
8. Andres M Cifuentes-Bernal, Vu Vh Pham, Xiaomei Li, Lin Liu, Jiuyong Li, and Thuc Duy Le. A pseudo-temporal causality approach to identifying miRNA-mRNA interactions during biological processes. *Bioinformatics*, 10 2020. btaa899. URL: <https://doi.org/10.1093/bioinformatics/btaa899>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa899/33937099/btaa899.pdf>, doi:10.1093/bioinformatics/btaa899.
9. Jeffrey N Law, Shiv D Kale, and T M Murali. Accurate and efficient gene function prediction using a Multi-Bacterial network. *Bioinformatics*, 10 2020. btaa885. URL: <https://doi.org/10.1093/bioinformatics/btaa885>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa885/33894881/btaa885.pdf>, doi:10.1093/bioinformatics/btaa885.
10. K.C. Akdemir, V.T. Le, J.M. Kim, and et al. Somatic mutation distributions in cancer genomes vary with three-dimensional chromatin structure. *Nat Genet*, 10 2020. btaa885. URL: <https://doi.org/10.1038/s41588-020-0708-0>, arXiv:<https://www.nature.com/articles/s41588-020-0708-0#citeas>, doi:10.1038/s41588-020-0708-0.
11. Kexin Huang, Cao Xiao, Lucas M Glass, and Jimeng Sun. MolTrans: Molecular interaction transformer for drug target interaction prediction. *Bioinformatics*, 10 2020. btaa880. URL: <https://doi.org/10.1093/bioinformatics/btaa880>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa880/33937102/btaa880.pdf>, doi:10.1093/bioinformatics/btaa880.
12. Julie Jiang, Li-Ping Liu, and Soha Hassoun. Learning graph representations of biochemical networks and its application to enzymatic link prediction. *Bioinformatics*, 10 2020. btaa881. URL: <https://doi.org/10.1093/bioinformatics/btaa881>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa881/33880019/btaa881.pdf>, doi:10.1093/bioinformatics/btaa881.

13. Gregory W Schwartz, Yeqiao Zhou, Jelena Petrovic, Maria Fasolino, Lanwei Xu, Sydney M Shaffer, Warren S Pear, Golnaz Vahedi, and Robert B Faryabi. Toomanycells identifies and visualizes relationships of single-cell clades. *Nature Methods*, 17(4):405–413, 2020. URL: <https://www-nature-com.eproxy.lib.hku.hk/articles/s41592-020-0748-5>.
14. Geoff Fudenberg, David R Kelley, and Katherine S Pollard. Predicting 3d genome folding from dna sequence with akita. *Nature Methods*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41592-020-0958-x>.
15. Birgit Kriener, Rishidev Chaudhuri, and Ila R Fiete. Robust parallel decision-making in neural circuits with nonlinear inhibition. *Proceedings of the National Academy of Sciences*, 117(41):25505–25516, 2020. URL: <https://www.pnas.org/content/117/41/25505.short>.

2.4.16 2020-10-05

1. Monika Litviňuková, Carlos Talavera-López, Henrike Maatz, Daniel Reichart, Catherine L Worth, Eric L Lindberg, Masatoshi Kanda, Krzysztof Polanski, Matthias Heinig, Michael Lee, and others. Cells of the adult human heart. *Nature*, pages 1–10, 2020. URL: <https://www.nature.com/articles/s41586-020-2797-4>.
2. Alexandra Maslova, Ricardo N Ramirez, Ke Ma, Hugo Schmutz, Chendi Wang, Curtis Fox, Bernard Ng, Christophe Benoist, Sara Mostafavi, and others. Deep learning of immune cell differentiation. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/09/23/2011795117>.
3. Ziqi Zhang and Xiuwei Zhang. Inference of multiple trajectories in single cell rna-seq data from rna velocity. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/early/2020/10/02/2020.09.30.321125>, doi:10.1101/2020.09.30.321125.
4. Simone Ciccolella, Camir Ricketts, Mauricio Soto Gomez, Murray Patterson, Dana Silverbush, Paola Bonizzoni, Iman Hajirasouliha, and Gianluca Della Vedova. Inferring Cancer Progression from Single-Cell Sequencing while Allowing Mutation Losses. *Bioinformatics*, 08 2020. btaa722. URL: <https://doi.org/10.1093/bioinformatics/btaa722>, arXiv:<https://academic.oup.com/bioinformatics/advance-article-pdf/doi/10.1093/bioinformatics/btaa722/33658591/btaa722.pdf>, doi:10.1093/bioinformatics/btaa722.
5. Xin Chen, Zhaowei Yang, Wanqiu Chen, Yongmei Zhao, Andrew Farmer, Bao Tran, Vyacheslav Furtak, Malcolm Moos, Wenming Xiao, and Charles Wang. A multi-center cross-platform single-cell rna sequencing reference dataset. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/early/2020/09/21/2020.09.20.305474>, arXiv:<https://www.biorxiv.org/content/early/2020/09/21/2020.09.20.305474.full.pdf>, doi:10.1101/2020.09.20.305474.
6. Zhisong He, Agnieska Brazovskaja, Sebastian Ebert, J. Gray Camp, and Barbara Treutlein. Css: cluster similarity spectrum integration of single-cell genomics data. *Genome biology*, 21(1):1–21, 2020.
7. Nitzan Nissimov, Zivar Hajiyeva, Sebastian Torke, Katja Grondey, Wolfgang Brück, Silke Häusser-Kinzel, and Martin S. Weber. B cells reappear less mature and more activated after their anti-cd20-mediated depletion in multiple sclerosis. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/09/29/2012249117>, arXiv:<https://www.pnas.org/content/early/2020/09/29/2012249117.full.pdf>, doi:10.1073/pnas.2012249117.
8. Rui Martiniano, Erik Garrison, Eppie R Jones, Andrea Manica, and Richard Durbin. Removing reference bias and improving indel calling in ancient dna data analysis by mapping to a sequence variation graph. *Genome biology*, 21(1):1–18, 2020.
9. Yuzhe Yang and Zhi Xu. Rethinking the value of labels for improving class-imbalanced learning. *NeurIPS 2020*, 2020. arXiv:<https://arxiv.org/pdf/2006.07529>.
10. Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, and Yanis Kalantidis. Decoupling representation and classifier for long-tailed recognition. *ICLR 2020*, 2020. arXiv:<https://arxiv.org/pdf/1910.09217>.

11. Dong Zhang, Hanwang Zhang, Jinhui Tang, Xiansheng Hua, and Qianru Sun. Causal intervention for weakly-supervised semantic segmentation. *NeurIPS 2020*, 2020. [arXiv:https://arxiv.org/pdf/2009.12547](https://arxiv.org/pdf/2009.12547).
12. Courtney J Spoerer, Tim C Kietzmann, Johannes Mehrer, Ian Charest, and Nikolaus Kriegeskorte. Recurrent neural networks can explain flexible trading of speed and accuracy in biological vision. *bioRxiv*, pages 677237, 2020.
13. Wen Jun Xie, Yifeng Qi, and Bin Zhang. Characterizing chromatin folding coordinate and landscape with deep learning. *PLOS Computational Biology*, 16(9):e1008262, 2020.
14. Camila PE de Souza, Mirela Andronescu, Tehmina Masud, Farhia Kabeer, Justina Biele, Emma Laks, Daniel Lai, Patricia Ye, Jazmine Brimhall, Beixi Wang, and others. Epiclomal: probabilistic clustering of sparse single-cell dna methylation data. *PLOS Computational Biology*, 16(9):e1008270, 2020.

2.4.17 2020-09-21

1. Xizhi Luo, Fei Qin, Guoshuai Cai, and Feifei Xiao. Integrating Genomic Correlation Structure Improves Copy Number Variations Detection. *Bioinformatics*, 08 2020. URL: <https://doi.org/10.1093/bioinformatics/btaa737>.
2. Ron Zeira and Benjamin J Raphael. Copy number evolution with weighted aberrations in cancer. *Bioinformatics*, 36(Supplement_1):i344–i352, 07 2020. URL: <https://doi.org/10.1093/bioinformatics/btaa470>.
3. Ye Yuan and Ziv Bar-Joseph. Deep learning for inferring gene relationships from single-cell expression data. *Proceedings of the National Academy of Sciences*, 116(52):27151–27158, 2019.
4. Thinh N Tran and Gary Bader. Tempora: cell trajectory inference using time-series single-cell rna sequencing data. *bioRxiv*, pages 846907, 2019.
5. Nicola F Müller, Ugnė Stolz, Gytis Dudas, Tanja Stadler, and Timothy G Vaughan. Bayesian inference of reassortment networks reveals fitness benefits of reassortment in human influenza viruses. *Proceedings of the National Academy of Sciences*, 117(29):17104–17111, 2020.
6. Yan Zhang, Yaru Zhang, Jun Hu, Ji Zhang, Fangjie Guo, Meng Zhou, Guijun Zhang, Fulong Yu, and Jianzhong Su. scTPA: a web tool for single-cell transcriptome analysis of pathway activation signatures. *Bioinformatics*, 36(14):4217–4219, 05 2020. URL: <https://doi.org/10.1093/bioinformatics/btaa532>, [arXiv:https://academic.oup.com/bioinformatics/article-pdf/36/14/4217/33547302/btaa532.pdf](https://academic.oup.com/bioinformatics/article-pdf/36/14/4217/33547302/btaa532.pdf), doi:10.1093/bioinformatics/btaa532.
7. Lipeng Jin, Chenyao Li, Tao Liu, and Lei Wang. A potential prognostic prediction model of colon adenocarcinoma with recurrence based on prognostic lncrna signatures. *Human genomics*, 14(1):1–11, 2020.
8. Lawrence T. C Ong, Grant P Parnell, Ali Afrasiabi, Graeme J Stewart, Sanjay Swaminathan, and David R Booth. Transcribed b lymphocyte genes and multiple sclerosis risk genes are underrepresented in epstein–barr virus hypomethylated regions. *Genes and immunity*, 21(2):91–99, 2019.
9. Amrit Dhar, Duncan K. Ralph, Vladimir N. Minin, and Frederick A. Matsen, IV. A bayesian phylogenetic hidden markov model for b cell receptor sequence analysis. *PLOS Computational Biology*, 16(8):1–27, 08 2020. URL: <https://doi.org/10.1371/journal.pcbi.1008030>, doi:10.1371/journal.pcbi.1008030.
10. Long Vo Ngoc, Cassidy Yunjing Huang, California Jack Cassidy, Claudia Medrano, and James T Kadonaga. Identification of the human dpr core promoter element using machine learning. *Nature*, 2020. URL: <https://www.nature.com/articles/s41586-020-2689-7>.
11. Xiaodong Liu, Many Others, and Jose M. Polo. Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature*, 2020. URL: <https://www.nature.com/articles/s41586-020-2734-6>.
12. Ori Maoz, Gašper Tkacik, Mohamad Saleh Esteki, Roozbeh Kiani, and Elad Schneidman. Learning probabilistic representations with randomly connected neural circuits. *PNAS*, 2020. URL: <https://doi.org/10.1073/pnas.1912804117>.

13. Guillaume Holley and Páll Melsted. Bifrost: highly parallel construction and indexing of colored and compacted de Bruijn graphs. *Genome Biology*, 2020.
14. Yuan He, Surya B Chhetri, Marios Arvanitis, Kaushik Srinivasan, François Aguet, Kristin G Ardlie, Alvaro N Barbeira, Rodrigo Bonazzola, Hae Kyung Im, Christopher D Brown, and others. sn-spMF: matrix factorization informs tissue-specific genetic regulation of gene expression. *Genome Biology*, 21(1):1–25, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02129-6>.
15. Kevin J Gleason, Fan Yang, Brandon L Pierce, Xin He, and Lin S Chen. Primo: integration of multiple GWAS and omics QTL summary statistics for elucidation of molecular mechanisms of trait-associated SNPs and detection of pleiotropy in complex traits. *Genome Biology*, 21(1):1–24, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02125-w>.
16. Weihua Hu*, Bowen Liu*, Joseph Gomes, Marinka Zitnik, Percy Liang, Vijay Pande, and Jure Leskovec. Strategies for pre-training graph neural networks. In *International Conference on Learning Representations*. 2020. URL: <https://openreview.net/forum?id=HJIWWJSFDH>.
17. Hao Wang and Dit-Yan Yeung. A Survey on Bayesian Deep Learning. *arXiv e-prints (To appear in ACM Computing Surveys (CSUR) 2020)*, pages arXiv:1604.01662, April 2016. arXiv:1604.01662.
18. Dongjoon Lim and Mathieu Blanchette. EvoLSTM: context-dependent models of sequence evolution using a sequence-to-sequence LSTM. *Bioinformatics*, 36(Supplement_1):i353–i361, 07 2020. URL: <https://doi.org/10.1093/bioinformatics/btaa447>, arXiv:https://academic.oup.com/bioinformatics/article-pdf/36/Supplement_1/i353/33488780/btaa447.pdf, doi:10.1093/bioinformatics/btaa447.
19. Allen W Zhang and Kieran R Campbell. Computational modelling in single-cell cancer genomics: methods and future directions. *arXiv preprint arXiv:2005.01549*, 2020. URL: <https://arxiv.org/abs/2005.01549>.
20. Chenfei Wang, Dongqing Sun, Xin Huang, Changxin Wan, Ziyi Li, Ya Han, Qian Qin, Jingyu Fan, Xintao Qiu, Yingtian Xie, and others. Integrative analyses of single-cell transcriptome and regulome using maestro. *Genome biology*, 21(1):1–28, 2020. URL: <https://doi.org/10.1186/s13059-020-02116-x>.
21. Sergey Aganezov and Benjamin J Raphael. Reconstruction of clone-and haplotype-specific cancer genome karyotypes from bulk tumor samples. *Genome Research*, 30(9):1274–1290, 2020. URL: <https://genome.cshlp.org/content/30/9/1274>.
22. Simone Zaccaria and Benjamin J Raphael. Characterizing allele-and haplotype-specific copy numbers in single cells with chisel. *Nature Biotechnology*, pages 1–8, 2020. URL: <https://www.nature.com/articles/s41587-020-0661-6>.
23. Volker Bergen, Marius Lange, Stefan Peidli, F Alexander Wolf, and Fabian J Theis. Generalizing rna velocity to transient cell states through dynamical modeling. *Nature Biotechnology*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41587-020-0591-3>.
24. Junyue Cao, Wei Zhou, Frank Steemers, Cole Trapnell, and Jay Shendure. Sci-fate characterizes the dynamics of gene expression in single cells. *Nature Biotechnology*, pages 1–9, 2020. URL: <https://www.nature.com/articles/s41587-020-0480-9>.
25. Qi Qiu, Peng Hu, Xiaojie Qiu, Kiya W Govek, Pablo G Cámara, and Hao Wu. Massively parallel and time-resolved rna sequencing in single cells with scnt-seq. *Nature methods*, pages 1–11, 2020. URL: <https://www.nature.com/articles/s41592-020-0935-4>.
26. Scott A. Lujan, Matthew J. Longley, Margaret H. Humble, Christopher A. Lavender, Adam Burkholder, Emma L. Blakely, Charlotte L. Alston, Grainne S. Gorman, Doug M. Turnbull, Robert McFarland, Robert W. Taylor, Thomas A. Kunkel, and William C. Copeland. Ultrasensitive deletion detection links mitochondrial dna replication, disease, and aging. *Genome Biology*, 21(1):248, Sep 2020. URL: <https://doi.org/10.1186/s13059-020-02138-5>, doi:10.1186/s13059-020-02138-5.

2.4.18 2020-09-07

1. Alessia Centonze, Shuheng Lin, Elisavet Tika, Alejandro Sifrim, Marco Fioramonti, Milan Malfait, Yura Song, Aline Wuidart, Jens Van Herck, Anne Dannau, and others. Heterotypic cell–cell communication regulates glandular stem cell multipotency. *Nature*, 584(7822):608–613, 2020. URL: <https://www.nature.com/articles/s41586-020-2632-y>.
2. Eiji Miyauchi, Seok-Won Kim, Wataru Suda, Masami Kawasumi, Satoshi Onawa, Naoko Taguchi-Atarashi, Hidetoshi Morita, Todd D Taylor, Masahira Hattori, and Hiroshi Ohno. Gut microorganisms act together to exacerbate inflammation in spinal cords. *Nature*, pages 1–5, 2020. URL: <https://www.nature.com/articles/s41586-020-2634-9>.
3. Amir Ghasemian, Homa Hosseinmardi, Aram Galstyan, Edoardo M. Airoidi, and Aaron Clauset. Stacking models for nearly optimal link prediction in complex networks. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/09/03/1914950117>, arXiv:<https://www.pnas.org/content/early/2020/09/03/1914950117.full.pdf>, doi:10.1073/pnas.1914950117.
4. Alessandro Didonna, Ester Canto Puig, Qin Ma, Atsuko Matsunaga, Brenda Ho, Stacy J. Caillier, Hengameh Shams, Nicholas Lee, Stephen L. Hauser, Qiumin Tan, Scott S. Zamvil, and Jorge R. Oksenberg. Ataxin-1 regulates b cell function and the severity of autoimmune experimental encephalomyelitis. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/09/01/2003798117>, arXiv:<https://www.pnas.org/content/early/2020/09/01/2003798117.full.pdf>, doi:10.1073/pnas.2003798117.
5. David Bau, Jun-Yan Zhu, Hendrik Strobelt, Agata Lapedriza, Bolei Zhou, and Antonio Torralba. Understanding the role of individual units in a deep neural network. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/08/31/1907375117>.
6. Akshaya Ramesh, Ryan D Schubert, Ariele L Greenfield, Ravi Dandekar, Rita Loudermilk, Joseph J Sabatino, Matthew T Koelzer, Edwina B Tran, Kanishka Koshal, Kicheol Kim, and others. A pathogenic and clonally expanded b cell transcriptome in active multiple sclerosis. *Proceedings of the National Academy of Sciences*, 2020. URL: <https://www.pnas.org/content/early/2020/08/27/2008523117>.
7. Zhisong He, Agnieszka Brazovskaja, Sebastian Ebert, Gray Camp, and Barbara Treutlein. CSS: cluster similarity spectrum integration of single-cell genomics data. *Genome Biology*, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02147-4>.
8. Andrew E. Teschendorff, Tianyu Zhu, Charles E. Breeze, and Stephan Beck. Episcor: cell type deconvolution of bulk tissue dna methylomes from single-cell rna-seq data. *Genome Biology*, 2020. URL: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02126-9>.
9. James M McFarland, Brenton R Paoletta, Allison Warren, Kathryn Geiger-Schuller, Tsukasa Shibue, Michael Rothberg, Olena Kuksenko, William N Colgan, Andrew Jones, Emily Chambers, and others. Multiplexed single-cell transcriptional response profiling to define cancer vulnerabilities and therapeutic mechanism of action. *Nature Communications*, 11(1):1–15, 2020. URL: <https://www.nature.com/articles/s41467-020-17440-w>.
10. Anna S Nam, Ronan Chaligne, and Dan A Landau. Integrating genetic and non-genetic determinants of cancer evolution by single-cell multi-omics. *Nature Reviews Genetics*, pages 1–16, 2020. URL: <https://www.nature.com/articles/s41576-020-0265-5>.

2.5 Best3 in semester1-2020

2.5.1 Best 3

1. Simone Zaccaria and Benjamin J Raphael. Characterizing allele-and haplotype-specific copy numbers in single cells with chisel. *Nature biotechnology*, pages 1–8, 2020. URL: <https://www.nature.com/articles/s41587-020-0661-6>.
2. Volker Bergen, Marius Lange, Stefan Peidli, F Alexander Wolf, and Fabian J Theis. Generalizing rna velocity to transient cell states through dynamical modeling. *Nature biotechnology*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41587-020-0591-3>.
3. Deepmind. ‘alphafold: a solution to a 50-year-old grand challenge in biology. 2020. URL: <https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>.

2.5.2 Rongting

1. CHISEL: It is really the relevant one to what I’m doing now. The integration of BAF and RDR of scDNA-seq data in chisel is a step forward in calling CNV more accurately. Though CaSpER also uses BAF and RDR information of scRNA-seq at the same time, the phasing strategy in chisel can be more helpful for getting more information.

2. Methods Review article in single cell data: The review outline the major challenges in analyzing single-cell cancer genomics data and survey the current computational tools available to tackle these. They further outline unsolved problems that may be considered as major opportunities for future methods development to help interpret the vast quantities of data being generated.

3. AlphaFold2: As the title of the share says, “game has changed in the field of solving protein structures!” It can be helpful in many research works. However, “This isn’t the end of something,” “It’s the beginning of many new things.”

1. Simone Zaccaria and Benjamin J Raphael. Characterizing allele-and haplotype-specific copy numbers in single cells with chisel. *Nature Biotechnology*, pages 1–8, 2020. URL: <https://www.nature.com/articles/s41587-020-0661-6>.
2. Allen W Zhang and Kieran R Campbell. Computational modelling in single-cell cancer genomics: methods and future directions. *arXiv preprint arXiv:2005.01549*, 2020. URL: <https://arxiv.org/abs/2005.01549>.
3. Robert F Service. ‘the game has changed.’ ai triumphs at protein folding. 2020. URL: <https://www.sciencemag.org/news/2020/11/game-has-changed-ai-triumphs-solving-protein-structures>.

2.5.3 Xianjie

1. Proteins prediction-Deep learning Cool stuff

2. scVelo Life process is assumed to be dynamically regulated. ScVelo uses kinetics to build the dynamical model trying to characterize and predict the change of gene expression level. I learned kinetics from the course of physical chemistry and it’s pleasantly surprised to find that how knowledge from textbook could apply to the real-world problem and benefit to the community.

3. Benefit from the CHISEL model, especially the EM part.

1. Andrew W Senior, Richard Evans, John Jumper, James Kirkpatrick, Laurent Sifre, Tim Green, Chongli Qin, Augustin Židek, Alexander WR Nelson, Alex Bridgland, and others. Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792):706–710, 2020. URL: <https://www.nature.com/articles/s41586-019-1923-7?fbclid=IwAR37LQHolvzYLj9Dj5wGbaH48oKcKFEX4jaGFwl1oxspEvxtlC6uyDgrCKg>.

2. Volker Bergen, Marius Lange, Stefan Peidli, F Alexander Wolf, and Fabian J Theis. Generalizing rna velocity to transient cell states through dynamical modeling. *Nature biotechnology*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41587-020-0591-3>.
3. Simone Zaccaria and Benjamin J Raphael. Characterizing allele-and haplotype-specific copy numbers in single cells with chisel. *Nature biotechnology*, pages 1–8, 2020. URL: <https://www.nature.com/articles/s41587-020-0661-6>.

2.5.4 Yuanhua

1. AlphaFold2: it addressed a long-history problem on protein structure prediction with near perfect performance, and opens many exciting opportunities for functional and interaction analysis in biology. It is also a very encouraging example that data driven (often black-box) methods have great potentials on revealing complex patterns in science.
2. CHISEL: the multi-step phasing of the gremlin SNPs effectively increases the power of using the allelic information to identify copy number variations. The modelling of the two-module data further motivates the integration pivot of multiple mutations and multiple omics data for cancer mutations.
3. scVelo: it offers an interesting extension of RNA velocity and nice implementation. The inferred splicing kinetic parameters through the dynamic model also inspires further investigation on its regulatory patterns.
1. Deepmind. ‘alphafold: a solution to a 50-year-old grand challenge in biology. 2020. URL: <https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>.
2. Simone Zaccaria and Benjamin J Raphael. Characterizing allele-and haplotype-specific copy numbers in single cells with chisel. *Nature biotechnology*, pages 1–8, 2020. URL: <https://www.nature.com/articles/s41587-020-0661-6>.
3. Volker Bergen, Marius Lange, Stefan Peidli, F Alexander Wolf, and Fabian J Theis. Generalizing rna velocity to transient cell states through dynamical modeling. *Nature biotechnology*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41587-020-0591-3>.

2.5.5 Julia

1. It provided insight on factors affecting tumor development and progression, which was very useful to me to reference to when I was writing my literature review.
2. Also used as a reference when i was writing my literature review. cause it provided insight on different computational analysis method used for analysing cell heterogeneity.
3. Even though it is not really relevant to what I’m doing for my fyp, but it is probably one of the article that i understand the most regarding the methodology.
1. Anna S Nam, Ronan Chaligne, and Dan A Landau. Integrating genetic and non-genetic determinants of cancer evolution by single-cell multi-omics. *Nature Reviews Genetics*, pages 1–16, 2020. URL: <https://www.nature.com/articles/s41576-020-0265-5>.
2. Gabriela S Kinker, Alissa C Greenwald, Rotem Tal, Zhanna Orlova, Michael S Cuoco, James M McFarland, Allison Warren, Christopher Rodman, Jennifer A Roth, Samantha A Bender, and others. Pan-cancer single cell rna-seq uncovers recurring programs of cellular heterogeneity. *Nature Genetics*, pages 807552, 2019. URL: <https://www.biorxiv.org/content/10.1101/807552v1.abstract>.
3. Tatiana A. Gurbich and Valery Vladimirovich Ilinsky. Classifycnv: a tool for clinical annotation of copy-number variants. *Scientific Reports*, 10(1):20375, Nov 2020. URL: <https://doi.org/10.1038/s41598-020-76425-3>, doi:10.1038/s41598-020-76425-3.

2.5.6 Jerry

1. Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations I like it because this paper provided a new angle – Network Analysis to look into single cell BCR repertoire, which can clearly visualize the condition of diseased and healthy subjects. This paper also explains quite a bit on the mechanism of BCR differentiation and mutation.

2. A pathogenic and clonally expanded B cell transcriptome in active multiple sclerosis This paper described the B cell clonal expansion of multiple sclerosis and proved the effectiveness of B cell depletion therapy. This paper inspired me to compare the B cell expansion between diseased patient and healthy subject.

3. Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs I like it because this paper applied transcriptomic analysis into many parts of our body, which can reveal a clear pathway of immune cell differentiation. This paper also provide marker genes of many cell types, which are useful for future analysis

1. Rachael JM Bashford-Rogers, Anne L Palser, Brian J Huntly, Richard Rance, George S Vassiliou, George A Follows, and Paul Kellam. Network properties derived from deep sequencing of human b-cell receptor repertoires delineate b-cell populations. *Genome research*, 23(11):1874–1884, 2013. URL: <https://genome.cshlp.org/content/23/11/1874.short>.
2. Akshaya Ramesh, Ryan D Schubert, Ariele L Greenfield, Ravi Dandekar, Rita Loudermilk, Joseph J Sabatino, Matthew T Koelzer, Edwina B Tran, Kanishka Koshal, Kicheol Kim, and others. A pathogenic and clonally expanded b cell transcriptome in active multiple sclerosis. *Proceedings of the National Academy of Sciences*, 117(37):22932–22943, 2020. URL: <https://www.pnas.org/content/117/37/22932.short>.
3. Shuai He, Linhe Wang, Yang Liu, Yiqi Li, Haitian Chen, Jinghong Xu, Wan Peng, Guowang Lin, Panpan Wei, Bo Li, and others. Single-cell transcriptome profiling an adult human cell atlas of 15 major organs. *bioRxiv*, 2020. URL: <https://www.biorxiv.org/content/10.1101/2020.03.18.996975v1.abstract>.

2.5.7 Weizhong

1. Momentum Contrast for Unsupervised Visual Representation Learning Reason : this work improve the contrastive learning by a simple queuing technique, intuitively and at the same time effectively.

2. Dissecting heterogeneous cell populations across drug and disease conditions with PopAlign Reason : This paper focuses on an important question : how to map cell clusters across different condition. The creativity is mainly that we only build one probability density function, making the changes of coefficient explainable.

3. Characterising the loss-of-function impact of 5' untranslated region variants in 15,708 individuals Reason Of great biological significance, this paper covers a large scale population genetics analysis on 5' UTR sequences, showing the strength of purification selection on the sequence motif.

1. Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9729–9738. 2020. URL: https://openaccess.thecvf.com/content_CVPR_2020/html/He_Momentum_Contrast_for_Unsupervised_Visual_Representation_Learning_CVPR_2020_paper.html.
2. Sisi Chen, Paul Rivaud, Jong H Park, Tiffany Tsou, Emeric Charles, John R Haliburton, Flavia Pichiorri, and Matt Thomson. Dissecting heterogeneous cell populations across drug and disease conditions with popalign. *Proceedings of the National Academy of Sciences*, 117(46):28784–28794, 2020. URL: <https://www.pnas.org/content/117/46/28784.short>.
3. Nicola Whiffin, Konrad J Karczewski, Xiaolei Zhang, Sonia Chothani, Miriam J Smith, D Gareth Evans, Angharad M Roberts, Nicholas M Quaipe, Sebastian Schafer, Owen Rackham, and others. Characterising the loss-of-function impact of 5' untranslated region variants in 15,708 individuals. *Nature communications*, 11(1):1–12, 2020. URL: <https://www.nature.com/articles/s41467-019-10717-9>.

2.5.8 Qiaochen

1. Strategies for Pre-training Graph Neural Networks Strategies for pretraining graph neural networks are summarized and validated, pretrained models can also be downloaded. Like Bert-triggered pretrained language models in the NLP field, this work has the potential to benefit many downstream tasks.
2. scVelo: Generalizing RNA velocity to transient cell states through dynamical modeling An interesting extension of steady-state estimation of RNA velocity to maximum likelihood estimation.
3. Rethinking the Value of Labels for Improving Class-Imbalanced Learning: A lite-weighted yet effective strategy to tackle class-imbalance problems in classification.
 1. Weihua Hu, Bowen Liu, Joseph Gomes, Marinka Zitnik, Percy Liang, Vijay Pande, and Jure Leskovec. Strategies for pre-training graph neural networks. *arXiv preprint arXiv:1905.12265*, 2019. URL: <https://openreview.net/forum?id=HJIWWJSFDH>.
 2. Volker Bergen, Marius Lange, Stefan Peidli, F Alexander Wolf, and Fabian J Theis. Generalizing rna velocity to transient cell states through dynamical modeling. *Nature biotechnology*, pages 1–7, 2020. URL: <https://www.nature.com/articles/s41587-020-0591-3>.
 3. Yuzhe Yang and Zhi Xu. Rethinking the value of labels for improving class-imbalanced learning. *Advances in Neural Information Processing Systems*, 2020. URL: <https://papers.nips.cc/paper/2020/file/e025b6279c1b88d3ec0eca6fcb6e6280-Paper.pdf>.

2.5.9 Yoyo

1. Jointly defining cell types from multiple single-cell datasets using LIGER The tool may be useful for analyzing scRNA datasets, although I haven't used it
2. Identification of a novel cancer stem cell subpopulation that promotes progression of human fatal renal cell carcinoma by single-cell rna-seq analysis. This paper involves a couple of fancy tools I'm interested in, e.g. RNA velocity, cell-cell communication analysis
3. Reprogramming roadmap reveals route to human induced trophoblast stem cells The scRNA-seq data is relevant to my research, and I am exploring this
 1. Jialin Liu, Chao Gao, Joshua Sodicoff, Velina Kozareva, Evan Z Macosko, and Joshua D Welch. Jointly defining cell types from multiple single-cell datasets using liger. *Nature Protocols*, 15(11):3632–3662, 2020. URL: <https://www.nature.com/articles/s41596-020-0391-8>.
 2. Xiu-wu Pan, Hao Zhang, Jia-xin Chen Da Xu, Wen-jin Chen, Si-shun Gan, Fa-jun Qu, Chuan-min Chu, Jian-wei Cao, Ying-hui Fan, Xu Song, and others. Identification of a novel cancer stem cell subpopulation that promotes progression of human fatal renal cell carcinoma by single-cell rna-seq analysis. *International journal of biological sciences*, 16(16):3149, 2020. URL: <https://www.ijbs.com/v16p3149.htm>.
 3. Xiaodong Liu, John F Ouyang, Fernando J Rossello, Jia Ping Tan, Kathryn C Davidson, Daniela S Valdes, Jan Schröder, Yu BY Sun, Joseph Chen, Anja S Knaupp, and others. Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature*, 586(7827):101–107, 2020. URL: <https://www.nature.com/articles/s41586-020-2734-6>.