Single-cell genomics to study developmental cell fate decisions in zebrafish

Moreno-Ayala R., Junker J.P.

Briefings in Functional Genomics
2021 NOV 06 ; 20(6): 420–426
2021 MAR 30 (first published online: final publication)
doi: 10.1093/bfgp/elab018

Publisher: Oxford University Press

Single cell genomics to study developmental cell fate decisions in zebrafish

Roberto Moreno-Ayala[1], Jan Philipp Junker[1,*]

[1] Max Delbrück Center for Molecular Medicine, Berlin Institute for Medical Systems Biology

[*] Correspondence: janphilipp.junker@mdc-berlin.de

Abstract

New developments in single cell genomics have transformed developmental biology in recent years by enabling systematic analysis of embryonic cell types and differentiation trajectories. Ongoing efforts in experimental and computational method development aim to reveal gene-regulatory mechanisms and to provide additional spatio-temporal information about developmental cell fate decisions. Here, we discuss recent technological developments as well as biological applications of single cell genomics, with a particular focus on analysis of developmental cell fate decisions. While the approaches described here are generally applicable to a broad range of model systems, we focus our discussion on applications in zebrafish, which has proven to be a particularly powerful model organism for establishing novel methods in single-cell genomics.

Keywords

zebrafish, cell fate decisions, development, single-cell genomics, spatial transcriptomics, lineage tracing

Highlights

- Single-cell transcriptomics allows systematic identification of cell types and the differentiation trajectories that produce them.
- The zebrafish has served as a powerful model organism for establishing novel methods in single-cell genomics.
- Spatial information can be added to single-cell transcriptomics data by three different approaches: Computational image reconstruction, high-throughput fluorescent in situ hybridization, and spatial barcoding of tissue sections prior to sequencing analysis.
- Explicit temporal information can be added by high-throughput lineage tracing based on CRISPR/Cas9 induced lineage barcodes.

Short author biographies

Jan Philipp Junker is a group leader at the Max Delbrück Center in Berlin. Using the zebrafish as their primary model system, his group develops and uses methods in single-cell genomics in order to understand cell fate decisions in health and disease.

Roberto Moreno-Ayala is a postdoctoral researcher in Jan Philipp Junker's laboratory working on early zebrafish developmental variability and its phenotypic outcomes.

Main text

One of the main outcomes of embryonic development is the acquisition of cell identity and function. Identifying and categorizing the many cell types present in an organism has been a slow and laborious process in the past. Single-cell genomics technologies constitute an important advancement for the characterization of the cellular heterogeneity in a sample by allowing identification of transcriptomic and chromatin accessibility profiles in thousands of single cells. Importantly, these approaches not only enable systematic identification of embryonic cell types, but they also yield insight into developmental differentiation trajectories, lineage trees, and regulatory mechanisms [1–3]. Single cell transcriptomics is by far the most advanced of the single cell omics technologies and will hence take up the largest part of this review. The zebrafish has been one of the protagonist models in the emergence of the single-cell technologies. In this review, we summarize the applications used to this day in this model organism.

Technologies for single-cell transcriptomics

Single-cell genomics experiments typically start with dissociation of the tissue of interest into a single cell solution, and the quality of the single cell suspension is a decisive factor for the success of the downstream experiment. Incomplete dissociation, loss of specific cell types, and triggering of cellular stress response are typical challenges in single-cell genomics experiments. The use of a psychrophilic protease during this critical step has been reported to alleviate possible artifacts [4]. Once dissociation has been optimized, individual cells need to be processed into sequencing libraries. The two most widely used experimental approaches are plate-based processing using liquid handling robotics and droplet microfluidics. Most early studies in single-cell transcriptomics were plate-based, i.e. cells are sorted into and lysed in individual wells of a microwell plate [5,6]. Droplet-based methods, in which cell lysis and reverse transcription happen in nanoliter-sized droplets containing reagents and cellular barcodes, have gained prominence in recent years due to their higher throughput and lower cost per cell [7,8]. However, both approaches have distinct advantages and disadvantages: while plate-based methods are limited to lower numbers of cells, they typically provide higher quality data and full transcript coverage, while current droplet microfluidics approaches capture only 3’ or 5’ tags of transcripts.

Understanding cell fate decisions by single-cell genomics

Single cell RNA sequencing (scRNA-seq) has emerged as a powerful method for systematic identification of cell types [9]: single cell profiles can be clustered by transcriptome similarity, and the identified clusters correspond to the different cell types in the sample. However, clustering results may differ depending on the algorithm and the metrics that used, which leads to a certain level of ambiguity in cell type identification. Characterizing cell fate dynamics during embryogenesis is an ongoing endeavor. By sampling embryos at different stages, and by ordering single cell transcriptomic profiles by similarity, a systematic landscape of developmental differentiation trajectories can be reconstructed (Figure 1A) [10,11]. In this way, the transcriptional changes that cells undergo during differentiation can be measured in a

systematic and continuous way, which may lead to identification of marker genes for previously uncharacterized intermediate states. By tracing back the earliest origin of an embryonic structure, the likely progenitors of this cell type and the branch points of cell fate decisions can be determined [10].

Identifying the gene regulatory mechanisms that underlie cell fate decisions is one of the major questions in developmental biology, and single-cell genomics data is a powerful basis for computational prediction of regulatory networks. However, inference of gene regulatory networks remains a challenging task, and methods based purely on transcriptomic data have only moderate performance [12]. Therefore, approaches that include transcription factor binding information or open chromatin data [13,14] are required for reliable identification of the gene regulatory networks that underlie developmental cell fate decisions.

Beyond analysis of wildtype animals, single-cell genomics also provides the means to better understand mutant phenotypes by comparing their cell state composition to wild-type in a systematic manner. Interestingly, no emergence of new transcriptomic profiles in the mutant cells was observed in Farrell et al. [10] and Wagner et al. [11], only the cell state frequencies changed and some of them were enriched, indicating that cell fates stay canalized in mutant individuals. However, it is likely that in most cases, perturbation leads not only to differences in cell type composition, but also to transcriptional changes within cell types (i.e. cells move to a different cell state). Detection of activated cell states can be challenging, since transcriptional changes may be too subtle to be detected, or they may be so massive that the cell type of origin is unclear. We anticipate that establishing new experimental and computational methods for identification of activated cell states in e.g. disease conditions will be a major focus of method development in single-cell genomics in the next few years. Another important finding from these studies is the the observation that independent lineages can give rise to cells exhibiting the same transcriptomic profile [11]. This convergence is a phenomenon also observed in the *Caenorhabditis elegans* atlas [15].

While methods for pseudo-temporal ordering of single cell transcriptomics efficiently orient cells along continuous trajectories, the directionality of the differentiation process cannot be inferred from this data. However, the recently developed method “RNA velocity” [16] aims to determine the direction of cells in gene expression space based on unspliced vs. spliced (i.e. “new” vs “old”) transcript molecules. Another emerging method for looking into the immediate future of cells is RNA metabolic labeling [17,18] which allows separation of old from new transcripts by experimentally introducing labels into RNA molecules during a defined time window. While single-cell RNA labeling is so far mostly based on cultured cells, it is likely that the method will be established in model organisms, including the zebrafish, in the future. RNA labeling coupled to single-cell readout has great potential in improving reconstruction of differentiation trajectories [19] as well as in regulatory network inference [20].

In summary, single-cell genomics has to potential to systematically reveal the differentiation trajectories as well as the gene regulatory mechanisms that underlie embryonic development. Furthermore, these approaches constitute a powerful framework for dissecting the effects of perturbations based on measurement of cell type and cell state changes. The zebrafish has served as an excellent model for

establishing experimental and computational approaches due to its experimental accessibility coupled to an extensive literature on mechanisms of cell fate decisions. As these technologies mature, we anticipate that the zebrafish will become a favorite model organism for studying developmental variability and mechanisms of robustness using single-cell approaches.

Lineage tracing

While pseudo-temporal ordering is a great approach for analyzing cell differentiation, it is important to consider its limitations: 1) Continuous sampling of time points is required, which makes it very challenging and costly to identify the embryonic origin of cells that are generated at later stages. 2) The approach is based on the assumption that differentiation trajectories are smooth (no sudden "jumps" in gene expression space) and that trajectories do not cross each other. 3) In complex systems, it is important to properly adjust computational parameters in order to match the reconstruction outcome to known developmental principles. While this works well in well-characterized systems like early development, the approach may be limited in situations where less ground-truth information is available. In many cases it is therefore important to directly measure the lineage origin of individual cells in addition to the transcriptome.

The tracking of the cellular relations during embryogenesis is one of the most important questions in developmental biology to this day. The superbly detailed description of the lineage tree in *Caenorhabditis elegans* [21] is difficult to match in vertebrates, due to the higher cell numbers and complexity of cellular behaviors, which makes it practically impossible to measure lineage trees by microscopy over extended periods of time: cell labeling has been used in zebrafish to establish fate maps, but the resolution and the information per sample are limited [22]. The use of reporters activated by recombination has offered substantial improvement in precision and efficiency compared to earlier approaches, but remains limited by the amount of possible color combinations [23]. Furthermore, these approaches cannot be directly combined with cell type identification by single cell transcriptomics.

Lineage reconstruction based on genome recording has proven to be an effective way to surpass these technical issues. Recording can be achieved by transposon-based barcoding [11] during early development, lentiviral delivery of barcodes [24] or by using CRISPR/Cas9-mediated genome editing. In the latter, insertions and deletions (indels) are generated at target sites in either reporter genes [25,26], engineered recording cassettes [27,28], or endogenous loci (using a list of selected suitable genomic sites [29]). These changes in the sequence can be used as cellular 'barcodes' that are stably transmitted to daughter cells upon cell division, and hence give information about cell identify that can be retrieved at a later point by sequencing and the lineage relations between individual cells can then be established by computational methods (Figure 1B). Furthermore, since the target 'barcodes' used in these systems are expressed as messenger RNA (mRNA), these methods offer the possibility of simultaneously obtaining the transcriptomic profile, and hence the cell type, of individual cells. Typically, the targeted sequences are amplified by PCR from the whole transcriptome library in order to generate a barcode library that can be sequenced very deeply. In practice, there are several important factors to consider

when designing a CRISPR/Cas9 lineage recording experiment: 1) The design of the target sequences needs to take gene coverage bias into account. Barcodes are typically created at the 3' end of transcripts, since many experimental methods in single-cell transcriptomics have a 3' bias. 2) Dropouts lead to loss of information, i.e. not all barcode sequences are detected. Computational analysis strategies need to be optimized to deal with sparse detection. 3) Indels might affect the degradation of a transcript, and depending on the promoter used there may be cell type dependent differences in detection efficiency. Careful data analysis is need to detect such effects, which otherwise might lead to artefacts in lineage tree reconstruction.

A limitation for these approaches is that in order to have a complete picture of the lineage trees, new 'barcodes' should be generated every cell division and the information from all the cells should be captured. So far, this has not been possible with current designs of genomic recording systems, since the diversity of possible barcodes and the number of target sites per cell are limiting factors. By computer simulations, assuming ideal conditions, it has been calculated that between 30-100 targets would be needed in order to reconstruct a lineage for a zebrafish embryo with 65,000 cells (pharyngula stage) with good accuracy [30]. Furthermore, the temporal control of lineage recording is suboptimal in most published approaches, since it is determined by injection of the reagent into the zygote or driven by stimuli such as heatshock, both of which provide only a small window of activity. However, the simoultaneous use of different CRISPR/Cas and anti-CRISPRs systems, activated by a different set of promoters, might provide a good strategy to regulate the timing and increase the complexity of the lineage recording [31].

Despite these remaining limitations, methods for lineage recording have great potential for improving our understanding of cell fate decisions. For instance, a recent study of haematopoiesis combining lineage barcodes and scRNA-seq readout revealed clonal fate biases not detectable by single cell transcriptomics [32], suggesting that scRNA-seq alone is not sufficient for determining when a cell takes a decision. Furthermore, CRISPR lineage tracing has recently been used to measure the origin of transient cell types that are generated during zebrafish heart regeneration [33] and to reconstruct metastatic dissemination of cancer [34]. In summary, we expect that the combination of single-cell RNA-seq with lineage recording will develop into a general approach for identifying not only the origin of cell types, but also the time point of cell fate decisions.

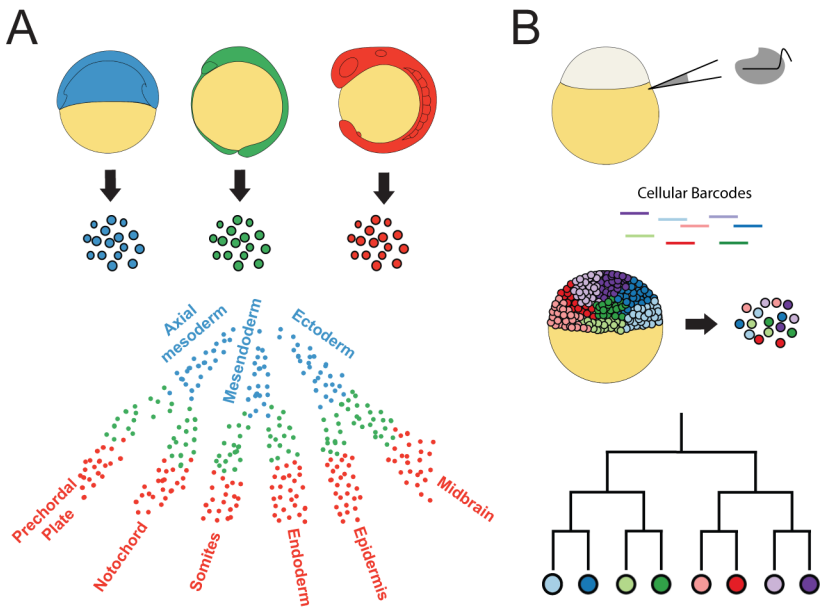

Figure 1. Cell fate dynamics and lineage. (A) A cell atlas of early development: shield (6 hpf, blue), Bud (10 hpf, green) and 10-somite (14 hpf, red) embryos are dissociated into a single cell suspension and the transcriptomics profiles are used to build cell fate trajectories for some embryonic structures. (B) Workflow of a linage tracing experiment: a 1-cell stage embryo is injected with Cas9/gRNA to introduce indels in specific loci of the genome during the first cell divisions (colored lines). These cellular barcodes will give information about the history of each cell to construct a lineage tree.

Spatial transcriptomics

An obvious requirement for single cell genomics is the dissociation of the sample, and hence, spatial information is lost. However, a remedy to this barrier is to map individual cell transcriptomic profiles to an approximate location by computationally inferring the position in the embryo using known landmark genes obtained by whole-mount *in situ* hybridization as reference[35]. As a proof of principle, the authors of this study used a two-dimensional space for 50% epiboly stage embryos (anteroposterior/dorsoventral, since the right-left axis has not been established and genes are expressed homogenously in the depth axis) to successfully map the transcriptomic profile of the individual cells (Figure 2A). Another method (novoSpaRc) reported to use fewer landmark genes to reconstruct the expression patterns from the same data set by matching similarities in the transcriptomic profile to the physical neighborhood in a virtual embryonic space [36]. Even a coarse *de novo* reconstruction (with no landmark genes) for a two-dimensional space is possible, as shown for an example using the Drosophila embryo.

However, for samples with a more complex or more variable spatial architecture, methods that provide direct experimental measurements of spatial information are required. A variety of technologies have been developed in recent years, ranging from direct in-situ sequencing [37] to microdissection of tissue (Figure 2B) [38]. Other methods add molecular barcodes encoding spatial information directly in tissue slices. These approaches are either based on arrays of barcoded polyT primers

spotted on a surface [39] or on barcoded beads that are positioned on a surface [40]. Another class of methods uses sequential rounds of single-molecule FISH to detect transcript molecules by microscopy. While these methods were previous limited to low numbers of genes, transcriptome-wide transcription imaging was recently reported [41]. Importantly, imaging-based techniques have much higher transcript recovery rates than sequencing based approaches, as they are based on probe hybridization rather than on the relatively inefficient reverse transcription reaction. Which approach is most suitable for a given biological question depends on the required spatial resolution, the necessary transcript recovery rate, and on the expected complexity of the spatial patterns. Furthermore, approaches based on landmark genes are suitable for reproducible samples like embryos, but are less useful for variable structures like tumors. Finally, these methods are vastly different in experimental complexity.

The different approaches in spatially-resolved transcriptomics allow us to identify the location of cell types in the tissue and to analyze spatial tissue remodeling in disease conditions, which is important for a large range of questions. However, as the technologies for spatial transcriptomics are maturing, applications beyond descriptive spatial atlases are becoming possible. In particular, we anticipate that transcriptome-wide spatial methods will soon allow detailed insights into cell-cell communication, by enabling a systematic analysis of ligand-receptor interactions between pairs of cells that are in direct physical contact.

Remaining challenges and emerging technologies

Analysis at the single-cell level has emerged as a powerful tool to address fundamental biological questions. However, being a relatively new technique, it is understandable that there is still a lot of room for improvement. One prevalent issue with scRNA-seq is the low transcript detection efficiency: typically, only around 1% of the cellular mRNA are recovered in a large-scale scRNA-seq experiment, which leads to considerable sampling noise [42]. While, state-of-the-art plate-based methods can have detection efficiencies that are one order of magnitude higher [43], hybridization-based approaches are the method of choice if precise quantification of lowly expressed genes is required. However, it is important to note that, to some degree, low transcript recovery can be compensated by analyzing large numbers of cells by scRNA-seq. For example, a large single-cell atlas of 2 million mouse embryo cells offers superb cell type resolution despite very sparse transcript coverage [44].

Some notable discrepancies in the Ensembl and RefSeq zebrafish annotations have been described and the source is an incomplete or lack of annotation on the 3' untranslated region (UTR) of many genes: 20% and 6%, respectively [45]. Since scRNA-seq usually relies on 3' end-sequencing, this is a relevant issue to have in mind. The authors of this work generate a more complete zebrafish transcriptome annotation with higher 3' UTR coverage using existing bulk RNA-seq datasets at several stages of embryogenesis. They demonstrate that this new annotation helps identify more clusters and capture information of more cells in scRNA-seq datasets when compared to the Ensembl annotation.

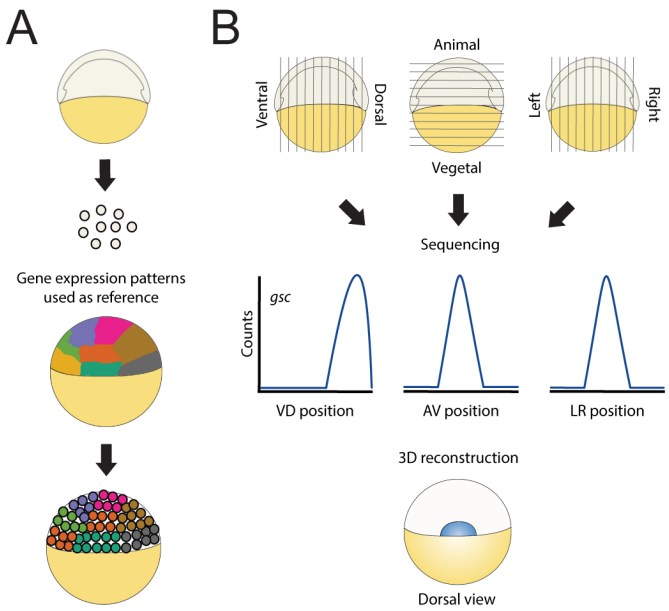

Figure 2. Spatial transcriptomics in the zebrafish embryo. (A) A 50% epiboly stage embryo is dissociated and the transcriptomic profile of the individual cells is compared to known gene expression patterns to map their approximate location. (B) Tomo-seq: embryos from the same stage are sliced along the three main axis and the individual slices are sequenced. The information can be used to reconstruct the spatial pattern of any gene, e.g. the dorsally expressed *goosecoid* (*gsc*).

Except for a few cases, cell type annotation after cluster identification remains a challenging task: The differentially expressed genes in detected cell type clusters often do not correspond to established cell type marker genes, partially because the low transcript detection rate impedes capture of informative transcription factors, as their expression is generally lower than that of structural genes. Furthermore, the identification of new cell types is hindered by the ambiguous definition of “cell type” versus “cell state”. Relying on the activity of transcription factors or gene regulatory networks more than the whole set of differentially expressed genes to identify cell types might offer an alternative to approach this issue [32,13]. But an evolutionary definition and characterization of cell types would help in this matter [46]: just like the sequencing of many genomes allowed the identification of orthologous genes, finding the same cell types in closely related species would provide a good validation for their existence, especially for rare cell types, and would provide a more precise blueprint of their transcriptomic profile. Recent efforts in this direction comparing the cellular composition of the hypothalamus of the zebrafish to the cavefish Mexican tetra have reported interesting findings on how orthologous cell types can exhibit a different regulatory network of effector genes, such as neuropeptides [47].

While scRNA-seq is by far the most advanced technology in single cell genomics, measurement of other parameters is rapidly catching up. Besides DNA methylation and protein detection, single cell open chromatin profiling is becoming particularly widely used. Single cell ATAC-seq, a transposase-based method for open chromatin profiling, can now routinely be performed in thousands of cells due to new protocols for combinatorial barcoding of single cells as well as droplet based methods [48]. Applications include atlases of chromatin accessibility in mouse [49] and in drosophila

development [50]. Recently, scATAC-seq was applied to zebrafish embryos at 24 hours post fertilization [51]. This data can also be directly used for transgenic reporter gene design by allowing a more rational choice of which genomic fragment to use to drive expression of fluorescent reports.

Conclusions

The emergence and adoption of single-cell technologies by many laboratories around the world in the last few years has been a very rapid process. Within a few years, single cell genomics, and in particular single cell transcriptomics, has turned from a highly specialized method into a widespread approach. As the cost of single-cell sequencing decreases, and as computational analysis tools become easier to use, single-cell genomics will become even more widely used. We anticipate that future developments regarding inclusion of spatial and temporal information in single cell genomics will make this approach even more attractive for studying developmental cell fate decisions, and the zebrafish will certainly continue to have a prominent place in the ongoing single-cell revolution.

References

1. Griffiths JA, Scialdone A, Marioni JC. Using single-cell genomics to understand developmental processes and cell fate decisions. *Mol Syst Biol*. 2018;14(4):1-12. doi:10.15252/msb.20178046
2. McKenna A, Gagnon JA. Recording development with single cell dynamic lineage tracing. *Dev*. 2019;146(12):1-10. doi:10.1242/dev.169730
3. Kester L, van Oudenaarden A. Single-Cell Transcriptomics Meets Lineage Tracing. *Cell Stem Cell*. 2018;23(2):166-179. doi:10.1016/j.stem.2018.04.014
4. Adam M, Potter AS, Potter SS. Psychrophilic proteases dramatically reduce single cell RNA-seq artifacts: A molecular atlas of kidney development. *Development*. 2017;1:dev.151142. doi:10.1242/dev.151142
5. Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Type. 2014;(February).
6. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171-181. doi:10.1038/nprot.2014.006
7. Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202-1214. doi:10.1016/j.cell.2015.05.002
8. Klein AM, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 2015;161(5):1187-1201. doi:10.1016/j.cell.2015.04.044
9. Grün D, Van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell*. 2015;163(4):799-810. doi:10.1016/j.cell.2015.10.039
10. Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science (80-)*. 2018;360(6392):eaar3131. doi:10.1126/science.aar3131
11. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science (80-)*. 2018;360(6392):981-987.

doi:10.1126/science.aar4362
12. Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat Methods*. 2020;17(2):147-154. doi:10.1038/s41592-019-0690-6
13. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: Single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14(11):1083-1086. doi:10.1038/nmeth.4463
14. Kamimoto K, Hoffmann CM, Morris SA. CellOracle: Dissecting cell identity via network inference and in silico gene perturbation. *bioRxiv*. 2020. doi:10.1101/2020.02.17.947416
15. Packer JS, Zhu Q, Huynh C, et al. A lineage-resolved molecular atlas of C. Elegans embryogenesis at single-cell resolution. *Science (80-)*. 2019;365(6459). doi:10.1126/science.aax1971
16. La Manno G, Soldatov R, Zeisel A, et al. RNA velocity of single cells. *Nature*. 2018;560(7719):494-498. doi:10.1038/s41586-018-0414-6
17. Erhard F, Baptista MAP, Krammer T, et al. scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature*. 2019;571(7765):419-423. doi:10.1038/s41586-019-1369-y
18. Hendriks G-J, Jung LA, Larsson AJM, et al. NASC-seq monitors RNA synthesis in single cells. 2018. doi:10.1101/498667
19. Lange M, Bergen V, Klein M, et al. CellRank for directed single-cell fate mapping. *bioRxiv*. 2020. doi:10.1101/2020.10.19.345983
20. Cao J, Zhou W, Steemers F, Trapnell C, Shendure J. Sci-fate characterizes the dynamics of gene expression in single cells. *Nat Biotechnol*. 2020;38(8):980-988. doi:10.1038/s41587-020-0480-9
21. Sulston J, Schierenberg E, White J. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol*. 1983;100(1):64-119. http://ukpmc.ac.uk/abstract/MED/6684600.
22. Kimmel CB, Law RD. Cell lineage of zebrafish blastomeres. III. Clonal analyses of the blastula and gastrula stages. *Dev Biol*. 1985;108(1):94-101. doi:10.1016/0012-1606(85)90012-0
23. Pan AY, Freundlich T, Weissman TA, et al. Zebrabow: Multispectral cell labeling for cell tracing and lineage analysis in zebrafish. *Dev*. 2013;140(13):2835-2846. doi:10.1242/dev.094631
24. Biddy BA, Kong W, Kamimoto K, et al. Single-cell mapping of lineage and identity in direct reprogramming. *Nature*. 2018;564(7735):219-224. doi:10.1038/s41586-018-0744-4
25. Alemany A, Florescu M, Baron CS, Peterson-Maduro J, Oudenaarden A Van. Whole-organism clone tracing using single-cell sequencing. *Nature*. 2018. doi:10.1038/nature25969
26. Spanjaard B, Hu B, Mitic N, et al. Simultaneous lineage tracing and cell-type identification using CRISPR – Cas9-induced genetic scars. 2018;(March). doi:10.1038/nbt.4124
27. McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science (80-)*. 2016;353(6298). doi:10.1126/science.aaf7907
28. Raj B, Wagner DE, McKenna A, et al. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat Biotechnol*. 2018;36(5):442-450. doi:10.1038/nbt.4103

29. Chen Z, Ye C, Liu Z, Deng S, He X, Xu J. Lineage recording of zebrafish embryogenesis reveals historical and ongoing lineage commitments. 2020. doi:10.1101/2020.07.15.203760
30. Salvador-Martínez I, Grillo M, Averof M, Telford MJ. Is it possible to reconstruct an accurate cell lineage using CRISPR recorders? *bioRxiv*. 2018:373357. doi:10.1101/373357
31. Takasugi PR, Drage EP, Kanishka SN, Higbee MA, Gagnon JA. Orthogonal CRISPR-Cas genome editing and efficient inhibition with anti-CRISPRs in zebrafish embryos. 2020:1-9.
32. Weinreb C, Rodriguez-fraticelli A, Camargo FD, Klein AM. Lineage tracing on transcriptional landscapes links state to fate during differentiation. 2020;3381(January):1-15.
33. Hu B, Lelek S, Spanjaard B, et al. Cellular drivers of injury response and regeneration in the adult zebrafish heart. *bioRxiv*. 2021:2021.01.07.425670. https://doi.org/10.1101/2021.01.07.425670.
34. Simeonov KP, Byrns CN, Clark ML, et al. Single-cell lineage and transcriptome reconstruction of metastatic cancer reveals selection of aggressive hybrid EMT states. 2020:1-36. doi:10.1101/2020.08.11.245787
35. Satija R, Farrell J a, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33(5). doi:10.1038/nbt.3192
36. Nitzan M, Karaiskos N, Friedman N, Rajewsky N. Gene expression cartography. *Nature*. 2019;576(December). doi:10.1038/s41586-019-1773-3
37. Lee JH, Daugharthy ER, Scheiman J, et al. Sequencing in Situ. *Science*. 2014;343(March):1360-1363. doi:10.1126/science.1250212
38. Junker JP, Noël ES, Guryev V, et al. Genome-wide RNA Tomography in the Zebrafish Embryo. *Cell*. 2014;159(3):662-675. doi:10.1016/j.cell.2014.09.038
39. Ståhl PL, Salmén F, Vickovic S, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science (80-)*. 2016;353(6294):78-82. doi:10.1126/science.aaf2403
40. Rodriques SG, Stickels RR, Goeva A, et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science (80-)*. 2019;363(6434):1463-1467. doi:10.1126/science.aaw1219
41. Eng CHL, Lawson M, Zhu Q, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature*. 2019;568(7751):235-239. doi:10.1038/s41586-019-1049-y
42. Grün D, Kester L, van Oudenaarden A. Validation of noise models for single-cell transcriptomics. *Nat Methods*. 2014;11(6):637-640. doi:10.1038/nmeth.2930
43. Hagemann-Jensen M, Ziegenhain C, Chen P, et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat Biotechnol*. 2020;38(6):708-714. doi:10.1038/s41587-020-0497-0
44. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496-502. doi:10.1038/s41586-019-0969-x
45. Lawson ND, Li R, Shin M, et al. An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes. *Elife*. 2020;9:1-76. doi:10.7554/ELIFE.55792
46. Arendt D, Bertucci PY, Achim K, Musser JM. Evolution of neuronal types and families. *Curr Opin Neurobiol*. 2019;56:144-152.

doi:10.1016/j.conb.2019.01.022
47. Shafer MER, Sawh AN, Schier AF. Gene family evolution underlies cell type diversification in the hypothalamus of teleosts. *BioRxiv*. 2020:0-2.
48. Chen X, Miragaia RJ, Natarajan KN, Teichmann SA. A rapid and robust method for single cell chromatin accessibility profiling. *Nat Commun*. 2018;9(1):5345. http://www.nature.com/articles/s41467-018-07771-0.
49. Cusanovich DA, Hill AJ, Aghamirzaie D, et al. A Single-Cell Atlas of In Vivo Mammalian Chromatin Accessibility. *Cell*. 2018;174(5):1309-1324.e18. doi:10.1016/j.cell.2018.06.052
50. Cusanovich DA, Reddington JP, Garfield DA, et al. The cis-regulatory dynamics of embryonic development at single-cell resolution. *Nature*. 2018;555(7697):538-542. doi:10.1038/nature25981
51. McGarvey AC, Kopp W, Vučićević D, et al. Single-cell-resolved dynamics of chromatin architecture delineate cell and regulatory states in wildtype and cloche/npas4l mutant zebrafish embryos. 2020.