Visualization of Single-Cell Transcriptomics from Zebrafish Pigment Cells

Diogo Esteves[1][0000−0001−5741−5686] & David Henriques[2][0000−0002−9477−292X]

[1] Department of Informatics, University of Minho, Portugal
pg28935@uminho.pt
[2] CSIC IIM, Vigo, Spain
davidh@iim.csic.es

Abstract. Zebrafish (*Danio rerio*) represent an essential model organism for studying the mechanisms of disease and vertebrate development. Their unique coloring patterns offer special insights into genetic control and cellular differentiation. To investigate the complexity of pigment cell development in zebrafish, this work uses single-cell transcriptomics, which provides hitherto unheard-of precision into the dynamics of cellular heterogeneity and gene expression. Using state-of-the-art tools such as scRNA-seq and sophisticated analytical techniques like UMAP and t-SNE, we want to clarify the processes involved in pigment cell development while emphasizing the function of transcription factors and genetic networks. Our work aims to provide a deeper understanding of the genetic origins of pigmentation, which could have implications for biomedicine, in addition to filling important knowledge gaps in developmental biology. This comprehensive approach is required for the investigation of genetic variation and cellular development within this well-studied model organism.

1 Introduction

Studying the pigmentation of *D. rerio* cells offers valuable insight into genetic and developmental processes, in addition to being aesthetically pleasing. *D. rerio* is a crucial model organism in biology, understanding its genetic regulation and cellular differentiation can be visually inferred from its distinct pigmentation patterns, which are controlled by various types of pigment cells. As recent research has shown [13, 2], these patterns originate from a highly multipotent progenitor. Although single-cell transcriptomics offers a detailed view of cellular heterogeneity and gene expression dynamics, conventional bulk mRNA sequencing techniques fail in capturing the cellular complexity and dynamic nature of living tissues [5].

A recent technique, called single-cell transcriptomics, makes it possible to examine gene expression at the individual cell level, revealing the variety and dynamics of cellular states [5]. According to Saunders *et al.* [9], this method is crucial for tracking the lineage and differentiation pathways of pigment cells in *D. rerio*, offering previously unseen clarity [2].

2 State of the Art

Sur *et al.* (2023)'s study provides a detailed analysis of shared signatures and transcriptional diversity during the development of *Danio rerio*, highlighting the complexity of cell development trajectories captured by advanced technology of single cell transcriptomy [14]. These findings are parallel to those observed in studies on sensory nodes in vertebrates, where Vermeiren *et al.* (2020) discussed how transcription programs can direct functional specialization, suggesting similar underlying genetic mechanisms that may be in play in the differentiation of pigment cells in zebrafish [15].

2.1 Single Cell Transcriptomics

With the rapid advancement of single-cell transcriptomics, which provides insights into cellular heterogeneity and gene expression patterns within organisms, the field has gone from being an innovative technique to an essential part of contemporary biological research. Novel approaches that have deepened our understanding of cellular processes have fueled this advancement [4].

2.2 Advancements in Single-cell Technologies

Thanks to technological developments like the 10x Genomics Chromium system, which have enabled thorough molecular profiling at high throughput, the pigmentation of zebrafish and the underlying genetics have progressed [10]. In a short period of time, single-cell transcriptomics has evolved from a novel technique to a standard in biological research. This technique has made it possible to thoroughly examine gene expression at the level of individual cells, giving previously unachievable insights into the heterogeneity of cells within organisms.

Innovations such as Drop-seq and the 10x Genomics Chromium system have been essential to open a window into the cellular heterogeneity and transcriptional landscapes of developing organisms [5]. These advances have been essential in deciphering the complex mechanisms behind the pigmentation of zebrafish and have improved our understanding of the effects of both genetics and the environment on the differentiation of pigment cells [7].

2.3 Algorithms for the analysis of single cell data

The proliferation of single-cell data has necessitated the development of advanced analytical methods to interpret the complex datasets generated by scRNA-seq. Dimensionality reduction techniques such as t-SNE and UMAP are now essential for reducing high-dimensional data into a form that can be understood in order to help identify different populations and cell states [4, 5].

These techniques facilitate the identification and distinction of various cell populations and states in a variety of samples. Because UMAP can handle larger datasets without compromising the data's global structure, it is especially well

suited for integrating datasets under a variety of experimental conditions. However, careful parameter selection and data pre-processing are required to yield meaningful insights [5]. Moreover, the comprehensive integration of single-cell

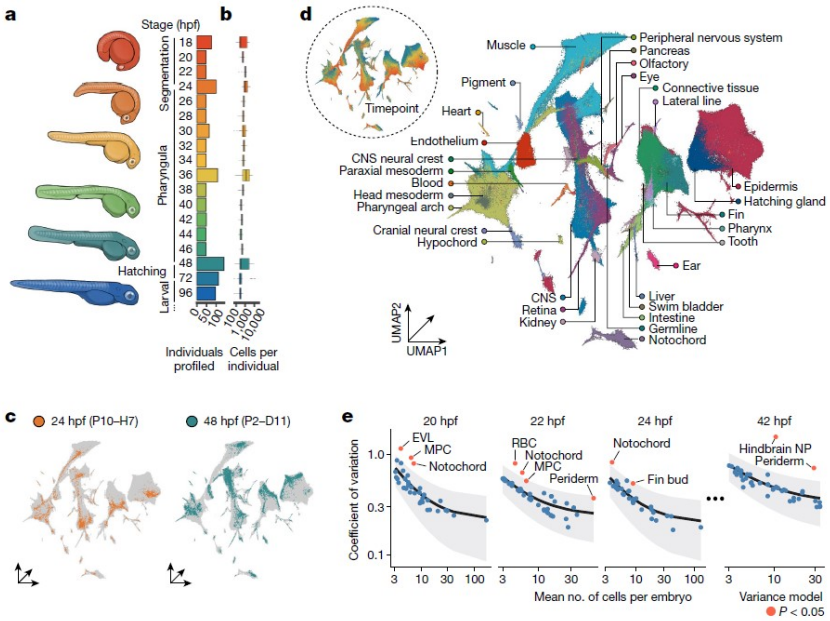

Fig. 1. (a) Development stages of zebrafish from segmentation to larval stages with hours post-fertilization (hpf). (b) Number of cells profiled per individual at each stage. (c) UMAP projections at 24 hpf and 48 hpf showing developmental cell transitions. (d) UMAP visualization highlighting the diversity of cell types and their developmental pathways. (e) Analysis of cell type variance at different developmental timepoints, illustrating the dynamics of cell type distribution. Image adapted from Saunders *et al.*, 2023 [9].

data, as discussed by Stuart *et al.* [12], highlights the importance of merging various types and data sources to enhance the understanding of complex biological systems. This integration is important for the advance in our ability to decipher the multifaceted nature of genetic regulation and the development of cellular behavior. Stuart *et al.* emphasize the utility of combining single-cell transcriptomics data with other data types to provide a more holistic view of cellular functions and states, which is crucial for our analysis of zebrafish pigmentation patterns.

Variations of t-SNE and UMAP have been implemented in several tools such as Slingshot and Monocle together with pseudo-time analysis [7, 11]. These tools are a good match to explore the developmental pathways of zebrafish pigment

cells by methodically arranging cells according to their similarity thanks to these methodologies [5, 4].

3 Objective

The main goal of this study is to close the knowledge gaps on the dynamics of gene expression during pigment cell development in *D. rerio* and cellular heterogeneity. By doing this, we hope to clarify the processes underlying pigment cell development and identify the regulatory systems at play.

In addition to exploring the cellular differentiation trajectories, a critical objective of this study is to investigate the dynamics of transcription factors during these processes. Kenny *et al.* (2022) show that TFAP2 paralogs facilitate the access to MITF chromatin in pigmentation and cell proliferation genes, highlighting the significant role of these factors in modulation of gene expression during pigment cell differentiation.

These findings underscore the importance of analyzing the function of transcription factors in the development of pigment cells in *Danio rerio*.

3.1 Visualize Cellular Differentiation Trajectories

We intend to use computational tools such as UMAP and t-SNE in conjunction with trajectory inference techniques like Monocle [7] to visualize and analyze cellular differentiation trajectories leading to pigment cell development in *D. rerio*. This will enable us to distinguish between different cellular states and transitions during the differentiation process, revealing the intricate regulatory mechanisms driving pigment cell formation.

We will use tools such as Slingshot [11] and Monocle, along with dimensionality reduction techniques and pseudo-time analysis, to accomplish our goals. With the help of these techniques, we will be able to map out the developmental pathways of pigment cells by methodically arranging cells according to how similar they are.

Our goal is to improve our understanding of the mechanisms that regulate zebrafish pigmentation by identifying cell-type lineages and highlighting important gene markers indicative of pigment cell differentiation through the use of these bioinformatics tools.

3.2 Extract Information on Transcription Factor Dynamics

Our objective is to examine and describe modifications in the expression of transcription factors (TF) along the recognized cellular developmental pathways. Finding the important transcription factors (TF) that control pigment cell differentiation will provide information on the genetic networks that govern this developmental process [13, 8, 1].

With potential applications in disease treatment, cosmetic research, and environmentally friendly aquaculture practices, these findings may open the door for novel genetic interventions that manipulate pigment cell formation and function.

4 Tasks

4.1 Visualization of Cellular Differentiation Trajectories

High-throughput single-cell RNA sequencing will be used to investigate pigment cell development in detail, with an emphasis on identifying important transcription factors and genes involved in cell differentiation [10, 4]:

- **Single-Cell RNA Sequencing** Our goal is to produce a dataset that outlines each cell's transcriptional landscape as it differentiates into *D. rerio* pigment cells. Using high-throughput single-cell RNA sequencing technologies, *D. rerio* samples from different embryonic stages are processed, with a focus on tissues rich in pigment [10].

- **Dimensionality Reduction and Clustering** We will use dimensionality reduction methods like UMAP and t-SNE on the scRNA-seq data in order to identify different cell populations and states within the intricate dataset. Indicative of similar biological states or functions, this will assist in identifying cell populations [4, 5].

- **Trajectory Inference Analysis** We intend to map pigment cell developmental trajectories, following their lineage from progenitors to mature states, using trajectory inference tools such as Monocle. This will allow us to identify regulatory genes and important transition states involved in pigment cell development, as well as to analyze the temporal progression of cells along the differentiation pathways [7].

4.2 Transcription Factor Dynamics

In order to analyze the temporal progression of cells along differentiation pathways and identify important transition states and regulatory genes involved in the development of pigment cells, we will use trajectory inference tools such as Monocle [7].

- **Differential Expression Analysis** To find TFs and genes that are differentially expressed, we will carry out differential expression analysis both across cell clusters and along inferred trajectories. The results of this analysis will identify potential players in the regulation of pigment cell differentiation. Using software intended for RNA-seq data analysis in a single cell, we will employ computational packages [5, 4].

– **Validation of Key Transcription Factors** We will use in situ hybridization and immunocytochemistry techniques to validate the spatial and temporal expression patterns of important TFs in *D. rerio* tissue sections in order to confirm the role of identified TFs in pigment cell differentiation. This stage establishes the empirical foundation of the study by guaranteeing that computational predictions faithfully capture biological reality [6].

In the next phase of the project, where we focus on the dynamics of transcription factors, we take the insights of Kenny *et al.* (2022) will guide our analyzes. Specifically, we will seek to validate the influence of identified transcription factors, such as TFAP2, on tissue samples of *Danio rerio*, using in situ hybridization and immunocytochemical techniques to confirm [3] expression patterns.

During data analysis, we will apply trajectory inference techniques, as highlighted by Street *et al.* (2018), to better understand cell lines and development paths. This approach will help us identify key transitions in pigment cell development, supplementing our analysis of transcription factors with a more comprehensive view of cell differentiation.

5 Planning and Execution

Following the initial data collection, the project enters a rigorous data analysis phase that lasts from weeks 1 through 4. In this time frame, we utilize dimensionality reduction techniques such as t-SNE and UMAP to reveal the various cellular environments concealed in zebrafish pigmentation patterns. In parallel, we track the lineage and maturation of individual pigment cells using trajectory analysis tools such as Monocle, giving a dynamic picture of their evolution over time.

Weeks five through seven are when we will work on the dynamics of transcription factors. In this second phase, we want to find the transcription factors that are necessary for the differentiation of pigment cells. To identify crucial regulatory mechanisms, the gene expression profiles of different cell types and states are compared.

We synthesize our computational results in advance of the May 29 presentation, which concludes week 8. The purpose of this synthesis is to provide strong evidence that the transcription factors that we have discovered are essential for the growth of pigment cells.

Weeks 9 and 10 comprise the project's final phase, which culminates in the final submission on June 19. We complete our documentation and write detailed reports during this period.

Our goal is to successfully communicate our findings through these reports, which provide an overview of the project's insights and overall results. This systematic strategy ensures a well-coordinated project, with each step being completed within the allocated weeks for maximum efficiency.

References

1. Fabian, P., Tseng, K.C., Thiruppathy, M., Arata, C., Chen, H.J., Smeeton, J., Nelson, N., Crump, J.G.: Lifelong single-cell profiling of cranial neural crest diversification in zebrafish. Nature communications **13**(1), 13 (2022)
2. Howard IV, A.G., Baker, P.A., Ibarra-García-Padilla, R., Moore, J.A., Rivas, L.J., Tallman, J.J., Singleton, E.W., Westheimer, J.L., Corteguera, J.A., Uribe, R.A.: An atlas of neural crest lineages along the posterior developing zebrafish at single-cell resolution. Elife **10**, e60005 (2021)
3. Kenny, C., Dilshat, R., Seberg, H.E., Van Otterloo, E., Bonde, G., Helverson, A., Franke, C.M., Steingrímsson, E., Cornell, R.A.: Tfap2 paralogs facilitate chromatin access for mitf at pigmentation and cell proliferation genes. PLoS genetics **18**(5), e1010207 (2022)
4. Kulkarni, A., Anderson, A.G., Merullo, D.P., Konopka, G.: Beyond bulk: a review of single cell transcriptomics methodologies and applications. Current opinion in biotechnology **58**, 129–136 (2019)
5. Nayak, R., Hasija, Y.: A hitchhiker's guide to single-cell transcriptomics and data analysis pipelines. Genomics **113**(2), 606–619 (2021)
6. Phipson, B., Sim, C.B., Porrello, E.R., Hewitt, A.W., Powell, J., Oshlack, A.: propeller: testing for differences in cell type proportions in single cell data. Bioinformatics **38**(20), 4720–4726 (2022)
7. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., Trapnell, C.: Reversed graph embedding resolves complex single-cell trajectories. Nature methods **14**(10), 979–982 (2017)
8. Saunders, L.M., Mishra, A.K., Aman, A.J., Lewis, V.M., Toomey, M.B., Packer, J.S., Qiu, X., McFaline-Figueroa, J.L., Corbo, J.C., Trapnell, C., et al.: Thyroid hormone regulates distinct paths to maturation in pigment cell lineages. Elife **8**, e45181 (2019)
9. Saunders, L.M., Srivatsan, S.R., Duran, M., Dorrity, M.W., Ewing, B., Linbo, T.H., Shendure, J., Raible, D.W., Moens, C.B., Kimelman, D., et al.: Embryo-scale reverse genetics at single-cell resolution. Nature **623**(7988), 782–791 (2023)
10. Srivatsan, S.R., McFaline-Figueroa, J.L., Ramani, V., Saunders, L., Cao, J., Packer, J., Pliner, H.A., Jackson, D.L., Christiansen, L., et al.: Massively multiplex chemical transcriptomics at single-cell resolution. Science **367**(6473), 45–51 (2020)
11. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., Dudoit, S.: Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC genomics **19**, 1–16 (2018)
12. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., Satija, R.: Comprehensive integration of single-cell data. cell **177**(7), 1888–1902 (2019)
13. Subkhankulova, T., Camargo Sosa, K., Uroshlev, L.A., Nikaido, M., Shriever, N., Kasianov, A.S., Yang, X., Rodrigues, F.S., Carney, T.J., Bavister, G., Schwetlick, H., Dawes, J.H.P., Rocco, A., Makeev, V.J., Kelsh, R.N.: Zebrafish pigment cells develop directly from persistent highly multipotent progenitors. Nature Communications **14**(1), 1258 (2023)
14. Sur, A., Wang, Y., Capar, P., Margolin, G., Prochaska, M.K., Farrell, J.A.: Single-cell analysis of shared signatures and transcriptional diversity during zebrafish development. Developmental Cell **58**(24), 3028–3047 (2023)

15. Vermeiren, S., Bellefroid, E.J., Desiderio, S.: Vertebrate sensory ganglia: common and divergent features of the transcriptional programs generating their functional specialization. Frontiers in Cell and Developmental Biology **8**, 587699 (2020)