Laboratory of Molecular Genetics
Department of Biomedicine
Medical University Lublin
Collegium Universum
Chodzki 1, 20-093 Lublin, Poland

Akceptuję

Review a PhD thesis by Eugeniusz Tralle

Ph review of a thesis titled " Cell lineage tracing in zebrafish heart development"

Eugeniusz Tralle submitted a PhD thesis manuscript titled "Cell lineage tracing in zebrafish heart development" for a review. The thesis has a traditional structure and consists of the Introduction and two chapters with their own substructure of Materials and Methods, followed by Results and a chapter-specific Discussion and Future Work. The last chapter provides general Conclusions to the thesis.

Two experimental chapters are not linked in any obvious way and presumably describe two parallel project the Candidate has been pursuing during his PhD.

The first experimental chapter describes the search for enhancers driving the heart-specific expression of isl1a gene. To accomplish that, the Candidate used a zebrafish-based enhancer screen to test several candidate enhancers identified by genome analysis. Next, both transient and transgenic lines were explored to investigate the specificity of the enhancers. The second experimental chapter describes the lineage analysis of the developing zebrafish cardiac precursors using single cell RNA Seq and software-based lineage analysis. Here, the Candidate generated two sources of the single cell pools. One wild type and one with morphants having reduced expression of nkx2.5/nkx2.7.

The submitted thesis is well written and methodologically rich. Nonetheless, as a reviewer I have several comments that I feel require clarification. I divided my comments into the corresponding sections of the thesis.

List of abbreviations

- Why the list is NOT in an alphabetical order

Introduction

The introduction is divided in several sections. It begins with the general description of the zebrafish model. Next it describes the heart development and the ontology of the second heard field. Finally, it describes the concept of single cell RNA sequencing. I feel however that the thesis misses a few essential points:

- Entire second chapter is devoted to the search for enhancers. The Introduction however misses completely the explanation what the enhancers are, how do they work how are they detected and how they can be studied. This is a significant omission since the reader is later on faced with the data analysis, not being able to appreciate the importance of the finding. I would like to hear fomr the Candidate a concise enhancer introduction
- I am surprised by one panel of the Figure 2. The transgene expression at 15 hpf (incidentally, which stage that is?) is not corresponding to neither the endogenous expression not the transgenic expression published in http://dx.doi.org/10.1016/j.ydbio.2014.12.019 . I would like to know where is this discrepancy coming from?
- The part about the single cell sequencing is rather poor. It does not explain the principle of the method, latest developments like spatial transcriptomics not the pitfalls associated with the method. I would like to hear more about these aspects of the scRNA seq.

Aims

The Aims section is a little bit chaotic. While the Candidate describes the both lines of research, I fail to see any link between them and the Aims section would be an ideal place to link them. There are several ways to approach it, I would like to hear about at least one

Materials and Methods

The section is written in general carefully but a few issues remain

- Section 2.1.2- I understand the need to identify the SHF enhancers but it is not stated how they were initially defined. H3K27ac is a general signature all over the genome. How the Candidate defined SHP enhancers in the search? Was it a global search or only around the *Isl1a* gene?
- Were the enhancer constructs linearized for the co injection with Tol2? I seem to fail to find that information. It is important that they are linear in such a protocol.
- I am surprised that the post injection selection was made only 24h after the injection and not after 8 hours where the GFP is already visible.

Results Chapter 2

- I am missing the Topologically Active Domain (TAD) and/or 4C analysis which would show which region is likely to be involved in the gene regulation, simplifying the search and adding

to the biological credibility of the result. I would like to see such a result. While such data mayb be missing for the zebrafish, they are available for the mouse and human. Taking under the consideration the conservation of gene regulation, I would consider such data relevant.

- Extensive work was done on the effect of GFP alone which is a pity since it took a lot of time and such date are widely available. The reporter construct used by the candidate is commonly used. Why so much time and effort went into this work?
- Fig 8- what is the faint green expression outside of the heart. It does not look like a background but rather a low level but specific epression
- Fig 9. There is only about 30% overlap between the myl7 and I3 driven expression with the latter being much broader. Please, identify the positive tissues. A bright field picture would be very helpful here.
- Fig 11. The negative error bars suggest negative data points. Please, explain that.
- Stating on page 42 "However, the absence of a specific SHF marker did not allow us to rule out the possibility that additional cells other than those of the SHF are also present within the GFP+ population." leaves me a bit confused. So, has the Candidate isolate a specific enhancer or an enhancers with a partial specificity?
- It appears that the MEF2C is the key transcription factor binding to the I3 enhancer and regulating the gene expression. I would like to know how to prove that hypothesis

Results chapter 3

- There is a nkx2.5 crispant available for 3 years now (at least). Why it was not used instead of the morphant?
- I am missing several morpholino controls. Their absence will not allow for these results to be published in a good journal. Which controls these are and why they were not done? Especially, since the Fig. 16 shows that the morpholinos are very inefficient. What is the explanation for that?
- Instead of using sGESTALT system, one could do lineage tracing based on the scRNASeq data. How this could be done?
- This chapter leaves me with a feeling of an unfinished story. The tool was generate and what now?

Candiate

Laboratory of Molecular Genetics
Department of Biomedicine
Medical University Lublin
Collegium Universum
Chodzki 1, 20-093 Lublin, Poland

I am puzzled by the fact that two publications submitted by the candidate as a part of this PhD. One of them (joint first author, a publication from 2021) I about liver injury and the second is about the innate immune response. Neither of them is a part of this PhD. Conversely, the topic of the PhD, heart development, has not been published by the Candidate. Therefore I would like to hear the explanation about this apparent subject diversity.

Rozprawa doktorska spełnia warunki określone w art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668.

Podsumowując, wnioskuję o dopuszczenie mgr Eugeniusza Tralle do dalszych kroków wymaganych przez proces doktorski w Radzie Dyscypliny Nauk Medycznych Warszawskiego Uniwersytetu Medycznego

Prof. Przemysław Tylżanowski

Tylzanowski Lab
przemko@kuleuven.be