Doc. Hassan Hashimi
České Budějovice
29 April, 202

RE: Review of Andrea Hauserová's bachelor's thesis

The topic of bachelor's thesis of Andrea Hauserová is nicely summarized by its title: "Analyzing the role of p38 mitogen activated protein kinases and their effectors on mouse blastocyst maturation". Here, Andrea describes her contribution to an ongoing project in the Bruce lab investigating the role of the p38-MAPK in pre-implantation mouse embryo development by phenotypic analysis of pharmacologically-inhibited p38-MAPK and RNAi of p38-MAPK effectors Nucleostemin (NS) and an RNA helicase DDX21. Other proteins were also investigated, such as CDX2 and Tead4, although it was not clear how these fit into the main line of investigation nicely stated in the thesis title. The assays performed were based on sophisticated image analysis determining blastocoel volume, cell lineage numbers, lineage-specific gene expression quantification. These data were acquired by confocal microscopy, with acquired z-stacks 3D reconstruction of the assayed embryos. Andrea gives credit to colleagues for demanding methods (*e.g.* injection of dsRNA into blastomeres) that were done prior to her phenotypic analysis, which is a further sign that she acquired the presented data herself.

I will first state what impressed me about the thesis. Clearly Andrea has worked hard in acquiring the presented data, and I am sure also other data that she chose not to include. She has mastered sophisticated microscopy and image analysis methods and has shown the ability to do statistical analysis. The written English is excellent, with only a few grammatical mistakes, such as the occasional missing article. She seems to understand her topic as well, giving a quite detailed introduction. And of course, she made a valuable contribution to a manuscript that was under revision in the journal Communication Biology at the time of writing her thesis. What a great achievement at such an early stage of her career. I congratulate her on this big and much deserved accomplishment.

While I had a very good overall impression of the thesis, there were some flaws that should also be mentioned. The caveat with this part of the report is that pre-implantation mouse embryo is not my topic of expertise. While I know generally how this fascinating system works, I do not know much about the molecular details, mostly addressed in the Introduction, and so my comments here should be considered in this context.

So, while I was impressed with the written English and volume of information in introduction, I felt that the Introduction was more of a list of acts about important molecules in pre-implantation development than what an Introduction should be: a synthesis of past knowledge put in such a way to show how the work extends that knowledge. I admit my drowning in information in this section is certainly due to my lack of expertise in this topic. Nevertheless, I feel that the gold standard for any scientific text is that it can be easily understood by a non-specialist like myself. So the Introduction did not help me to understand why Tead4 and CDX2 were examined in the Results section. Also, it did not really help me to understand why Gata4 and Nanog were assayed in the p38-MAPK inhibited embryos. Were they markers of specific cell lineages? Were they used as outputs of specific processes affected by p38-MAPK?

I should also say that this is a pretty common mistake in theses and even article introductions-the unfocused recounting of information in the field. I would say that a bachelor's thesis is a perfect place to learn, so I do not count this against Andrea. I just mention this for her future work. One way I have found that addresses this problem is to write the Introduction AFTER writing the other

sections, so you know what information to give to the reader instead of thwoing all your knowledge at the reader and hoping that something sticks.

However, the next criticisms are ones that I would like Andrea to address in her rebuttal to my report:

1) In Figure 13, changes in the blastocoel volume are addressed upon p38-MAPK inhibition. However, it is not clear how this was measured. In section 3.6, she sates the 'inner circumference' was measured. Fine, but a representative picture of the embryos measured where this 'inner circumference' is demarked would have been helpful to have seen. This was done by Andrea in Figure 5, showing Gata4 and Nanog IFAs.
2) Andrea claims that DDX21 shifts from a nucleoplasm to nucleolar localization upon p38-MAPk inhibition in Figures 14 and 15. First of all, she does not explain how she determines this. I know from prior experience with other organisms, the nucleolus is the dark part of DAPI stained nuclei. This is not common knowledge and I am not sure if this rule applies also to pre-implantation embryos. Thus, how the reader should identify the nucleolus should have been explained in the text and pointed out in the figure. I should say I was not convinced by this claim in Figure 15. This phenomenon was better shown in Figure 24. In fact, I would have used this Figure instead of Figures 14 and 15 for this reason and also economy (*i.e.* less data evaluated by your reviewer).
3) Sections 4.4 and 4.5 addresses the volume of blastocysts. Again, as I said in point 1 above, a representative image would be appropriate to show here as well as what was measured if not immediately obvious from such an image. Also, section 4.5 is entitled "Clonal downregulation of Tead4, a TE specific transcription factor, also results in smaller blastocysts" and yet Figure 30 demonstrates the opposite conclusion. Which one should I take. Is there an interpretation of this Figure that I am missing?
4) There is no Abbreviation section. Given how many were used in this thesis, I would have found this very helpful.
5) N values (number of samples measured) are often omitted from Figures (*e.g.* Figure 8). I realize I could count the dots, but it is standard to report this number in the legend or even within the figure.

Some minor comments also just for the future:

1) You mention many times your article under revision throughout the text. Once again, this is a great achievement and I congratulate you. But I found this distracting, often breaking the flow of the sentence, making me forgot what the author was trying to convey in the first place. I would have just mentioned this once at the beginning and then cited the manuscript normally. This is perfectly fine since it is available to all on BioRxiv.
2) Table 1 with antibodies. Nice job. I would have found it helpful if you added what the antibodies are supposed to do (*e.g.* cell lineage marker)
3) In section 1.1.5, the author relies heavily on a review by Fuller-Pace for support of various facts about this and other helicases. Better would be to cite primary articles in this case. The reviewer can at least get a clue from cited the article titles.
4) Andrea often reported the P-values of statistically insignificant differences, despite defining these as 'ns' in Table 3. I would use ns or not show the comparison.

To conclude, I was very impressed with Andrea's work in her bachelor's thesis. I think she demonstrates very high potential for a career in science. Nevertheless, I urge her to take my advice about trying to make a relevant synthesis of knowledge instead of painting a very broad picture of various molecules. Such text is hard to read, especially for laymen. All errors I point out here are also advice for her future work. I would recommend that she gets the grade 1 minus with a chance to get the grade1 based on her presentation and answers to the following questions.

I have also sent to Andrea the PDF of her thesis with small comments peppered throughout about English and some very minor issues I chose not to mention in this report.

Sincerely,

Hassan Hashimi

Questions for the candidate:

1) Can Andrea please in simple terms explain the relationship between the molecules examined in her thesis: p38-MAPK, DDX2, NS, CDX2 and Tead4. Also, please summarize what is the significance of Gata4 and Nanog.
2) She mentions that there are 4 isoforms of p38-MAPK. Does the inhibitor used here (SB220025) affect all the isoforms equally? What is the mechanism of p38-MAPK inhibition by this compound? If you needed to dissect the function of the different isoforms, how would you go about doing this (assuming the inhibitor affects all isoforms equally)?
3) In Table 3, Andrea lists degrees of statistical significance by Student's t-test. The two most significant categories are given the same designation "extremely significant", which to me undermines having such a distinction. Can Andrea explain what does the P-value actually signify in layman's terms?
4) In Figure 6 and other figures, Andrea refers to the whiskers as "errors" instead of "error bars" or "standard deviation". What is the middle bar in these figures, mean or mediam values? What is the difference between the two values? What are other ways to show distribution instead of "standard deviation"?