Dear Magdalena,

We would like to submit a revised version of the comparative transcriptome manuscript. As you can see we have shortened it as you requested to three pages. In particular we have enclosed a completed Nature length macro to show you exactly how we have done this.

We have only two figures in our resubmission. One of the figures was well received by the referees and the other is a shortened version of another figure that received some praise from the referees in the first round.

We have also attached a response to the referees in the revision and we hope that you find this helpful. We are confident that the manuscript will meet your approval and we will await your thoughts.

Yours sincerely,

Mark Gerstein

**Detailed Point by Point Responses**

# Reviewer 3

### -- Ref3.1.1 – Lack of Novelty in Figure 1, 2 & 4 --

|  |  |
| --- | --- |
| Reviewer  Comment | If I take this a figure at a time, Fig 1 just shows that there has been a 5-10-fold increase in data quantity; Fig 2 shows that little has changed in terms of gene count or amount of ncRNA (e.g. they say this is in line with what had been previously shown in humans). Fig 4 shows that histone marks are useful in predicting transcriptional state in metazoa. None of these findings is novel/sufficiently new to merit publication in Nature. |
| Author  Response | We agree with the referee about figure 1 and figure 2. They have been moved to ED figures. However, we do feel we have made a point of the HM model in figure 4 that goes beyond previous work. It is true that many studies (including our previous studies) have been done to investigate the quantitative relationship between gene expression and histone modification. But none of them have done this work cross-species, generating "universal" models that are applicable across contexts. Our results are not at all obvious from looking at the performance of the models on individual organisms. To address the referee’s remarks, we have shrunk figure 4 and specifically emphasized the universal model aspect. |

### -- Ref3.1.2 – Figure 3b Top Panel, Fitting Line --

|  |  |
| --- | --- |
| Reviewer  Comment | Fig 3 is potentially more interesting, but I have real issues with Fig 3b, both top and bottom panels along with their analysis. Top panel: how did they fit the purple line to these data? It seems very heavily reliant on a single datapoint, yet ignores all the 13 datapoints clustered at the top left of the plot. Why? This alone looks very suspect to my eyes. |
| Author  Response | The purple line is not a fitting line, but indicates the secondary stage alignment between worm embryo and fly pupae. We have now edited the figure and indicated it with a dashed line and used the same color as the primary alignment. |

### -- Ref3.1.3 – Figure 3b Bottom Panel, Gap --

|  |  |
| --- | --- |
| Reviewer  Comment | In the bottom panel, they highlight a gap in brown - I could have placed the brown panel in several other places in that graph equally well - how did they choose this gap? Is this significantly different to random? |
| Author  Response | The brown bars are not placed arbitrarily, but correspond to the phylotypic stages in each of the two model organisms. The region where the two overlapping brown bars intersect corresponds to the gap in the alignment signal and this is consistent with the observation that the divergence of gene expression across hourglass modules is minimized during the phylotypic stage. But given the referee’s remarks, we have de-emphasized this panel and it is now an insert in the revised figure. |

### -- Ref3.2.1 – Shadow of modENCODE Papers --

|  |  |
| --- | --- |
| Reviewer  Comment | Several of these authors were senior authors on modENCODE papers. They show the increase in data quantity over those papers in Fig 1 (which is significant, but incremental) and emphasise the uniform nature of the processing which allows cross-species analysis. Had those papers not appeared, this paper would certainly have been publishable in  Nature and would certainly have represented a massive advance. However - they have been published. That first major step has already been taken. That the data were not uniformly processed in the first  modENCODE papers was a flaw of those studies and this paper is in some sense what those papers should have been, but rectifying what was a previous flaw does not make this study a huge advance, just a better, more sensible analysis. I truly wish this paper had been in place of the original modENCODE papers - it is more rigorous, the data processing is better, and it is better written and presented. However, it is no longer a major advance in my opinion. |
| Author  Response | We thank the referee for the compliment on our data processing and the data analysis. However, we respectfully disagree with the referee that the previous modENCODE papers overshadow our current study. We think the comparative analysis aspect of this work fundamentally sets it apart from the previous papers that each focused on a single organism. Also in this study, the combination of the large quantity of data (including about 50% new data) and the comparative analyses such as ortholog clustering and universal model is an important step beyond the previous studies. |

### 

### -- Editor’s comments --

|  |  |
| --- | --- |
| Reviewer  Comment | The comparative transcriptome paper (2012-12-15978B-Z; with Mark as the corresponding author on our system) will be resubmitted in condensed format, focusing on the novelty mentioned by Referee no.3; the revised version will include 2 display items and will not exceed 3 pages in print (please use the attached macro to estimate the length of your revised manuscript) |
| Author  Response | We have revised and shrunk the manuscript such that it conforms to the 3-page limit and contains 2 main figures only. This revised manuscript focuses on the novel results , with a brief introduction highlighting the value of the data resource. |