

# Genome Research:

## Comments about structure/Formatting

### Fixes to plots

- **Reviewer 1, comment 3** - suggests relabeling/rescaling y axis of mediators across tissues histogram (Figures 5) -currently in  $\log_{10}(x+1)$  scale )

“Figure 5: a re-labeled y-axis that simply shows the raw counts at the y-ticks would be helpful. It is currently difficult to see how many cases are present in the different categories. Probably most of this figure (from 7 tissues out) is showing a single case per category?”

- **Reviewer 1, comment 4** - spell out 'cor' in figure 4

“Figure 4: spell out “cor” in some of the axis labels”

### Defining Cis and Trans

- **Reviewer 2, comment 1** - suggests defining Cis/Trans early on in the paper and a definition that includes inter vs intra chromosomal interactions

“It would be helpful to explicitly define what is meant by a cis- and trans- gene much earlier in the paper. It seems trans-genes can be on the same chromosome as the eQTL although it can also be on a different chromosome. It would be helpful to see a further categorization of trans- genes based on inter- and intra chromosomal interactions. Further, when the gene and eQTL are on the same chromosome, how much does the distance impact the tendency of a transgene to participate in a particular type of trio.”

### More explanation of MRPC

- **Reviewer 2, comment 6** - Suggests doing more to explain how MRPC incorporates the PMR (even though the method is published...)

“The main method MRPC has been previously published but it would be beneficial to explain how exactly Mendelian Randomization is incorporated into the PC algorithm.”

## Comments about results/validation

### Suggestions about GO analysis

- **Reviewer 2, comment 2** - suggest including more about gene ontology, pathway enrichment for identified mediator genes

“There is very little in terms of interpreting the lists of genes identified in this analyses. It may be helpful to see if the genes participating in specific types of trios are enriched for specific gene ontology processes or other pathways. Similarly, it would be helpful to know what types of processes show up as tissue-specific versus tissue-generic. Right now most of the categorization is at the level of ncRNA vs protein-coding genes. This is not very helpful to understand in terms what processes are regulated.”

- **Reviewer 2, comment 4** - similar to other comments, wants to know if mediators are enriched for regulatory proteins

“ For the mediator trio, is there any enrichment of ”regulatory” proteins to be the mediating genes? This would be interesting and even more, if the regulators had known tissue-specific roles.”

- **Reviewer 3, comment 2** - wants to know if mediators in a given tissue are enriched for tissue-specific regulatory elements

“Among 1323 mediation relationships, are the genes enriched for cell type-specific expression genes or transcription factors. The former might indicate that aspects of the mediation are identifying cell-type-specificity of an eQTL and not a trans-gene mediation. ”

### Identifying validated cases of mediation

- **Reviewer 3, comment 3** - concerned about replicated or validated mediation trios from other studies - perhaps we can find a few examples?

“Given the number of papers that have now looked at aspects of cis- trans-eQTL networks using mediation based approaches, I would now expect that a portion of the discoveries here would be validated or replicated.”

- **Reviewer 3, comment 6** - suggests including information about trans-gene distance from eQTL

“Limited information was provided about the trans-gene distance from the cis-gene.”

### Suggestions/comments about HiC analysis

- **Reviewer 1, comment 2** - confusion about HiC analysis rationale

“I did not follow the rationale for the HiC analyses. Trans regulation is expected to involve a gene product (such as the transcription factor example the authors give themselves in the introduction, or a more indirect mechanism), rather than distant enhancer-promoter contacts. “Distant” is not the

same as “trans”. Even if an enhancer is very far from the gene it regulates, as long as it is on the same DNA molecule, it acts in cis by definition (see <https://www.nature.com/articles/nrg1964> for a refresher). Only interchromosomal DNA contacts would be trans-acting in this context. Indeed, the fact that there was no enrichment of HiC data in the trans-mediation trios illustrates this fact.”

- **Reviewer 2, comment 5** - HiC analysis monte-carlo steps were unclear

“The Hi-C analysis feels a bit preliminary. The authors obtained HiC data for four cell lines and asked if the specific trans-gene eQTL pair could be supported by significantly higher number of reads compared to background. The monte carlo-based method of calculating the p-values based on the neighborhood counts was not clear...”

- **Reviewer 2, comment 5** - suggests that we use a different procedure for the HiC analysis: ” Can the authors use a standard Hi-C peak called and check if the trans-gene and eQTL pair is overlapping?”

“...Can the authors use a standard Hi-C peak called and check if the trans-gene and eQTL pair is overlapping? Perhaps doing this test for all eQTLtrans-gene would have more power than trying to support individual pairs. Also the results can vary depending upon the peak caller so it is recommended to try multiple peak callers for HiC data (e.g. Hiccup and Fithic).”

### Interest in providing more analysis for M3 trios

- **Reviewer 2, comment 3** - suggests more explanation is needed for trios inferred to be conditionally independent (M3) - seems unconvinced that observing most trios being M3 is consistent with literature

“ There is little attempt to explain large proportion of trios that are identified to be conditionally independent. The authors say that there observation is consistent with what is observed in the literature, but does this hold if we were to consider trios of only cis-genes or trans-genes, or more generally not require the trios to have one cis and one trans gene? A related question is how many cis-genes or trans-genes are associated with a specific eQTL. One would expect that the cis-gene set might be enriched in some pathways. Similarly for the trans-gene set, it might be helpful is there is an enriched pathway”

### Concern about small latent factors during normalization

- **Reviewer 3, comment 1** - concerned about latent factors that could have influenced our analysis (must not have understood our confounding variable selection and adjustment)

“As GTEx is bulk tissue RNA-sequencing, how do the authors account for smaller latent factors outside of PEER factors that could influence their mediation analysis? ”

## Concerns about analysis reproducibility

- **Reviewer 2, comment 7** - concerned that no code is provided for reproducing analyses

“A number of analyses are done, but no code or implementation is provided to reproduce them.”

## Comments I don't understand

- **Reviewer 3, comments 4** - unsure what is meant by this comment

“When there is trans-gene mediation is the trans-gene more likely to be correlated with other genes in the genome than trans-genes for cis-gene mediations? I would like to see evidence that these are not just correlated to some other covariance in the data.”

- **Reviewer 3, comments 5** - unsure what is meant by this comment

“Line 142: By using thresholds here for sharing, is the estimate of cell-type specificity influenced by winner's curse?”