Review of: "Transcriptomic similarity in the mouse and human brain"

Melissa Thalhammer[1], Christian Sorg[1]

1 Technische Universität München

Potential competing interests: The author(s) declared that no potential competing interests exist.

Summary of the article

Although mouse models are indispensible for neuroscientific research, preclinical studies often fail to be translated into a clinical setting since we lack full understanding of murine brain equivalents in the human brain. Thus, current research aims at the direct quantitative comparison using a common data space. The established approach to retrieve a common space exploiting connectivity is not applicable to mice since neuroanatomical correlates are missing. Another way to derive a common space has been to explot spatial gene expression patterns of homologous genes. However, direct comparisons of human and murine brain maps would require a higher degree of morphometric similarity. Hence, this work aimed at the generation of a common space from spatial gene expression data with increased resolution, which is achieved by using a supervised machine learning model.

Using the AMBA and AHBA as a proxy for murine and human regional gene expression patterns, respectively, the first asked whether regional gene expression of homologous genes is in principle a suitable measure to assess general similarities between the two species' brains. Even simple Pearson correlation to connect the two expression profiles resulted in a similarity matrix that corresponded to well known neuroanatomical regions in both species.

However, this simplified model lacks the ability to fully resolve functionally distinct regions based on their gene expression profiles due to its limited resolution. Hence, weighting genes based on their ability to discriminate different brain regions to create a latent common space is suggested as a more promising approach. This latent common space was retrieved by training a multi-layer perceptron classifier and was confirmed to increase resolution in two control experiments.

Applying the multi-layer peceptron classification approach to concrete scientific questions, the authors showed that their model can be used to answer questions related to the degree of similarity between the two species' brains. First, cortical sensorimotor regions are more similar between humans and mice than supramodal cortices, which is in line with prior literature. Second, several striatal areas were identified that seem to be highly conserved between both species.

In summary, this work provides a promising approach for comparing murine and human cortical similarity in a latent common space, which is based on regional gene expression of homologous genes.

Suggested improvements

The clustering is solely based on 2624 homologous genes of over 20,000 protein coding genes in humans. Thus, the whole assessment of genome similarity is based on a small proportion of all genes that are crucial in functioning of both organisms.

Regarding the results part, some minor parts are not entirely clear. In Figure 1A, depiction of regions (i.e., also in colors of which some seem to be rather similar to the correlation scales) is not optimal. The authors state that “similarity matrix exhibits broad patterns of positive correlation between the mouse and human brains”. First, I doubt that a maximum value of Pearson r = 0.404 mentioned for cerebellar nuclei cannot be considered “relatively high correlation”. Which range is considered as highly similar for mice and human? Furthermore, they do not explain what the (also big part) of negative correlations in Fig 1B means. A general issue of applying Pearson correlation to obtain the similarity matrix might be the spatial autocorrelation of brain maps (see Markello and Misic 2021), which might skew p-values.

Regarding Figure 5A and B, the claimed findings from the results sections are not clearly depicted in these figures. At first sight, the difference in average maximal correlation illustrated in Figure 5B does not seem statistically significant since the variance seems quite large.

When looking at Figure 5C, one might wonder why human clusters do not separate as clearly. Might training the model on mouse data be a reason? Might this even bias the model towards murine data? Also, how initial cluster size is defined could be discussed in more detail since this is crucial for the outcome of hierarchical clustering.

In Figure 6A, the bars’ colors are not distinguishable easily at first sight. Referring to the sentence: “We find that the variance in correlation calculated over all mouse targets is much lower ($\sigma = 0.04$) compared with the equivalent variances for the caudate ($\sigma = 0.09$) and putamen ($\sigma = 0.10$), indicating less specificity to any one mouse striatal target.”, for me, it remains puzzling why a lower amount of variance over all mouse targets indicates less specificity rather than the other way round (i.e., less variance hints to higher precision).

Again, claiming to offer “a first quantitative whole-brain comparison between the two species” goes too far in my opinion. It is indeed a whole-brain analysis, but only homologous genes are included, neglecting a large proportion of the two species’ genome.

Overall, the article provides detailed information on the applied methods and aims at the reproducibility of findings by providing the analysis code. Its clear structure and helpful illustrations guide the reader through the work.