Mouse Human Mapping Paper 1

Antoine Beauchamp

October 01, 2021

Abstract

Introduction

Animal models play an indispensable role in neuroscience research, both in the context of preclinical translation and in that of basic science. While numerous species have been used to model the human brain, the mouse has emerged as the most prominent of these, due to its rapid life cycle, straightforward husbandry, and amenability to genetic engineering **(References)**. Mouse models have proven to be extremely useful for understanding the diverse features of the brain, from its molecular neurobiological properties to its large-scale network properties **(References)**. While studying the mouse brain is a worthwhile endeavour in its own right, the primary purpose of using the mouse as a model system is to understand something about the human brain. Consequently it is crucial that neuroscientific findings discovered using mouse models be translatable to humans. In order to accomplish this translation, a correspondence must be established between the brains of the two species. Historically this has been done by designating certain pairs of neuroanatomical regions as homologous. Such homologues are defined on the basis of a variety of biological and physical properties, including cytoarchitecture, myeloarchitecture, connectivity, and morphology **(References) (Include mention of evolution and embryonic development)**. This approach has worked well to identify general correspondences between mouse and human brains, but it is not without limitations. In particular, defining neuroanatomical homologues in this way is an entirely qualitative exercise, in which atlas labels from one species are matched to labels in the other species. Additionally, such semantic matches are usually taken to be one-to-one, eliminating the possibility that a given area in one brain may map onto multiple areas in the other. Finally, many brain regions don't have an obvious homologue in the other species. In such instances, there is no way to assess the degree of dissimilarity between species. These limitations highlight the need to develop new ways of comparing brains across species **(Can probably make these sentences on limitations better)**.

Over the last decade, researchers in the field of comparative neuroscience have begun to explore ways to make more formal comparisons between the brains of different species. This has been made possible by advances in data acquisition technologies, particularly magnetic resonance imaging (MRI), as well as advances in computing power and data analytic methodology. Thus far, this line of research has primarily explored homologies between the brains of humans and non-human primates. In 2013, Mars et al. first used Passingham's notion of a connectivity fingerprint (Passingham, Stephan, and Kötter 2002) to identify the macaque homologue of the human temporoparietal junction (Mars et al. 2013). The connectivity fingerprint

is a signature that characterizes how a given region of interest is connected to a set of pre-specified target regions. These targets are chosen to be homologous between the species of interest, where the definition of homology is based on prior information. Since each region can be uniquely identified by its connectivity fingerprint, it is possible to match fingerprints across species to identify regional homologues. Mars et al. generated their connectivity fingerprints using resting-state functional MRI (rs-fMRI) data. By building these fingerprints for voxels in the human temporoparietal junction and matching them to the fingerprints of voxels in the macaque brain, they were able to determine which macaque region was most likely to be homologous. Since then it has been shown that connectivity fingerprints can be used to perform a number of different cross-species comparisons, including matching the fingerprints for a set of pre-specific regions of interest in one species to a template of the brain in the other species, and matching a connectivity fingerprint across a specific region of the brain (e.g. the cortex) in order to identify the closest cross-species match (Mars et al. 2016). In 2018, Mars et al. extended the idea of a connectivity fingerprint to a whole-brain connectivity blueprint, in which a connectivity fingerprint is generated for every region in the brain (Rogier B. Mars et al. 2018). In this case, rather than building the connectivity fingerprints using functional connectivity to pre-specified target regions assumed to be homologous, they used a given region's connection to white matter tracts common among all higher primates. While this line of research has been focused primarily on cross-species comparisons between humans and non-human primates, the development of novel methods has encouraged neuroscientists to think in new ways about exploring cross-species homologies. More recently, Balsters et al. applied the concept of connectivity fingerprints to assess the degree of homology in the striatum of mice, non-human primates and humans, finding that the nucleus accumbens was well conserved across species, but that the human striatum contained a large number of voxels with no obvious homologues in the mouse (Balsters et al. 2020).

Together these results highlight a framework for making formal comparisons between the brains of different species (Rogier B Mars, Passingham, and Jbabdi 2018). Rather than relying on the definition of homologous neuroanatomical pairs, the degree of similarity between voxels or regions in the brain is assessed directly using an intermediate common space. In order to serve as a bridge between the different species, the common space is constructed using quantitative maps of some underlying homologous biological feature. For instance, in the case of rs-fMRI data (Mars et al. 2013), the space is built using a seed region's functional connections to a set of target regions, which are assumed a priori to be homologous between the species. With white matter tractography data (Rogier B. Mars et al. 2018), the space is built using a region's connections to homologous white matter tracts. While cross-species comparisons can be accomplished using connectivity profiles, connectivity maps are by no means the only kind of data that can be used to define a common space. In particular, the availability of whole-brain spatial gene expression data sets provides an opportunity to build such a space using the expression patterns of homologous genes (Lein et al. 2007; Hawrylycz et al. 2012). Using such data sets, it has been shown that certain broadly defined regions in the mouse and human brains exhibit similarity on the basis of their gene expression profiles (Myers 2017).

Here we examine the similarities between mouse and human brains using a common space constructed from the regional expression of homologous genes. **(Say something about many to many mappings)**
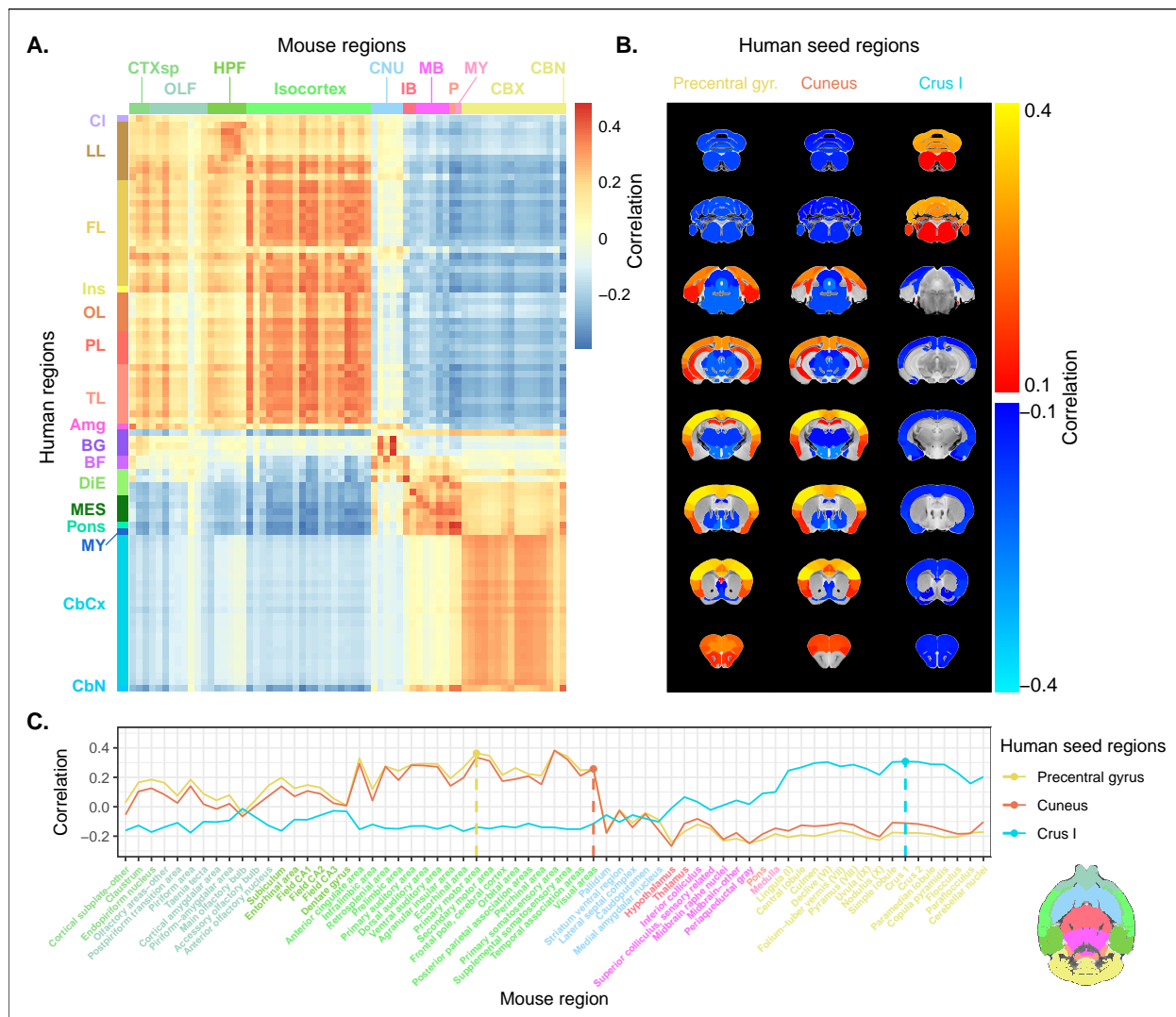
(Close off Introduction)

Results

Homologous genes capture broad similarities in the mouse and human brains

We first examined the pattern of homologies that emerged when comparing mouse and human brain regions on the basis of their gene expression profiles. We constructed a gene expression common space using widely available data sets from the Allen Institute for Brain Science: the Allen Mouse Brain Atlas (AMBA) (Lein et al. 2007) and the Allen Human Brain Atlas (AHBA) (Hawrylycz et al. 2012). These data sets provide whole-brain coverage of expression intensity for thousands of genes in the mouse and human genomes. For our purposes we filtered these gene sets to retain only mouse-human homologous genes. This subset was obtained using the NCBI HomoloGene system **(Reference)**. Prior to analysis, we ran both data sets through a preprocessing pipeline that included quality control checks, normalization procedures, and aggregation of the expression values under a set of atlas labels (see Materials and methods). The result was a gene expression blueprint in either species, describing the normalized expression of 2624 homologous genes across 67 mouse regions and 88 human regions **(How were these obtained?)**. We quantified the degree of similarity between all pairs of mouse and human regions in these blueprints using the Pearson correlation coefficient, resulting in a mouse-human similarity matrix (Figure 1, panel A).

As described previously by Myers (Myers 2017), we find that the similarity matrix exhibits broad patterns of positive correlation values between the mouse and human brains. These clusters of similarity correspond to coarse neuroanatomical regions that are generally well-defined in both species. For instance, we observe that the mouse isocortex is broadly similar to the human cerebral cortex, with the exception of the hippocampal formation, which forms a unique cluster. Similarly the mouse and human cerebellar hemispheres cluster together, while the cerebellar nuclei are set apart. These broad clusters of similarity are evident in the correlation matrix, but the ability to resolve regional matches on a finer scale is limited when using all homologous genes in this way. This is especially true for regions within the cortex and the cerebellum, which exhibit a high degree of internal homogeneity. For instance, the similarity profiles of the human precentral gyrus and cuneus both exhibit a plateau of similarity to the mouse isocortex. Within this plateau, the profiles of the two seeds are similar. The same is true of cerebellar cortical regions, such as the human crus 1, which is broadly similar to all regions in the mouse cerebellar cortex (Figure 1, panels B and C). Thus the regional expression patterns of mouse-human homologous genes can be used to identify general similarities between the brains of the two species, but the ability to identify finer scale matches is limited.

Results 2 (Placeholder title)

Using a common space constructed from the expression of 2624 homologous genes, we observed clusters of similarity between coarsely defined regions of the mouse and human brains. While the distribution of correlation values within these broad regions is relatively homogeneous, the data show hints of local variation among the finer regions that make up these clusters, suggesting that the set of homologous genes contains information about mouse-human matches on a finer scale (Figure 1, panel C). We proceeded to investigate whether we could enhance these finer regional matches by constructing a new set of variables from the original gene set. Our primary goal was to transform the initial gene space into a new common space that would improve the locality of mouse-human matches. However while we sought a transformation that would allow us to recapitulate known mouse-human neuroanatomical homologues, we also wanted to avoid directly

Figure 1: Mouse-human similarity in gene expression space. (A) Similarity matrix displaying the correlation between 67 mouse regions and 88 human regions based on the expression of 2624 homologous genes. Column annotations feature 11 coarsely defined mouse regions: Cortical subplate (CTXsp), olfactory areas (OLF), hippocampal formation (HPF), isocortex, cerebral nuclei (CNU), interbrain (IB), midbrain (MB), pons (P), medulla (MY), cerebellar cortex (CBX), cerebellar nuclei (CBN). Row annotations feature 16 coarsely defined human regions: Claustrum (Cl), limbic lobe (LL), frontal lobe (FL), insula (Ins), occipital lobe (OL), parietal lobe (PL), temporal lobe (TL), amygdala (Amg), basal ganglia (BG), basal forebrain (BF), diencephalon (DIE), mesencephalon (MES), pons, myelencephalon (MY), cerebellar cortex (CbCx), cerebellar nuclei (CbN). Broad patterns of similarity are evident between coarsely defined brain regions, while correlation patterns are mostly homogeneous within these regions. (B) Mouse brain slices showing similarity profiles for the human precentral gyrus, cuneus and crus I. Correlation patterns for the precentral gyrus and cuneus are highly similar to one another and broadly similar to most isocortical regions. The similarity profile for the crus I exhibits homogeneous similarity to the mouse cerebellum. (C) Anatomically-ordered line charts displaying the similarity profiles for the seed regions in (B). Dashed lines indicate the canonical mouse homologue for each human seed. Annotation colours correspond to atlas colours from the AMBA and AHBA for mouse and human regions respectively.

4

encoding such homologies in the transformation. Using this information as part of the optimization process for the transformation would run the risk of driving the transformation towards mouse-human matches that are already known. While we are interested in being able to recover such matches, we are equally interested in identifying novel and unexpected associations between neuroanatomical regions in the mouse and human brains, e.g. one-to-many correspondences. Given these criteria, our approach to identifying an appropriate transformation was to train a multi-layer perceptron (MLP) classifier on the data from the AMBA. The classifier was tasked with predicting the 67 labels in our mouse atlas from the voxel-wise expression of the 2624 homologous genes (Figure 2, panel A). While the model could have been trained using the data from either species, we chose to use the AMBA because it provides continuous coverage of the entire brain and is thus better suited to this purpose. In training the MLP to perform this classification task, we effectively optimize the network architecture to identify a transformation from the input gene space to a space that encodes information about the delineation between mouse brain regions. To extract this transformation, we removed the final predictive layer from the trained neural network. The resulting architecture defines a transformation from the input gene space to a lower-dimensional gene expression latent space. We then applied this transformation to the mouse and human gene expression blueprints to obtain representations of the data in the latent space (Figure 2, panel B). Finally, we used these gene expression latent space blueprints to compute the new similarity matrix (Figure 2, panel C). Since the algorithm used to train the MLP features an inherent degree of stochasticity, we repeated this training and transformation process 500 times to generate a distribution of latent spaces and similarity matrices over training runs.

To assess whether the latent space representations of the data improved the resolution of the mouse-human matches, we considered two criteria. The first was whether the similarity profiles of the mouse atlas regions were more localized within the corresponding broad regions of interest, compared with their similarity profiles in the original gene space. The second criterion was whether the degree of similarity between canonical neuroanatomical homologues improved in this new common space. The former tells us about our ability to extract finer-scale signal in these profiles, while the latter informs us about our ability to recover expected matches in this finer-scale signal. To evaluate these criteria, we computed ranked similarity profiles for every region in the mouse brain, ordered such that a rank of 1 indicates the most similar human region. In addition, given the difference in overall correlation values between the input gene space and gene expression latent space similarity matrices, we scaled the similarity profiles to the interval $[0, 1]$ in order to make comparisons between the spaces.

We evaluated the locality criterion by examining the decay rate of the head of the similarity profiles. We reasoned that the broad plateau of similarity to a coarse brain region, as seen in the anatomically-ordered similarity matrices and profiles (Figure 1, panels A and C; Figure 2, panel C), would correspond to a similar plateau at the head of the rank-ordered profiles. Moreover, the emergence of local signal would appear as an increase in the range between the peaks and troughs within the broad plateau. In the rank-ordered profiles, this would correspond to a faster rate of decay at the head of the profile. In order to quantify this decay for a given similarity profile, we computed the rank at which the profile fell to a scaled similarity value of 0.75. This was calculated for the similarity profiles of each mouse region in the initial gene space, as well as each of the 500 gene expression latent spaces. As a measurement of performance between the two representations of the data, we then took the difference in this rank between each of the latent spaces and the original gene space (Figure 3, panel A). This was done such that a negative rank difference indicated an improvement in the latent space.
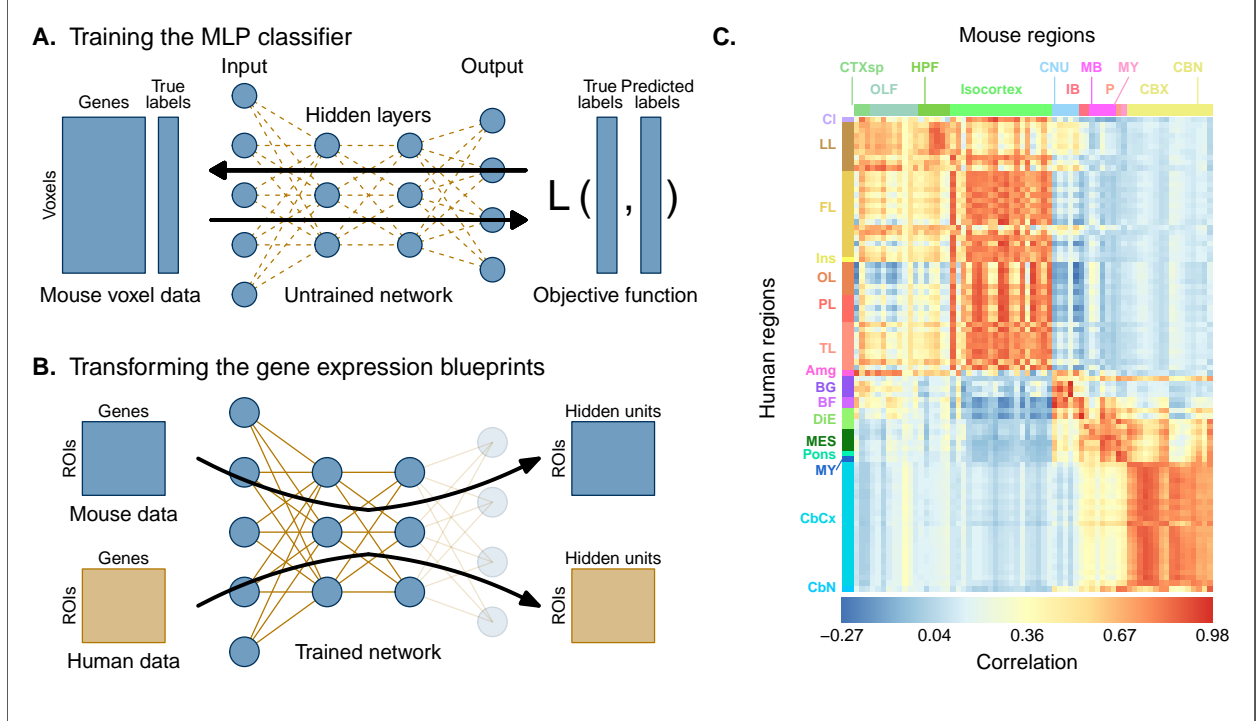
Figure 2: Mouse-human similarity in MLP gene expression latent space. (A) Schematic illustrating how the MLP classifier was trained. Voxel-wise expression maps from 2624 homologous genes in the AMBA were used to train the neural network to classify each mouse voxel into one of 67 atlas regions. Blue circles represent hidden units or neurons in the network while golden dashed lines represent the connections or weights between units in different layers. The network generates predicted labels from the input data (black arrow pointing forwards), which are then passed to an objective function to evaluate the network's performance. This information is propagated backwards through the network to improve the weights (black arrow pointing backwards). This process is repeated until an optimal solution is found. (B) Schematic illustrating how the MLP was used to create a new common space. Once the network was trained, the output layer was removed from the architecture. The mouse and human regional gene expression matrices were passed through the network (black arrows), resulting in matrices with columns corresponding to the hidden units in the final hidden layer of the MLP. These hidden units form the axes of the new gene expression latent space. The training and transformation process was repeated 500 times. (C) Similarity matrix displaying the gene expression latent space correlation between 67 mouse regions and 88 human regions averaged over 500 MLP training runs. Similar brains exhibit very high correlation values.

Examining the structure-wise distributions of these rank differences, we found that for the majority of regions in our mouse atlas, the MLP classification approach resulted in either an improvement in the amount of locality within a broad region, or no difference from the original gene space (Figure 3, panel B). Only a few regions performed considerably worse in the latent spaces, notably the main and accessory olfactory bulbs, as well as the cerebellar nuclei. Regions within the cortical subplate and olfactory areas (e.g. endopiriform nucleus, postpiriform transition area) gained the largest improvement from the classification approach. While the effects are smaller, the similarity profiles of regions belonging to the isocortex and cerebellar cortex also saw an improvement in locality. In contrast, regions belonging to the cerebral nuclei, the diencephalon, midbrain and hindbrain did not see much improvement in this new common space. For many such regions the degree of locality appears to be worse in this space, though only be a small number of ranks (e.g. striatum ventral region, thalamus, midbrain raphe nuclei). To evaluate how often the MLP training and transformation process returned an improvement, we computed the proportion of training runs over which the rank difference was negative or null for every mouse region (Figure 3, panel C). This elucidates the patterns seen in Figure 3, panel B, where cerebral and cerebellar regions exhibit a high proportion of runs resulting in improvement, while many subcortical regions are not improved in most training samples. Particularly, we find very low proportions for the thalamus, dentate gyrus, and striatum ventral region.

The decay rates of the similarity profiles indicated that the MLP classifier approach improved our ability to identify finer scale matches for a number of brain regions. However, this does not mean that those improved matches are biologically meaningful. Our second criterion for evaluating the performance of the MLP addresses whether this improvement in locality captures what we would expect in terms of known mouse-human homologies. To this end, we examined the degree of similarity between canonical mouse-human matches, both in the initial gene expression space and in the set of latent spaces. We began by establishing a list of 37 canonical mouse-human homolous pairs. For each of these regions in the mouse brain, we examined how the rank of the canonical human match changed in the rank-ordered similarity profiles between the latent spaces and the original gene space (Figure 4, panel A). The lower the rank, the more similar the canonical pair, with a rank of 1 indicating maximal similarity. We additionally computed the difference between the rank of the expected human match in the latent space and in the original common space. This was done for each of the 500 latent spaces. For each region in the mouse brain with a canonical human homologue, we used these differences to compute the proportion of latent spaces that returned an improvement or null difference when compared with the original gene space (Figure 4, panel B). We find that for most regions in the mouse brain, the MLP classification approach either improves the seed region's similarity to its expected match or performs as well as the comparison using all homologous genes. The improvement is most pronounced for regions in the cortical subplate and isocortex, particularly the frontal pole and visual areas. The gene expression latent spaces do not provide an improvement for many regions in the sub-cortex, since the initial common space was already recapitulating the appropriate homologues with maximal similarity. In such cases, the classification approach performs as well as the original approach. While many regions in the cerebellum feature some improvement in the latent spaces, the variation in the rank of the human homologue is often quite large, indicating some instability in the neural network's ability to recover these matches. By examining the identity of the top ranked human match for a given cerebellar seed region (e.g. declive (VI)) across different latent space samples, we find that the top match is always a cerebellar region (Figure 4, panel C). This suggests that the variability observed in the cerebellar regions corresponds to a shuffling in the positions of the cerebellar targets on the human side. The same is true for hippocampal regions. While the canonical human match is not always the top match, the highest ranked
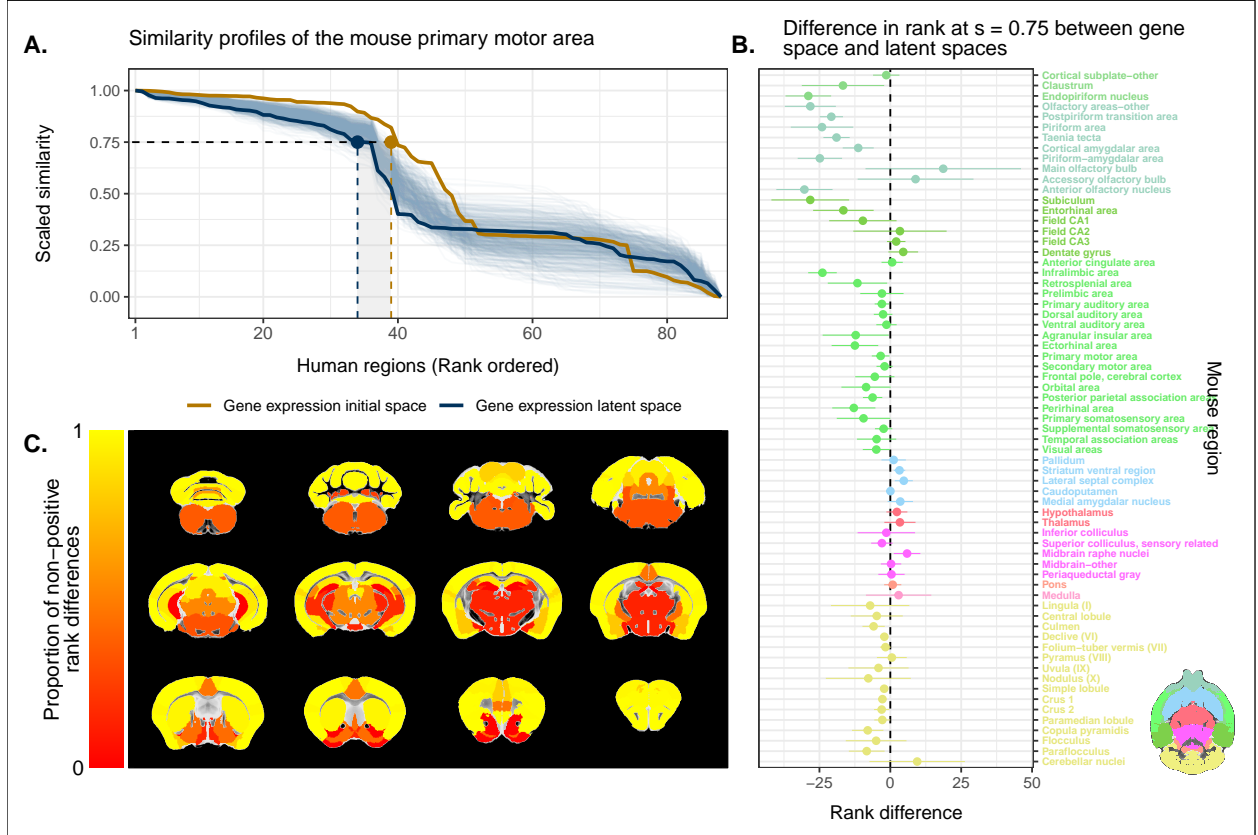
Figure 3: Quantifying improvement in locality in gene expression latent space. (A) The amount of local signal within a broadly similar region of the brain for a finer seed region's (e.g. primary motor area) similarity profile can be quantified by the decay rate of the head of the rank-ordered profile. Decay rate was quantified by computing the rank at a similarity of 0.75. This metric was compared between the initial gene expression space (orange line) and every gene expression latent space resulting from repeated training of the MLP (every blue line is a training outcome, heavy blue line serves as an example). A negative difference between these rank metrics indicates an improvement in locality in the latent space. (B) Structure-wise distributions of differences in rank at a similarity of 0.75 between the initial gene expression space and the gene expression latent spaces. Points and error bars represent mean and 95% confidence interval. Dashed black line at 0 indicates the threshold for improvement in one space over the other. Colours correspond to AMBA annotations as in Figures 1 and 2. (C) Proportion of MLP training runs resulting in an improvement or null difference in the gene expression latent space compared with the initial space. Cortical and cerebellar regions exhibit high proportions of improvement, while subcortical regions are less likely to be improved by the classification process.

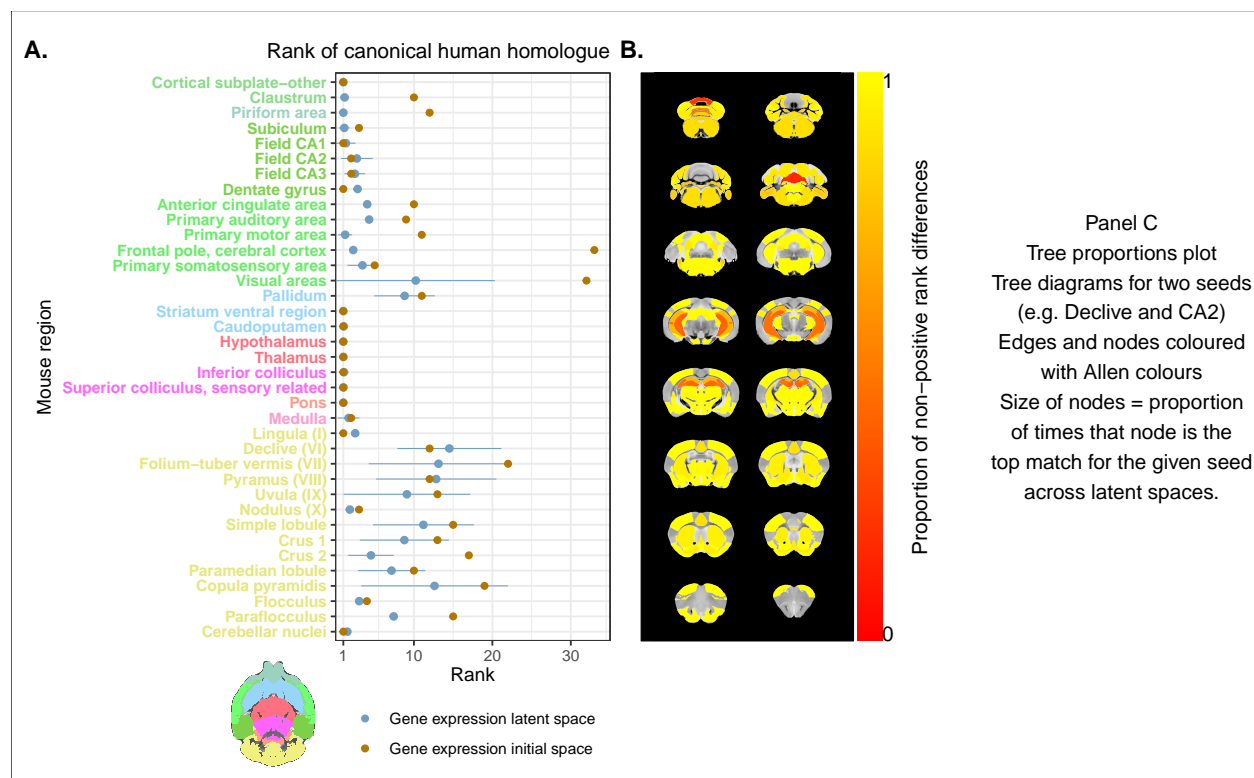region always belongs to the hippocampal formation.

Figure 4: Recovering canonical homologues in gene expression space. (A) Comparing the ranks of canonical human matches for mouse seed regions between the initial gene expression space and gene expression latent spaces. Points and error bars represent mean and 95% confidence interval. Axis colours correspond to AMBA annotations. (B) Proportion of latent spaces resulting in an improvement or null difference compared with the initial gene space. Uncoloured areas correspond to regions with no establish canonical human match. (C) Proportion of latent spaces returning each tree node as the top match for the given seed.

These results demonstrate that the MLP classification approach improves our ability to resolve finer scale mouse-human neuroanatomical matches within the broadly similar regions obtained using the initial gene expression space. By training a classifier to predict the atlas labels in one species, we were able to generate a new common space that amplified the amount of local signal within broadly similar regions while also improving our ability to recover known neuroanatomical homologues.

Sensorimotor areas of the mouse isocortex are more strongly conserved than association areas

Having validated our ability to resolve finer scale matches using the gene expression latent space, we proceeded to investigate hypotheses about conservation of brain areas between the mouse and human. We hypothesized that sensorimotor areas of the cortex would be more conserved than association areas **(Need some literature justification and contextualization here)**. We quantified the degree of conservation for a given cortical seed region in one species as the maximal correlation value across all target regions in the other species. We evaluated this quantity for every region in the mouse isocortex in each of the 500 gene expression latent spaces (Figure 5, panel A). We further probed the association between cortex type and conservation by computing the average maximal correlation per cortex type in each of the gene expression latent spaces (Figure 5, panel B). On average, we find that mouse sensorimotor cortical areas exhibit a greater

degree of conservation than association areas, with the primary motor and somatosensory areas being the most conserved. Despite the stochasticity that emerges due to the training process of the MLP classifier, mouse sensorimotor cortical areas consistently exhibit larger maximal correlation values than association areas.
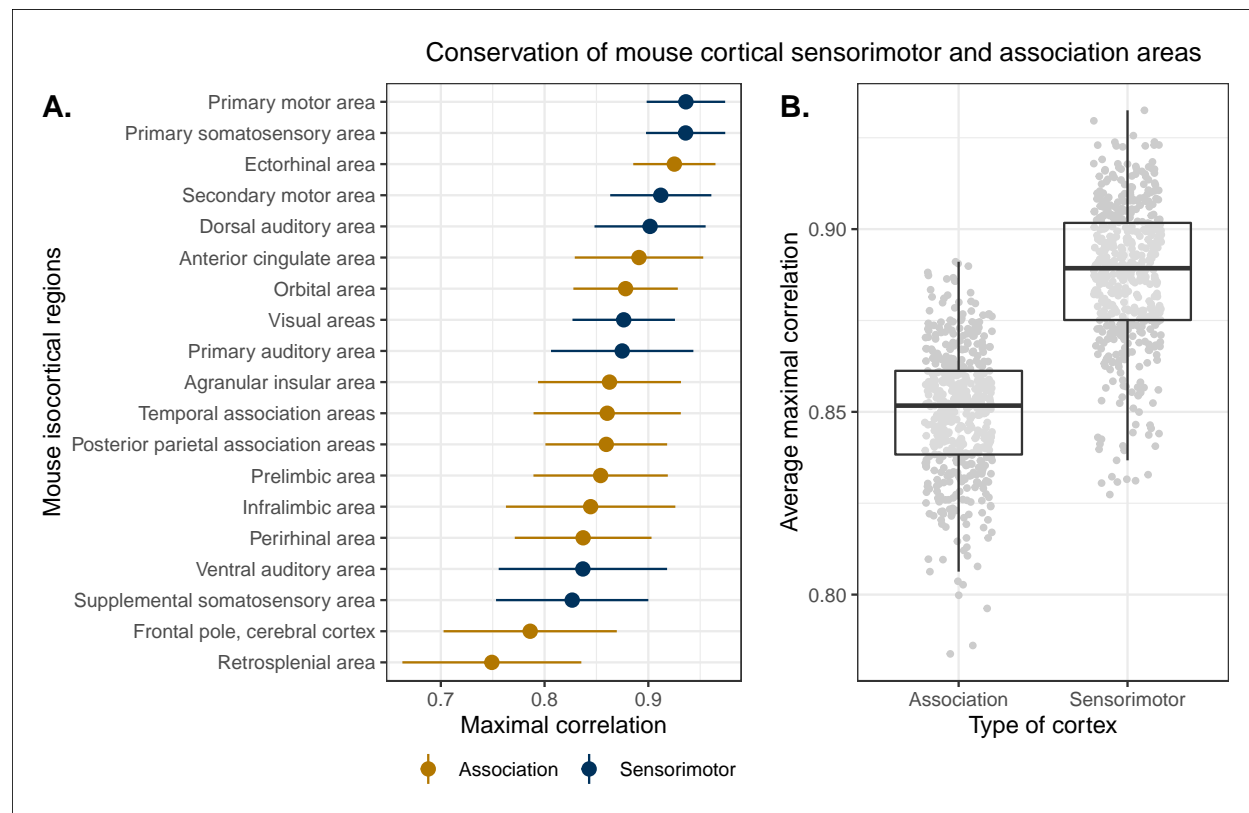

Figure 5: Conservation of mouse cortical areas. The gene expression latent space can be used to evaluate evolutionary hypotheses such as whether sensorimotor or association cortical areas are more conserved between the mouse and human. (A) Conservation was evaluated using the maximal correlation value of seed region. Points and error bars represent mean and 95% confidence interval over latent space samples. Sensorimotor areas tend to have higher maximal correlation values on average. (B) Average maximal similarity per cortex type in each gene expression latent space.

Discussion

(To be filled in. Need some sense about how to contextualize these results before I can write this.)

Materials and methods

Mouse gene expression data

For the mouse expression data, we used the whole brain in-situ hybridization (ISH) coronal data set from the Allen Mouse Brain Atlas (Lein et al. 2007). **(Sentence or two about how ISH MINC files were generated from the raw AMBA data)**. The ISH images for 4346 genes were imported and masked to

form a gene-by-voxel expression matrix. We pre-processed this data by first applying a `log2` transformation for consistency with the human data set. Subsequently we filtered out genes for which more than 20% of voxels contained missing values. For those genes with replicated ISH experiments, we averaged expression of each voxel across the experiments. Finally we applied a K-nearest neighbours algorithm to impute the remaining missing values. This was done using the genes as the variables. The result was a gene-by-voxel expression matrix with 3958 genes and 61316 voxels.

Human gene expression data

Human gene expression data was obtained from the Allen Human Brain Atlas (Hawrylycz et al. 2012). We used the microarray data from the brains of all 6 donors, each of which contains `log2` expression values for 58692 gene probes across numerous tissue samples. The data sets were fed through a pre-processing pipeline built following the recommendations from Arnatkeviciūtė et al. (2019) (Arnatkeviciūtė, Fulcher, and Fornito 2019). Specifically, once imported, we passed the data from each donor through a set of filters. The first filter removed gene probes that were not associated with an existing Entrez gene ID. The second filtering step used the probe intensity filter provided by the AHBA. For each donor, we only retained the probes for which more than 50% of samples passed the intensity filter. After filtering, we aggregated the expression values for probes that corresponded to the same gene. To do so, we computed the average expression per sample for probes corresponding to a given gene. This was done separately for each donor, and the averages were computed in linear space rather than `log2` space. Once the average gene expression values were obtained, we transformed the data back to `log2` space. Finally, we combined the gene-by-sample expression matrices across the different donors. In doing so, we retained only those genes present in the data sets from all 6 donors. The result was a gene-by-sample expression matrix with 15125 genes and 3703 samples.

Mouse atlases

We used a modified version of the DSURQE atlas (Dorr et al. 2008; Richards et al. 2011; Ullmann et al. 2013; Steadman et al. 2014; Qiu et al. 2018) at the Mouse Imaging Centre. The labels of the atlas are mapped to the finest regions in the AMBA ontology, resulting in a hierarchical neuranatomical tree. We used this tree to prune the hierarchy to the desired level of granularity. White matter and ventricular segmentations were removed entirely. The remaining grey matter regions were pruned so that the majority of leaf nodes contained enough voxels to be classified appropriately by the multi-layer perceptron. In doing so, we also attempted to maintain the same level of tree depth within a broad region, e.g. cerebellar regions were chosen at the same level of granularity. The leaf nodes of this tree were used as the labels in our custom atlas, resulting in a mouse atlas with 67 grey matter regions. We additionally generated an atlas with 11 broader regions for visualization and annotation purposes.

Human atlases

For our human atlas we used the ontology from the AHBA. We pruned the neuroanatomical hierarchy to correspond roughly to the level of granularity in our mouse atlas, resulting in 88 human brain regions. We additionally generated a set of 16 broad regions for visualization and annotation. White matter and ventricular regions for ommitted entirely.

Creating similarity matrices

We created the mouse and human gene expression blueprints from the human gene-by-sample expression matrix and the mouse gene-by-voxel expression matrix. We first intersected the gene sets in these matrices using a list of homologous genes obtained from the NCBI HomoloGene database. The resulting set contained 2624 homologous genes for both the mouse and human data. We then annotated each of the human samples with one of the 88 human regions in our atlas. The same was done for the mouse voxels with the 67 mouse atlas regions. Voxels and samples corresponding to white matter and ventricular regions were discarded in this process. These labelled expression matrices were subsequently normalized. For each matrix, we first standardized every gene across all voxels or samples. We then centered every voxel or sample across all genes. Finally, we generated the expression blueprints by averaging the expression of every homologous gene over the voxels or samples that corresponded to each atlas region. Using these blueprints, we computed the mouse-human similarity matrix by computing the Pearson correlation coefficient between all pairs of mouse and human regions.

Validating the multi-layer perceptron

Prior to training the multi-layer perceptron classifier, we applied a validation procedure to tune the hyperparameters of the network. We chose a model architecture in which each layer of the network was fully connected to the previous and subsequent layers. However we still had flexibility in setting a number of hyperparameters, namely the number of hidden layers in the network, the number of hidden units (i.e. neurons) per hidden layer, the dropout rate, and the amount of weight decay. Given that the majority of genes in the coronal AMBA data set are only sampled once over the entire mouse brain, we devised an ad hoc validation procedure to estimate the out-of-sample error of our model as we varied these hyperparameters. The process involved a combination of the coronal AMBA data set and the sagittal AMBA data set. Since the sagittal data set features more genes than the coronal data set, we first filtered the list of genes for those present in the coronal set. We then imported and pre-processed the sagittal data using the process described above for the coronal data set. Since the sagittal data set only covers one hemisphere of the mouse brain, we also repeated the importation and pre-processing process for the coronal data, but masking it using the sagittal brain mask so that only one hemisphere was retained. In applying this pipeline, we did not aggregate the expression of multiple ISH experiments for those genes featuring more than one experiment. Thus some of the genes in this coronal set had more than one sample. Once these gene-by-voxel expression matrices were built, we filtered each of them according to the list of mouse-human homologous genes and the human sample expression matrix. We then annotated the voxels in each of the mouse voxel expression matrices with one of the 67 regions in our atlas. In order to validate the MLP, we had to generate both a training set and a validation set. To do so, we applied the following procedure: For every gene in the homologous set, we first determined whether the gene had more than one sample in the coronal ISH data set. If this was the case, we randomly chose one of those samples for the training set and one of the remaining samples for the validation set. However if the gene only had a single sample in the coronal data set, we randomly chose either the coronal sample or the sagittal sample for the training set and the other for the validation set. Once the training and validation sets were generated, we trained the MLP on the training set and evaluated its performance on the validation set. Both the training and validation sets were normalized as described above prior to training. We repeated this entire procedure 10 times for every hyperparameter combination.

We varied the number of hidden layers between 3 and 5. We varied the number of hidden units between 100 and 1000, increasing them in increments of 100. For the dropout rate, we examined values of 0, 0.25 and 0.50. For the amount of weight decay, we examined values of 10^{-6} and 10^{-5}. Following this validation procedure, we found that the best performing model had 3 hidden layers, 200 neurons per layer, a dropout rate of 0, and a weight decay value of 10^{-6}. This model had an average accuracy of 0.926 on the training sets and of 0.526 on the validation sets.

Training the multi-layer perceptron

Following validation, we trained the MLP classifier on the full coronal ISH data set from the AMBA. The voxel-wise expression matrices were labelled and normalized as described previously. The hyperparameters of the network were set to the optimal values obtained from the validation procedure. The network was trained using a learning rate of 10^{-5} over 200 epochs. **(Confirm what the loss function was for this and the validation procedure. Maybe also mention activation function, optimizer, etc.)**.

Creating the gene expression latent space

The trained MLP was used to generate the new gene expression latent space. To do so, we removed the predictive output layer from the network architecture, thus allowing the hidden units of the third hidden layer to be the output of the network. We then applied this modified network to the mouse and human gene expression blueprints, resulting in matrices describing each brain region's weight over 200 hidden units. These matrices form the blueprints in the latent space.

References

Arnatkevičiūtė, Aurina, Ben D. Fulcher, and Alex Fornito. 2019. "A practical guide to linking brain-wide gene expression and neuroimaging data." *NeuroImage* 189 (January): 353–67. https://doi.org/10.1016/j.neuroimage.2019.01.011.

Balsters, Joshua Henk, Valerio Zerbi, Jerome Sallet, Nicole Wenderoth, and Rogier B. Mars. 2020. "Primate homologs of mouse cortico-striatal circuits." *eLife* 9: 1–24. https://doi.org/10.7554/eLife.53680.

Dorr, A E, J P Lerch, S Spring, N Kabani, and R M Henkelman. 2008. "High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice." *NeuroImage* 42: 60–69. https://doi.org/10.1016/j.neuroimage.2008.03.037.

Hawrylycz, Michael J, Ed S Lein, Angela L Guillozet-Bongaarts, Elaine H Shen, Lydia Ng, Jeremy A Miller, Louie N Van De Lagemaat, et al. 2012. "An anatomically comprehensive atlas of the adult human brain transcriptome." *Nature* 489: 391–99. https://doi.org/10.1038/nature11405.

Lein, Ed S., Michael J. Hawrylycz, Nancy Ao, Mikael Ayres, Amy Bensinger, Amy Bernard, Andrew F. Boe, et al. 2007. "Genome-wide atlas of gene expression in the adult mouse brain." *Nature* 445 (7124). Nature Publishing Group: 168–76. https://doi.org/10.1038/nature05453.

Mars, R. B., J. Sallet, F.-X. Neubert, and M. F. S. Rushworth. 2013. "Connectivity profiles reveal the relationship between brain areas for social cognition in human and monkey temporoparietal cortex." *Proceedings of the National Academy of Sciences* 110 (26): 10806–11. https://doi.org/10.1073/pnas.1302956110.

Mars, Rogier B, Richard E Passingham, and Saad Jbabdi. 2018. "Connectivity Fingerprints: From Areal Descriptions to Abstract Spaces." *Trends in Cognitive Sciences* 22 (11): 1026–37. https://doi.org/10.1016/j.tics.2018.08.009.

Mars, Rogier B., Stamatios N. Sotiropoulos, Richard E. Passingham, Jerome Sallet, Lennart Verhagen, Alexandre A. Khrapitchev, Nicola Sibson, and Saad Jbabdi. 2018. "Whole brain comparative anatomy using connectivity blueprints." *eLife* 7: 1–15. https://doi.org/10.7554/eLife.35237.

Mars, Rogier B., Lennart Verhagen, Thomas E. Gladwin, Franz Xaver Neubert, Jerome Sallet, and Matthew F S Rushworth. 2016. "Comparing brains by matching connectivity profiles." *Neuroscience and Biobehavioral Reviews* 60. Elsevier Ltd: 90–97. https://doi.org/10.1016/j.neubiorev.2015.10.008.

Myers, Emma M. 2017. "Molecular Neuroanatomy: Mouse-Human Homologies and the Landscape of Genes Implicated in Language Disorders." PhD thesis, Boston University.

Passingham, Richard E., Klaas E. Stephan, and Rolf Kötter. 2002. "The anatomical basis of functional localization in the cortex." *Nature Reviews Neuroscience* 3 (8). Nature Publishing Group: 606–16. https://doi.org/10.1038/nrn893.

Qiu, Lily R., Darren J. Fernandes, Kamila U. Szulc-Lerch, Jun Dazai, Brian J. Nieman, Daniel H. Turnbull, Jane A. Foster, Mark R. Palmert, and Jason P. Lerch. 2018. "Mouse MRI shows brain areas relatively larger in males emerge before those larger in females." *Nature Communications* 9 (1). Nature Publishing Group: 2615. https://doi.org/10.1038/s41467-018-04921-2.

Richards, Kay, Charles Watson, Rachel F Buckley, Nyoman D Kurniawan, Zhengyi Yang, Marianne D Keller, Richard Beare, et al. 2011. "Segmentation of the mouse hippocampal formation in magnetic resonance images." *NeuroImage* 58: 732–40. https://doi.org/10.1016/j.neuroimage.2011.06.025.

Steadman, Patrick E, Jacob Ellegood, Kamila U Szulc, Daniel H Turnbull, Alexandra L Joyner, R Mark Henkelman, and Jason P Lerch. 2014. "Genetic Effects on Cerebellar Structure Across Mouse Models of Autism Using a Magnetic Resonance Imaging Atlas HHS Public Access." *Autism Research* 7 (1): 124–37. https://doi.org/10.1002/aur.1344.

Ullmann, Jeremy F P, Charles Watson, Andrew L Janke, Nyoman D Kurniawan, and David C Reutens. 2013. "A segmentation protocol and MRI atlas of the C57BL/6J mouse neocortex." *NeuroImage* 78. Watson; Paxinos: 196–203. https://doi.org/10.1016/j.neuroimage.2013.04.008.