

Standardizing Human Brain Parcellations

Patrick E. Myers¹
Allison D. Lemmer¹

Ganesh C. Arvapalli¹
Eric W. Bridgeford^{1,2}

Sandhya C. Ramachandran¹
Joshua T. Vogelstein^{*1}

Paige F. Frank¹

Brain atlases are key for localizing neurobiological regions of interest. Using volumetric coordinate spaces, functional relationships can be assessed in conjunction with anatomical data. Over the past fifty years, many human brain atlases have been developed using a variety of methods. However, these atlases are stored in different formats, orientations, and coordinate spaces, making comparisons across atlases and studies difficult. We consolidate all the popular human brain atlases into a single location, and transform them all into the same standard format. This format served as the basis for a specification that we introduce to store future atlases. To demonstrate the utility of collecting all these atlases in a common specification, we conduct an experiment using the Healthy Brain Network data, quantifying the dependence between each parcel in each atlas with various phenotypic variables. The repository containing the atlases, the specification, as well as relevant transformation functions is available at <https://github.com/neurodata/neuroparc>.

Background & Summary

Brain atlases are systems of labelling the brain with recognizable regions that provide some structural or functional parcellation [1]. Atlases serve to help the user visualize and understand the brain's complex organization. Dividing a brain into determined regions can be a starting point for quantitative analysis. Differences between human brains can be quantified in terms of different sizes of regions. Diseases can be better understood by studying population differences in affected brain regions. Atlases enable researchers to localize activity in the brain to a shared region between different brains [2]. Atlases can be used as a reference for surgery [3]. Moreover, atlases provide a coordinate system that enables scientists and doctors to describe relation positions across brains, and compare across studies [1, 4].

Numerous atlases have been defined over the past fifty years, often with different reference brains and parcellation methods [5]. Different data modalities and scales have been used to parcellate the brain into regions, including (i) macroscale anatomy [6], (ii) functional activity [7], and (iii) cytoarchitecture [8]. To date, over 20 different adult human brain atlases are common in the literature. In addition to using different source data for parcellating the brain, they also use different reference images, scales, orientations, and formats to store these data.

Some efforts to consolidate these atlases is already underway. For example, Nilearn is a popular package for machine learning relating to neuroinformatics and neuroimaging [9]. The ease of use of Nilearn is excellent since it provides several single line command line interface functions to “fetch” both atlases and datasets. The problem comes with the lack of standardization. Nilearn's functions pull from the source without modification and the documentation of the atlases and datasets usually just provides links to the original publication. Thus, for each atlas that an investigator wants to use, they must figure out the atlas-specific specification and details. Moreover, if the investigator desires to compare across atlases, they must understand the metadata associated with both atlases, and be able to transform them into a common format.

We present Neuroparc, which consists of (1) a repository of the most commonly used atlases in neuroimaging, (2) an atlas specification that will enable researchers to both easily understand existing atlases, and generate new atlases compliant with this specification, and (3) a set of functions for transforming, comparing, and visualizing these atlases. We compile and visualize 24 different adult human brain atlases, and compare them using the Dice score between all pairs of atlases. To validate the atlases and their utility, we conduct an analysis using the Healthy Brain Network (HBN) data [10]. Specifically, for each parcel of each atlas, we evaluate the dependence between activity and various phenotypic variables, to identify which atlases are particularly informative about phenotype. To facilitate replication and extension of this work, all the data and code are available from .

UPDATED* Background & Summary

Developing a thorough understanding the brain's organization is one of the key challenges in human neuroscience (Glasser 2016) and critical for clinical translation (Fox Greicius 2010). Parcellation of the brain into functionally and structurally homogenous regions has seen impressive advances in recent years (Eickhoff, Yeo, Genon, 2018), and has facilitated the growth of the field of network neuroscience (Hagmann 2007, Zalesky 2010). Through a range of efforts using clustering (Bellec, Rosa-neto, Lyttelton, Benali, Evans, 2010; Craddock, James, Holtzheimer, Hu, Mayberg, 2012; Nikolaidis 2019; Thirion, Varoquaux, Dohmatob, Poline, 2014), multivariate decomposition (Beckmann, Mackay, Filippini, Smith, 2009; Varoquaux, Gramfort, Pedregosa, Michel, Thirion, 2011), gradient based connectivity (Cohen et al., 2008; Glasser et al., 2016; Gordon et al., 2014; Nelson et al., 2010; Wig et al., 2014; Xu et al., 2016), and multimodal neuroimaging (Glasser 2016), researchers have led to new fundamental insights into the organization of the brain and it's network properties (Margulies 2016). The dissemination of these brain parcellations has enabled researchers to investigate brain network associations with developmental (Dosenbach et al., 2010; Liem et al., 2017), cognitive (Finn et al., 2015; Shehzad et al., 2014), and clinical phenotypes (Abraham et al., 2017; Greene et al., 2016; Hong, Valk, Di Martino, Milham, Bernhardt, 2018) .

Today, researchers interested in understanding brain organization are presented with a range of possible brain atlases for performing network based analyses (Arslan 2018). While this range of options is a boon to researchers, the use of different parcellations across studies makes assessing reproducibility of brain-behavior relationships difficult (e.g. comparing across parcellations with different organizations and numbers of nodes; Zalesky et al., 2010). Furthermore, replicating results across a range of parcellations can be a challenge given that they lack a common data format, orientation, and availability. Creating a database that offers researchers access to a large range of atlases in a common format would facilitate replication across atlases and studies.

Some efforts to consolidate these atlases is already underway. For example, Nilearn is a popular package for machine learning relating to neuroinformatics and neuroimaging [9]. The ease of use of Nilearn is excellent since it provides several single line command line interface functions to 'fetch' both atlases and datasets. Nilearn includes a twelve of volumetric anatomically and functionally defined atlases, such as the Harvard-Oxford and AAL atlases. However, these atlases all belong to a limited range of currently available atlases, and the more recent gradient based, surface based, and multimodal parcellations have not yet been included into a central repository with the most common volumetric atlases.

In the current work we present a common platform for compile and visualize 24 different adult human brain atlases within an easily accessible python package, Neuroparc. We create a platform for common specification of atlases across a range of surface and volume based parcellations enabling easy comparison of brain-behavior relationships across a range of parcellations. Neuroparc provides detailed atlas specification requirements to enable new atlases to be added easily and a set of tools for transforming, comparing, and visualizing these atlases. Here, to provide users with an overview of the relationship between these parcellations we assess the spatial similarity between atlases, as measured by Dice coefficient. Furthermore, we provide an prototypical analysis of brain-behavior relationships in these atlases defined networks across a range of phenotypic variables from the large publicly available developmental sample, the Healthy Brain Network. All code used to create these analyses, as well as the Neuroparc package and documentation are all publicly available for easy extension and integration into currently existing frameworks.

UPDATED* REFS IN ORDER

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, 536 (7615), 171–178.

Fox, M. D., Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in Systems Neuroscience*, 4, 19.

Eickhoff, S. B., Yeo, B. T. T., Genon, S. (2018). Imaging-based parcellations of the human brain. *Nature Reviews Neuroscience*, 19(11), 672–686.

Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R., Thiran, J.-P. (2007). Mapping human

whole-brain structural networks with diffusion MRI. *PloS One* , 2 (7), e597.

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yülcel, M., Pantelis, C., Bullmore, E. T. (2010). Whole-brain anatomical networks: does the choice of nodes matter? *NeuroImage* , 50 (3), 970â983.

Bellec, P., Rosa-neto, P., Lyttelton, O. C., Benali, H., Evans, A. C. (2010). *NeuroImage* Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *NeuroImage* , 51 (3), 1126â1139.

Craddock, R. C., James, G. A., Holtzheimer, P. E., Hu, X. P., Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8), 1914â1928.

Nikolaïdis, A., Heinsfeld, A. S., Xu, T., Bellec, P., Vogelstein, J., Milham, M. (2019). Bagging Improves Reproducibility of Functional Parcellation (p. 343392). [https://doi.org/ 10.1101/343392](https://doi.org/10.1101/343392)

Thirion, B., Varoquaux, G., Dohmatob, E., Poline, J. B. (2014). Which fMRI clustering gives good brain parcellations? *Frontiers in Neuroscience* , 8 (8 JUL), 1â13.

Beckmann, C. F., Mackay, C. E., Filippini, N., Smith, S. M. (2009). Group comparison of resting-state FMRI data using multi-subject ICA and dual regression . 181.

Varoquaux, G., Gramfort, A., Pedregosa, F., Michel, V., Thirion, B. (2011). Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. *Information Processing in Medical Imaging: Proceedings of the ... Conference* , 22 , 562â573.

Cohen, A. L., Fair, D. A., Dosenbach, N. U. F., Miezin, F. M., Dierker, D., Van Essen, D. C., ... Petersen, S. E. (2008). Defining functional areas in individual human brains using resting functional connectivity MRI. *NeuroImage* , 41 (1), 45â57.

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., Petersen, S. E. (2014). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral Cortex* , 26(1), 288â303.

Nelson, S. M., Cohen, A. L., Power, J. D., Wig, G. S., Miezin, F. M., Wheeler, M. E., ... Petersen, S. E. (2010). A parcellation scheme for human left lateral parietal cortex. *Neuron*, 67 (1), 156â170.

Wig, G. S., Laumann, T. O., Cohen, A. L., Power, J. D., Nelson, S. M., Glasser, M. F., ... Petersen, S. E. (2014). Parcellating an individual subject's cortical and subcortical brain structures using snowball sampling of resting-state correlations. *Cerebral Cortex* , 24(8), 2036â2054.

Xu, T., Opitz, A., Craddock, R. C., Wright, M. J., Zuo, X.-N., Milham, M. P. (2016). Assessing Variations in Areal Organization for the Intrinsic Brain: From Fingerprints to Reliability. *Cerebral Cortex* , 26(11), 4192â4211.

Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., ... Smallwood, J. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (44), 12574â12579.

Dosenbach, N. U. F. N., Nardos, B., Cohen, A. L., Fair, D. a., Power, J. D., Church, J. a., ... Schlaggar, B. L. (2010). Prediction of Individual Brain Maturity using fMRI. *Science*, 329 (5997), 1358â1361.

Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J. M., ... Margulies, D. S. (2017). Predicting brain-age from multimodal imaging data captures cognitive impairment. *NeuroImage*, 148 , 179â188.

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., ... Finn, E. S. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, 18(October), 1â11.

Shehzad, Z., Kelly, C., Reiss, P. T., Cameron Craddock, R., Emerson, J. W., McMahon, K., ... Milham, M. P. (2014). A multivariate distance-based analytic framework for connectome-wide association studies. *NeuroImage*, 93(October 2015), 74â94.

Abraham, A., Milham, M. P., Di Martino, A., Cameron Craddock, R., Samaras, D., Thirion, B., Varoquaux, G. (2017). Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. *NeuroImage* , Vol. 147, pp. 736â745. [https://doi.org/ 10.1016/j.neuroimage.2016.10.045](https://doi.org/10.1016/j.neuroimage.2016.10.045)

Greene, D. J., Church, J. A., Dosenbach, N. U. F., Nielsen, A. N., Adeyemo, B., Nardos, B., ... Schlaggar, B.









































































| Atlas | # of regions | Horizontal | Sagittal | Coronal | Atlas | # of regions | Horizontal | Sagittal | Coronal |
|--------------|--------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------|--------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Hemispheric | 2 |  |  |  | Desikan | 70 |  |  |  |
| Tissue | 3 |  |  |  | DKT | 83 |  |  |  |
| Yeo-7 | 7 |  |  |  | AAL | 116 |  |  |  |
| Yeo-7-Lib | 7 |  |  |  | Glasser | 180 |  |  |  |
| Yeo-17 | 17 |  |  |  | CPAC200 | 200 |  |  |  |
| Yeo-17-Lib | 17 |  |  |  | Schaefer200 | 200 |  |  |  |
| HOS | 21 |  |  |  | Schaefer300 | 300 |  |  |  |
| Brodmann | 41 |  |  |  | Schaefer400 | 400 |  |  |  |
| HOC | 48 |  |  |  | Slab907 | 907 |  |  |  |
| JHU | 48 |  |  |  | Schaefer1000 | 1000 |  |  |  |
| PrincetonVis | 49 |  |  |  | Slab1068 | 1068 |  |  |  |
| PP264 | 58 |  |  |  | Talairach | 1105 |  |  |  |

Figure 1: A comparison of the regions present in the major atlases available in Neuroparc. These visualizations were made using MIPAV tri-planar views on the same slice numbers. The slice numbers are formatted as (H=Horizontal, S=Sagittal, C=Coronal). For most atlases, the slice numbers were (90, 108, 90). There are a few exceptions to the slice numbers listed here for visualization purposes: JHU: (90, 108, 109), Slab907: (95, 104, 95), Slab1068: (93, 105, 93) [8] [6] [11] [12] [13] [14] [15] [2] [16] [17] [7]

L. (2016). Multivariate pattern classification of pediatric Tourette syndrome using functional connectivity MRI. *Developmental Science*, 19 (4), 581–598.

Hong, S. J., Valk, S. L., Di Martino, A., Milham, M. P., Bernhardt, B. C. (2018). Multidimensional Neuroanatomical Subtyping of Autism Spectrum Disorder. *Cerebral Cortex*, 28(10), 3578–3588.

Arslan, S., Ktena, S. I., Makropoulos, A., Robinson, E. C., Rueckert, D., Parisot, S. (2018). Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. *NeuroImage*, 170, 5–30.

Methods

Data Compilation

Neuroparc contains atlases from several locations. As previously noted, there is no current standard for atlas storage, so all gathered datasets are converted into a single format. We define this human brain atlas specification

in detail within Data Records.

Atlases

Dice Coefficient

The Dice coefficient is a measure of similarity between two sets (reference Dice Paper). Specifically, it measures a coincidence index (CI) between two sets, normalized by the size of the sets. Let h be the number of points overlapping in the sets A and B , and a and b are the sizes of their corresponding sets, then CI is defined by

$$CI = \frac{2h}{a+b}. \quad (1)$$

This formula can easily be applied to a segmented image by finding the overlap between labelled regions.

$$CI_{ij} = \frac{2h_{ij}}{a_i + b_j} \quad (2)$$

where i is the region in image 1 and j is the region in image 2. The result is a similarity matrix, as shown in Figure 2. Since this map visualizes similarity between two regions in two atlases, the information provided by the Dice map can be used quantify which regions in a given atlas are most similar to regions in another atlas.

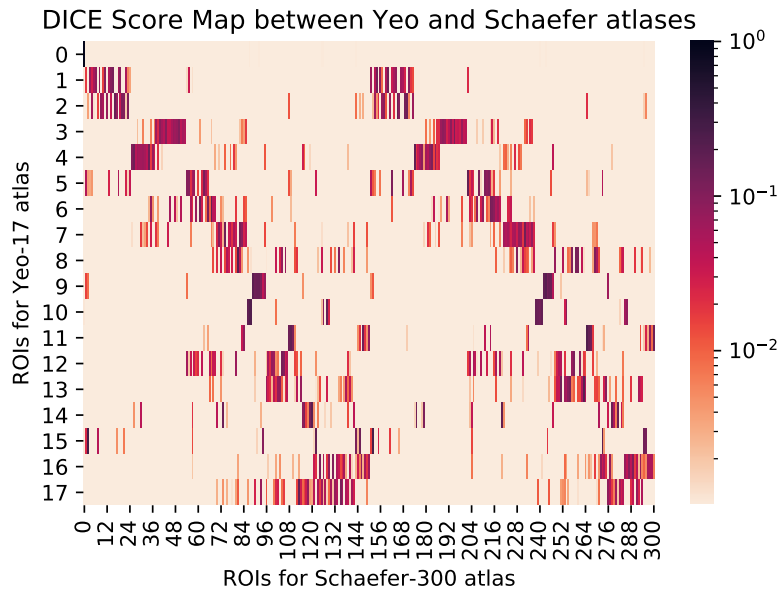


Figure 2: Dice Score Map between the Yeo 17 Networks atlas and the 300 parcellation Schaefer atlas. The strong symmetry of the Yeo atlas is apparent here.

Adjusted Mutual Information

Adjusted mutual information is another measure of the similarity of two labelled sets, quantifying how well a particular point can be identified as belonging to a region given another region. It differs from the Dice coefficient in that it tends to be more sensitive to region size and position relative to other regions. [18] But like the Dice coefficient, it is not dependent on a region's label [19]. Volumes that share many points are likely to have a higher mutual information score all else being equal [20]. To assure that all atlas comparisons were on the same scale, Neuroparc computes the adjusted mutual information score. Let $H(\cdot)$ denote entropy, N be the number of elements in total, and $H(MI_{AB})$ denote the expected mutual information for sets of size a and b . [21]

$$MI_{AB} = \sum_{a,b} P_{A,B}(a,b) \cdot \log \frac{P_{A,B}(a,b)}{P_a(a) \cdot P_b(b)} \quad (3)$$

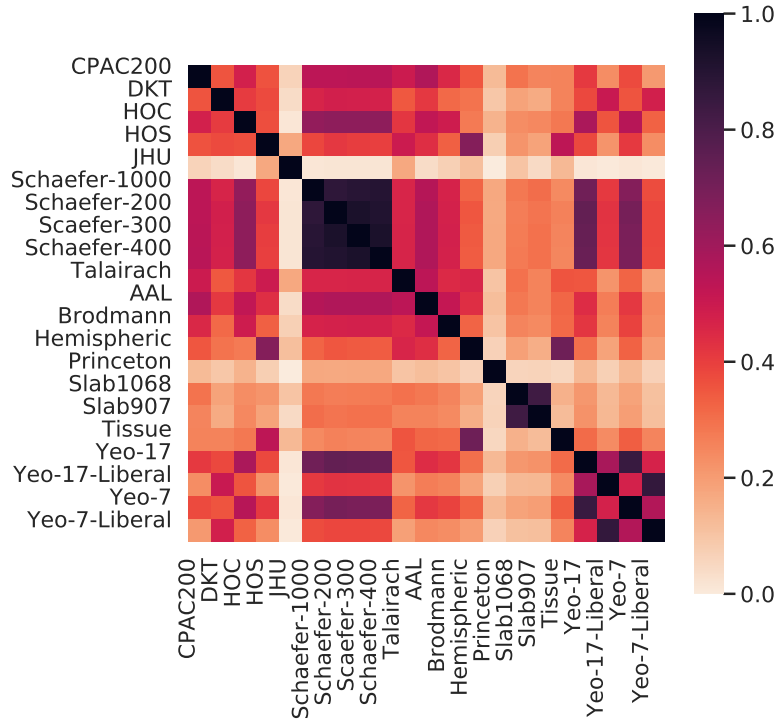


Figure 3: The adjusted mutual information between atlases contained within Neuroparc. The Schaefer atlases contain a high degree of mutual information as expected as they were built using the same algorithm.

$$AMI_{AB} = \frac{MI_{AB} - E(MI_{AB})}{avg(H(A), H(B)) - E(MI_{AB})} \quad (4)$$

Figure 3 shows the adjusted mutual information between all pairs of atlases. The information provided for this score is atlas-wide, while the Dice score was computed per region to generate a map. The similarity between groups of atlases, such as the various Schaefer atlases, and the Yeo liberal atlases, is immediately apparent.

Code availability

Code for processing is publicly available and can be found at the same GitHub link (<https://github.com/neurodata/neuroparc>) under the scripts folder. Examples of useful functions include center calculation for regions and scripts to save files in the NIfTI format. All code is provided under the Apache 2.0 License.

Visualizations are generated using both MIPAV 8.0.2 and FSLeys 5.0.10 to view the brain volumes in 2D and 3D spaces [22, 23]. Figure 1 is created using MIPAV triplanar views of each atlases with a striped LUT.

Data Records

All data records described in this paper are available primarily through the Github repository: <https://github.com/neurodata/neuroparc>. Several file types are contained in this location and are necessary for fully describing an atlas. Neuroparc introduces an atlas specification that includes a reference brain, an atlas file, and atlas metadata.

Reference Brain

To allow direct comparison between different atlases, a standard reference brain must be used for all involved atlases. Within Neuroparc, a single reference brain is provided, yielding a coordinate space. Neuroparc uses Montreal Neurological Institute 152 Nonlinear 6th generation reference brain, abbreviated MNI152NLin6 in the

file naming structure [24].

The brain is stored in a GNU-zipped NIfTI file format of a T1 weighted MRI. Three image resolutions are available in Neuroparc (1mm, 2mm, and 4mm) for flexibility. The naming convention for these files is: MNI152Nlin6_res-*<resolution>*_T1W.nii.gz. The format of the *resolution* input would be 1x1x1 for the 1 mm³ resolution.

Atlas Image

The atlas image is also a GNU-zipped NifTi file containing the parcellated reference brain according to the specifications of the atlas. The file indicates to which region each voxel belongs. Each region of interest (ROI) within this parcellated image is denoted by a unique integer ranging from 1 to n where n is the total number of ROIs. The naming convention for the atlas is: *<atlas_name>*_space-MNI152Nlin6_res-*<resolution>*.nii.gz. The *atlas_name* field is unique for each atlas image, ideally no more than two words long without a space in between (e.g. yeo-17, princetonvisual, HarvardOxford). For simplicity, only the 1 mm³ resolution parcellation is stored within the repository, but other resolutions can be calculated from the reference brain images.

Atlas Metadata

The metadata corresponding to the atlas is contained within a JSON file format. This file is split into two sections: region-wide and atlas-wide information.

The region-wide data must contain the number, label, center, and size for that region. The center and size can be calculated using provided code in Neuroparc. Although label must be specified, this information is not relevant for all atlases. In that case, NULL should be used for the labels of the regions. For hierarchical regions, the naming should be in order of largest region to smallest with a '.' in between each name. Optional fields in the region-wide data include description and color. Description can be used to provide more information than the region label if necessary. An example of this use is in the Yeo-7 Networks atlas. The label for this atlas is in the form '7Networks_2', but the description for that label is the corresponding functional network, 'Somatomotor' in this case. The color field must be given in the form [R, G, B] and is only used if the user wants to specify the colors of the regions upon visualization.

Brain-wide data must include the name, description, native coordinate space, and source of the atlas. The name field allows for more elaboration than in the name of the file. The description is more flexible, allowing the creator of an atlas to briefly describe important information for users of their atlas. The intended use case or the method of generation are examples of information provided in this field. Since all atlases in Neuroparc are stored in the same coordinate space, the coordinate space used during the creation of the atlas must be specified. Finally, the publication detailing the atlas should be included in the source field so users can have a more full understanding of the atlas being used. Optional fields for brain-wide data can all be calculated, including the number of regions, the average volume per region, whether the segmented regions are hierarchical, and if the atlas is symmetrical.

The naming convention for this file is as follows: *<atlas_name>*_space-MNI152Nlin6_res-1x1x1.json. Again, this metadata is only relevant to the 1 mm³ resolution, but other data is easily calculated when necessary.

The full description and format of the atlas specification is available within Neuroparc at https://github.com/neurodata/neuroparc/blob/devel/atlasses/Human/atlas_spec.md.

Technical Validation

All atlases included in Neuroparc have been pulled from reputable published sources (cite them), and modified to fit the above described atlas specification. Specifically, all of the atlas images were converted into the Montreal Neurological Institute 152 Nonlinear 6th generation coordinate space. All of the atlases were applied to the single T1 weighted MRI scan so that each atlas is directly comparable.

To demonstrate the validity of these atlases, we conduct the following case study. The Healthy Brain Network (HBN) is a relatively new dataset that consists of approximately 1,000 children and adolescents in New York City. HBN was created to study mental health and learning disorders. HBN includes phenotypic data in the form of tests on psychiatric, behavioral, cognitive, and lifestyle, as well as multimodal brain imaging, electroencephalography, digital voice and video recordings, genetics, and actigraphy. In this analysis, 'MGC', an independence test

implemented in the python package 'mgcpy' was used to test for correlation between functional connectomes and phenotypic data [25, 26]. For XXX individuals, we ran the NDMG pipeline on XXX individuals' resting state functional MRI data to obtain connectomes for each [27]. We then selected YYY assays to test for a dependence between the connectomes on various phenotypic properties. Doing so required carefully cleaning and purging the subject level questionnaire answers to eliminate missing data and spurious entries.

For each phenotypic test and each atlas, we ran MGC to test whether connectomes and cognitive assays were statistically dependent on one another. (Figure 4). Assays generally had either multiple atlases which found a significant correlation ($\alpha \leq 0.05$), or none. The assays which had a significant correlation for some atlases were APQ_SR, ASSQ, CELF_5_Screen, MFQ_P, PAQ_C, and SWAN. APQ_P, MFQ_SR, and PAQ_A which are the same tests but for a different age group as APQ_SR, MFQ_P, and PAQ_C. The APQ, ASSQ, CELF_5_Screen, MFQ, PAQ, and SWAN are in the categories Family Structure Stress and Trauma, ASD, Verbal Learning, Depression and Mood, Physical, and ADHD as given by the Child Mind Institute. There does not appear to be a strong connection from this alone between category of test and whether significant correlation is found. ACE, a questionnaire on the occurrence various traumatic childhood events, SCAREDP, an anxiety test for preschoolers, and PAQ_A, a questionnaire on physical activity all also showed extremely high p-values.

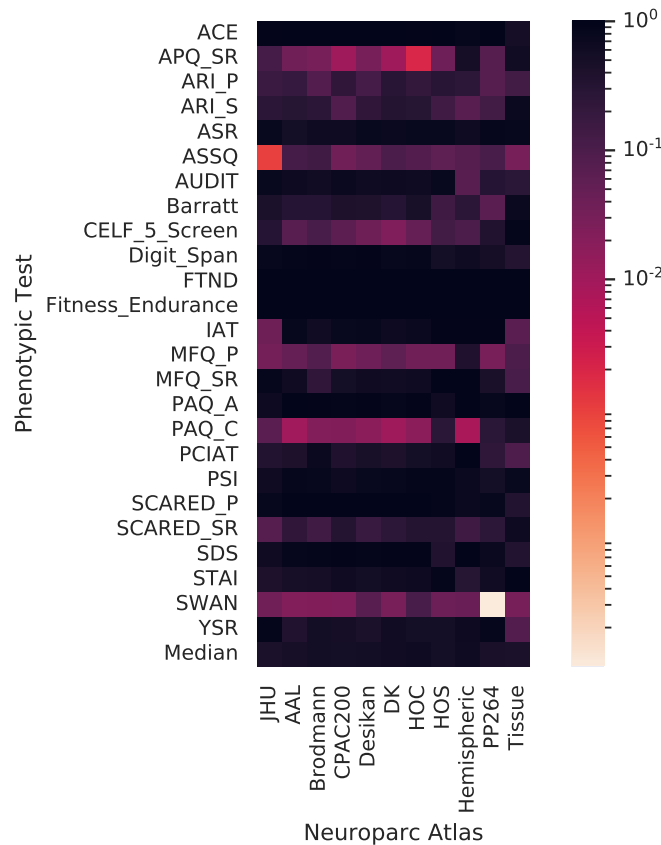


Figure 4: This bee-swarm plot shows the log of the p-value for the correlation between functional connectomes generated by different atlases and phenotypic tests. The differences in distributions demonstrate the importance of choosing the right atlas for the specific task.

Interestingly, the same tests but given in the regular version or the preschooler version frequently had very different results. This appears to indicate that such tests might have more correlation with brain connectivity and structure within different age groups.

Medians over all atlases show the results discussed above. Medians over all tests are all fairly high and not very distinct. The tissue atlas has the highest average p-value, and the HarOxCort atlas has the lowest, ranging from about 0.35 to 0.55. This does not account for whether or not the atlas is correctly finding correlations and is an

average over relatively few tests, so it is hard to interpret. It would be interesting to investigate this variation more, possibly when averaging over a higher number of tests than this analysis.

Usage Notes

The Usage Notes should contain brief instructions to assist other researchers with reuse of the data. This may include discussion of software packages that are suitable for analysing the assay data files, suggested downstream processing steps (e.g. normalization, etc.), or tips for integrating or comparing the data records with other datasets. Authors are encouraged to provide code, programs or data-processing workflows if they may help others understand or use the data. Please see our code availability policy for advice on supplying custom code alongside Data Descriptor manuscripts.

For studies involving privacy or safety controls on public access to the data, this section should describe in detail these controls, including how authors can apply to access the data, what criteria will be used to determine who may access the data, and any limitations on data use.

Acknowledgements

We would like to acknowledge generous support from National Science Foundation (NSF) under NSF Award Number EEC-1707298.

Affiliations

¹Department of Biomedical Engineering, Institute for Computational Medicine, Kavli Neuroscience Discovery Institute, Johns Hopkins University

²Department of Biostatistics, Johns Hopkins University

*corresponding author: jovo@jhu.edu

References

- [1] Alan C. Evans, Andrew L. Janke, D. Louis Collins, and Sylvain Baillet. Brain templates and atlases. *NeuroImage*, 62(2):911–922, August 2012. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2012.01.024. [1](#)
- [2] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15(1):273–289, January 2002. ISSN 1053-8119. doi: 10.1006/nimg.2001.0978. [1](#), [4](#)
- [3] Srivatsan Pallavaram, Benoit M. Dawant, Michael S. Remple, Joseph S. Neimat, Chris Kao, Peter E. Konrad, and Pierre-François D'Ágostino. Effect of brain shift on the creation of functional atlases for deep brain stimulation surgery. *5*(3):221–228. ISSN 1861-6429. doi: 10.1007/s11548-009-0391-1. URL <https://doi.org/10.1007/s11548-009-0391-1>. [1](#)
- [4] Michael I Miller, Laurent Younes, and Alain Trouvé. Diffeomorphometry and geodesic positioning systems for human anatomy. *Technology*, 2(1):36, March 2014. [1](#)
- [5] Jason W. Bohland, Hemant Bokil, Cara B. Allen, and Partha P. Mitra. The brain atlas concordance problem: quantitative comparison of anatomical parcellations. *PLoS One*, 4(9):e7200, September 2009. ISSN 1932-6203. doi: 10.1371/journal.pone.0007200. [1](#)
- [6] Rahul S. Desikan, Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, Anders M. Dale, R. Paul Maguire, Bradley T. Hyman, Marilyn S. Albert, and Ronald J. Killiany. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980, July 2006. ISSN 10538119. doi: 10.1016/j.neuroimage.2006.01.021. URL <https://linkinghub.elsevier.com/retrieve/pii/S1053811906000437>. [1](#), [4](#)

- [7] B. T. Thomas Yeo, Fenna M. Krienen, Jorge Sepulcre, Mert R. Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L. Roffman, Jordan W. Smoller, Lilla ZÁúllei, Jonathan R. Polimeni, Bruce Fischl, Hesheng Liu, and Randy L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165, September 2011. ISSN 1522-1598. doi: 10.1152/jn.00338.2011. 1, 4
- [8] Korbinian Brodmann. *Vergleichende Lokalisationslehre der Großhirnrinde : in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. 1909. 1, 4
- [9] Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Koskaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8:14, 2014. ISSN 1662-5196. doi: 10.3389/fninf.2014.00014. URL <https://www.frontiersin.org/article/10.3389/fninf.2014.00014>. 1, 2
- [10] Lindsay M. Alexander, Jasmine Escalera, Lei Ai, Charissa Andreotti, Karina Febre, Alexander Mangone, Natan Vega-Potler, Nicolas Langer, Alexis Alexander, Meagan Kovacs, Shannon Litke, Bridget O’Hagan, Jennifer Andersen, Batya Bronstein, Anastasia Bui, Marijayne Bushey, Henry Butler, Victoria Castagna, Nicolas Camacho, Elisha Chan, Danielle Citera, Jon Clucas, Samantha Cohen, Sarah Dufek, Megan Eaves, Brian Fradera, Judith Gardner, Natalie Grant-Villegas, Gabriella Green, Camille Gregory, Emily Hart, Shana Harris, Megan Horton, Danielle Kahn, Katherine Kabotyanski, Bernard Karmel, Simon P. Kelly, Kayla Kleinman, Bonhwang Koo, Eliza Kramer, Elizabeth Lennon, Catherine Lord, Ginny Mantello, Amy Margolis, Kathleen R. Merikangas, Judith Milham, Giuseppe Minniti, Rebecca Neuhaus, Alexandra Levine, Yael Osman, Lucas C. Parra, Ken R. Pugh, Amy Racanello, Anita Restrepo, Tian Saltzman, Batya Septimus, Russell Tobe, Rachel Waltz, Anna Williams, Anna Yeo, Francisco X. Castellanos, Arno Klein, Tomas Paus, Bennett L. Leventhal, R. Cameron Craddock, Harold S. Koplewicz, and Michael P. Milham. An open resource for transdiagnostic research in pediatric mental health and learning disorders. 4:170181. ISSN 2052-4463. doi: 10.1038/sdata.2017.181. URL <https://www.nature.com/articles/sdata2017181>. 1
- [11] Steven Giavasis, Daniel Clark, Caroline Froehlich, John Pellman, Sharad Sikka, Zarrar Shehzad, Ranjit Khanuja, Brian Cheung, Sebastian Urchs, Qingyang Li, Yaroslav Halchenko, Daniel Lurie, Rosalia Tunagaraza, Joshua Vogelstein, Asier Erramuzpe, Aimi Watanabe, Adam Liska, Daniel A Kessler, Chris Filo Gorgolewski, and R. Cameron Craddock. Fcp-indi/c-pac: Cpac version 1.0.0 beta, 2016. URL <https://zenodo.org/record/164638>. 4
- [12] Matthew F. Glasser, Timothy S. Coalson, Emma C. Robinson, Carl D. Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F. Beckmann, Mark Jenkinson, Stephen M. Smith, and David C. Van Essen. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, 2016. ISSN 1476-4687. doi: 10.1038/nature18933. 4
- [13] Jill M. Goldstein, Larry J. Seidman, Nikos Makris, Todd Ahern, Liam M. O’Brien, Verne S. Caviness, David N. Kennedy, Stephen V. Faraone, and Ming T. Tsuang. Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability. *Biological Psychiatry*, 61(8):935–945, April 2007. ISSN 0006-3223. doi: 10.1016/j.biopsych.2006.06.027. 4
- [14] Alexander Schaefer, Ru Kong, Evan M. Gordon, Timothy O. Laumann, Xi-Nian Zuo, Avram J. Holmes, Simon B. Eickhoff, and B. T. Thomas Yeo. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex (New York, N.Y.: 1991)*, 28(9):3095–3114, September 2018. ISSN 1460-2199. doi: 10.1093/cercor/bhx179. 4
- [15] J. Talairach and G. Szikla. Application of stereotactic concepts to the surgery of epilepsy. *Acta Neurochirurgica. Supplementum*, 30:35–54, 1980. 4
- [16] Setsu Wakana, Hangyi Jiang, Lidia M. Nagae-Poetscher, Peter C. M. van Zijl, and Susumu Mori. Fiber tract-based atlas of human white matter anatomy. *Radiology*, 230(1):77–87, January 2004. ISSN 0033-8419. doi: 10.1148/radiol.2301021640. 4
- [17] Liang Wang, Ryan E. B. Mrcuzek, Michael J. Arcaro, and Sabine Kastner. Probabilistic Maps of Visual Topography in Human Cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 25(10):3911–3931, October 2015. ISSN 1460-2199. doi: 10.1093/cercor/bhu277. 4

- [18] K. H. Zou, W. M. Wells, R. Kikinis, and S. K. Warfield. Three validation metrics for automated probabilistic image segmentation of brain tumours. *Stat Med*, 23(8):1259–1282, Apr 2004. 5
- [19] Dekang Lin. *An Information-Theoretic Definition of Similarity*. ICML '98. Kaufmann, San Francisco, Calif, 1998. ISBN 978-1-55860-556-5. URL <http://dl.acm.org/citation.cfm?id=645527.657297>. OCLC: 246383629. 5
- [20] H. B. Mitchell. Image Similarity Measures. In *Image Fusion*, pages 167–185. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-11215-7 978-3-642-11216-4. doi: 10.1007/978-3-642-11216-4_14. URL http://link.springer.com/10.1007/978-3-642-11216-4_14. 5
- [21] Kunihiro Tateoka. Assessment of Similarity Measures for Accurate Deformable Image Registration. *Journal of Nuclear Medicine & Radiation Therapy*, 03(04), 2012. ISSN 21559619. doi: 10.4172/2155-9619.1000137. URL <https://www.omicsonline.org/assessment-of-similarity-measures-for-accurate-deformable-image-registration-2155-9619.1000137.php?aid=18683>. 5
- [22] M. J. McAuliffe, F. M. Lalonde, D. McGarry, W. Gandler, K. Csaky, and B. L. Trus. Medical image processing, analysis and visualization in clinical research. In *Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001*, pages 381–386, July 2001. doi: 10.1109/CBMS.2001.941749. 6
- [23] Paul McCarthy. Fsleyes, 2019. URL <https://zenodo.org/record/1470761>. 6
- [24] Wilkin Chau and Anthony R. McIntosh. The talairach coordinate of a point in the MNI space: how to interpret it. 25(2):408–416. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2004.12.007. 7
- [25] Joshua T Vogelstein, Eric W Bridgeford, Qing Wang, Carey E Priebe, Mauro Maggioni, and Cencheng Shen. Discovering and deciphering relationships across disparate data modalities. 8:e41690. ISSN 2050-084X. doi: 10.7554/eLife.41690. URL <https://doi.org/10.7554/eLife.41690>. 8
- [26] Sambit Panda, Satish Palaniappan, Junhao Xiong, Ananya Swaminathan, Sandhya Ramachandran, Eric W. Bridgeford, Cencheng Shen, and Joshua T. Vogelstein. mgcpy: A comprehensive high dimensional independence testing python package. URL <http://arxiv.org/abs/1907.02088>. 8
- [27] Gregory Kiar, Eric W. Bridgeford, William R. Gray Roncal, Consortium for Reliability Reproducibility (CoRR), Vikram Chandrashekar, Disa Mhembere, Sephira Ryman, Xi-Nian Zuo, Daniel S. Margulies, R. Cameron Craddock, Carey E. Priebe, Rex Jung, Vince D. Calhoun, Brian Caffo, Randal Burns, Michael P. Milham, and Joshua T. Vogelstein. A high-throughput pipeline identifies robust connectomes but troublesome variability. page 188706. doi: 10.1101/188706. URL <https://www.biorxiv.org/content/10.1101/188706v4>. 8