CBPtools: A Python package for regional connectivity-based parcellation

Niels Reuter[1,2], Sarah Genon[2], Shahrzad Kharabian[2], Felix Hoffstaedter[1,2], Tobias Kalenscher[3], Kaustubh. R. Patil[1,2] & Simon B. Eickhoff[1,2]

[1]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany;
[2]Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, Jülich, Germany;
[3]Department of Comparative Psychology, Institute of Experimental Psychology, Heinrich-Heine University, Düsseldorf, Düsseldorf, Germany

n.reuter@fz-juelich.de

1. Connectivity-Based Parcellation Workflow

Regional connectivity-based parcellation (rCBP) is a widely used procedure for revealing the underlying brain organization of a region-of-interest (ROI) by mapping its whole-brain connectivity profiles [1].

- No standardized guidelines currently exist for applying rCBP
- No standardized software currently exist for applying rCBP
- Most studies utilizing it are based on in-house software.

Here, we outline and demonstrate **CBPtools**, an open-source plug-and-play yet customizable rCBP software package made in Python (3.5+). By introducing *CBPtools* we provide researchers the means to conduct reproducible, data-driven rCBP analyses on multiple neuroimaging modalities and large amounts of subject data.

Fig. 1 *CBPtools* workflow for applying rCBP to diffusion MRI (dMRI) or resting-state functional MRI (rsfMRI) data
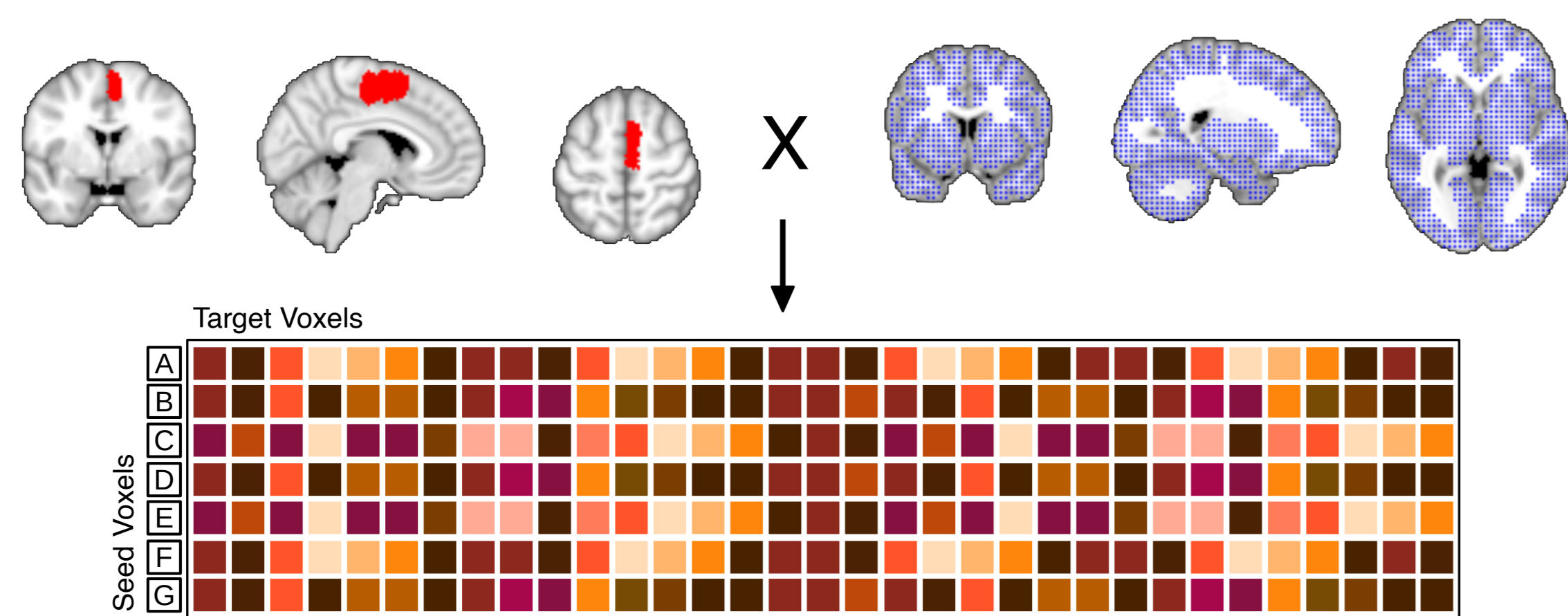
2. Connectivity Computation

Fig. 2 Computation of a seed-voxel by target-voxel connectivity matrix

Connectivity matrices are calculated for each subject using linear correlations for resting-state functional magnetic resonance imaging (MRI) and probabilistic tractography for diffusion MRI data
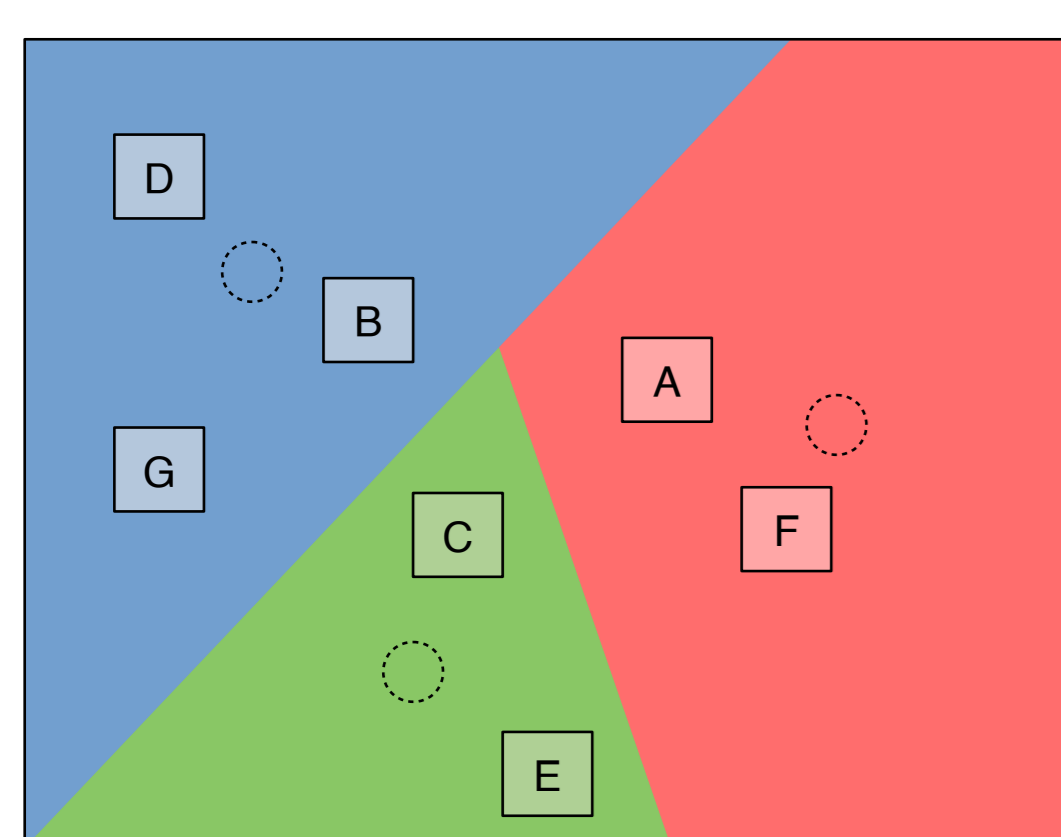
3. Clustering

Fig. 3 k-means clustering with $k = 3$. Voxels (squares) group around the 3 centroids (circles)

k-means clustering used to cluster connectivity matrices per subject and per k

- Randomly chooses k seed voxels as cluster centers (Fig. 3, circles)
- Assigns each seed voxel to closest centroid using Euclidean distance of its connectivity profile
- Move centroid to the mean of all assigned voxels and repeat
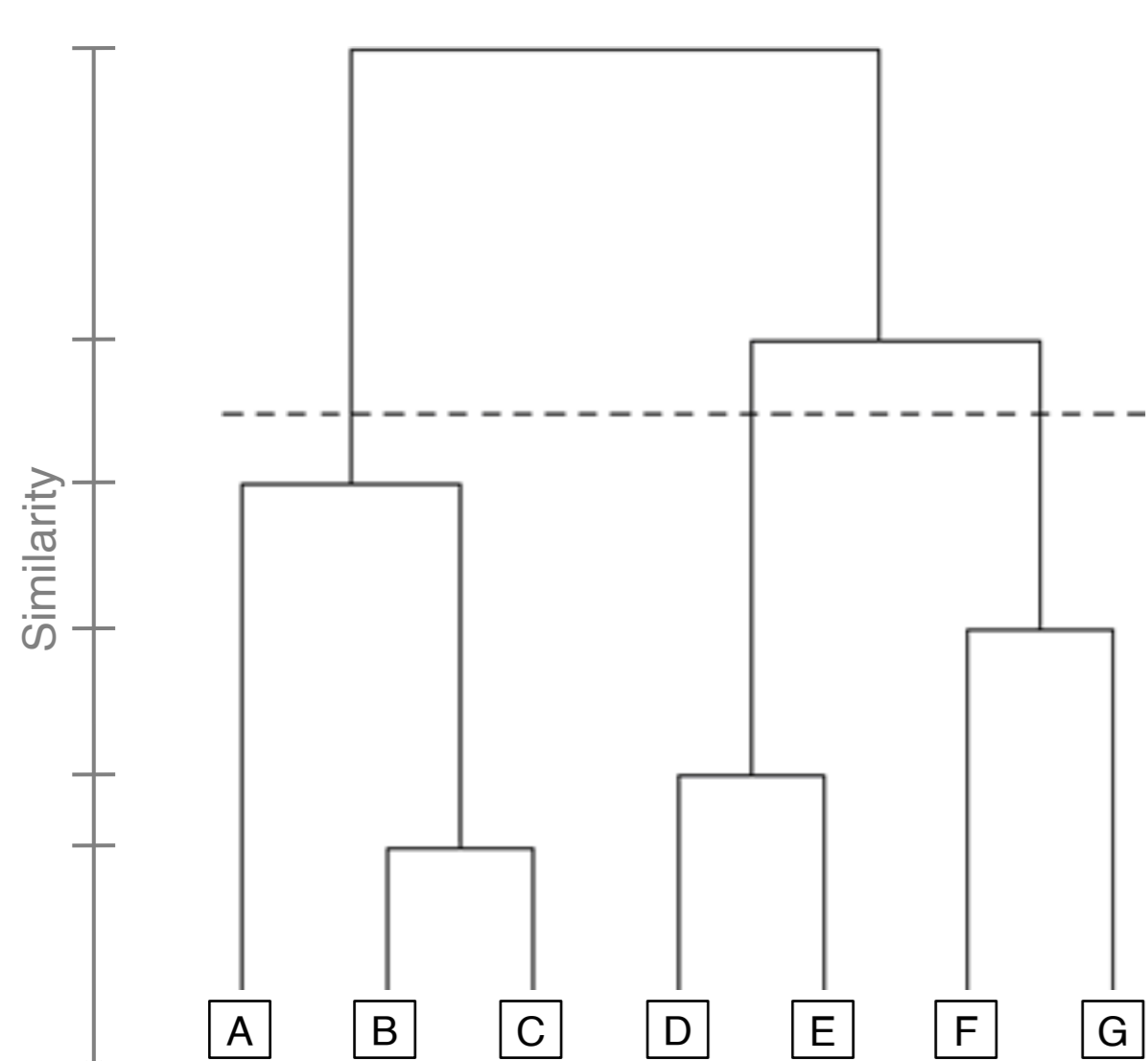
5. Grouping

Fig. 4 Hierarchical clustering with a 3-cluster cut-off. Voxels (squares) grouped at each step

Subject clustering results are used to construct a representative group clustering.

- **Hierarchical clustering** creates reference clustering of all subject clusterings
- Subject clusterings are **relabeled** to match reference clustering
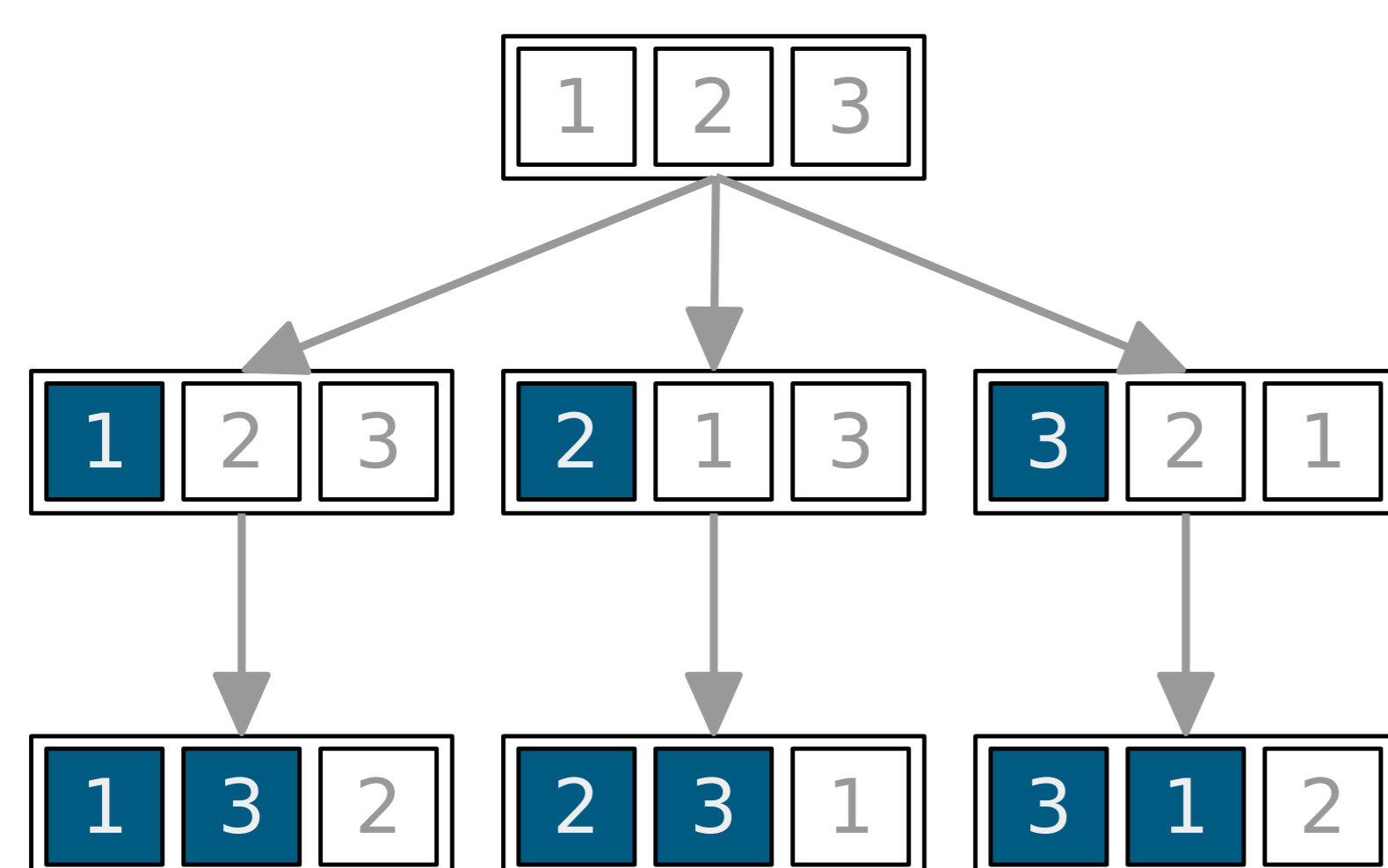- **Mode** of all relabeled subject clusterings taken as group clustering

Fig. 5 Relabeling through permutations. Here is a recursion tree for permutations of the labels "1, 2, 3"
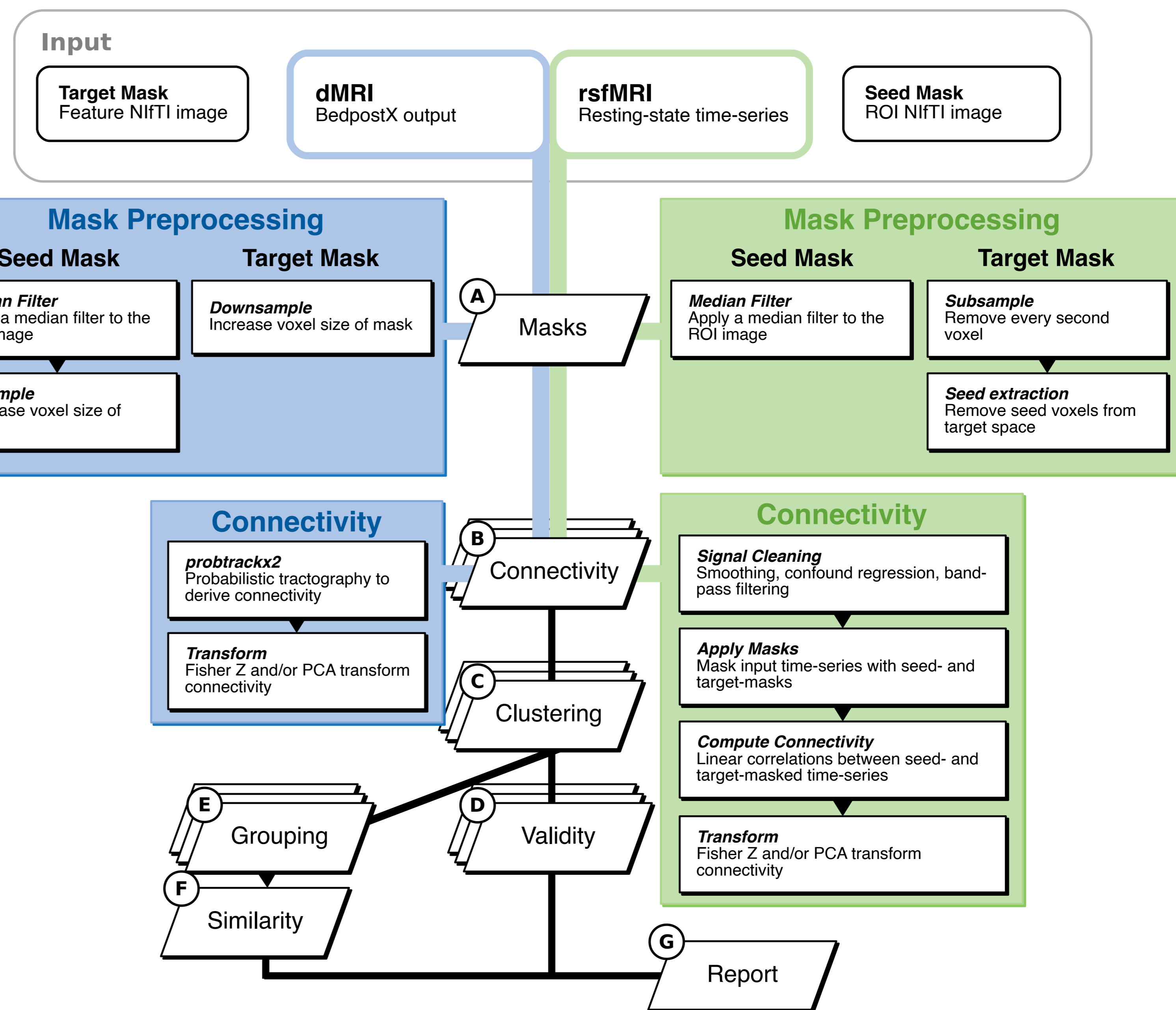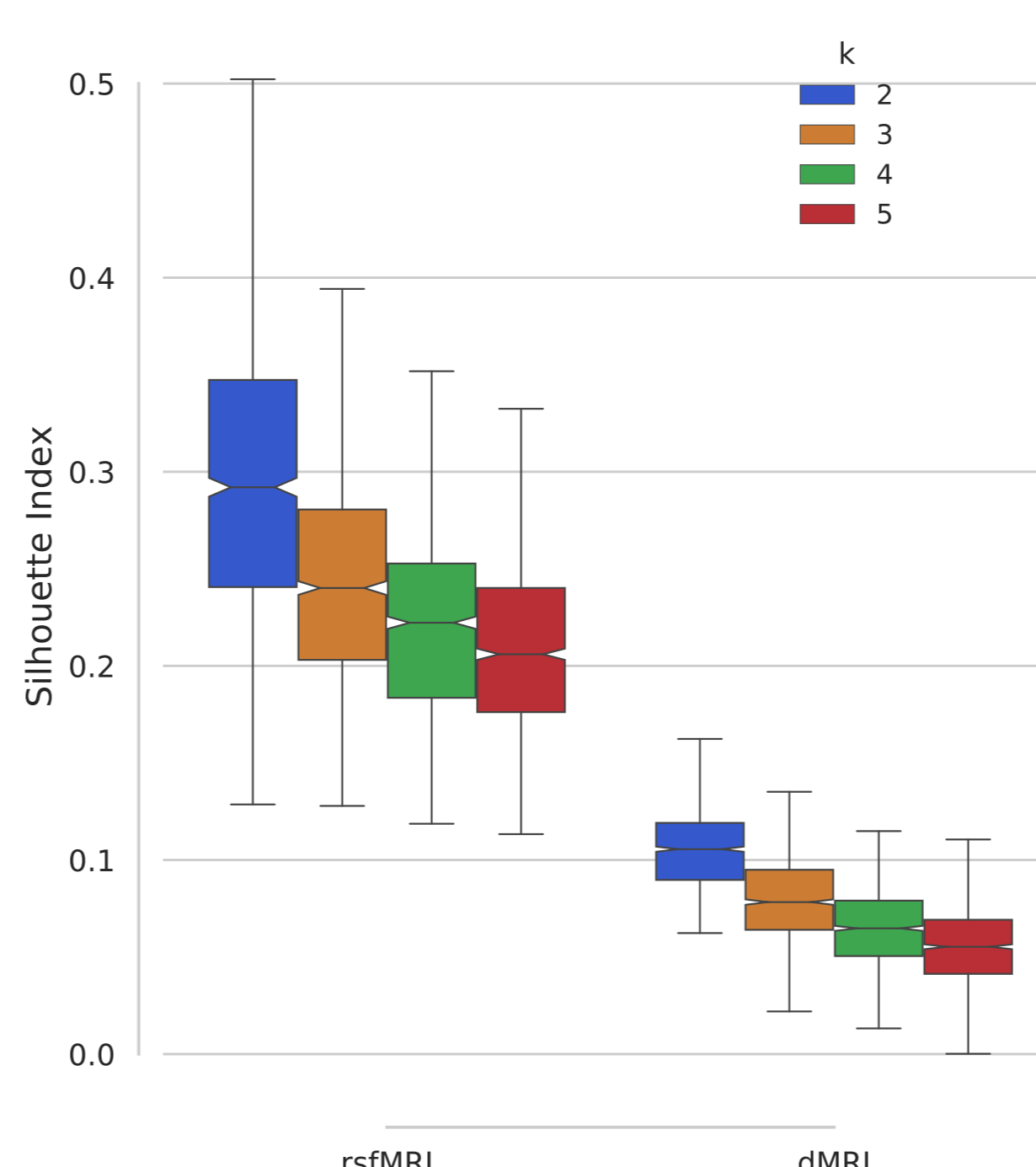
4. Validity & Similarity

Fig. 7 Silhouette index for rsfMRI and dMRI at k = 2, 3, 4, and 5

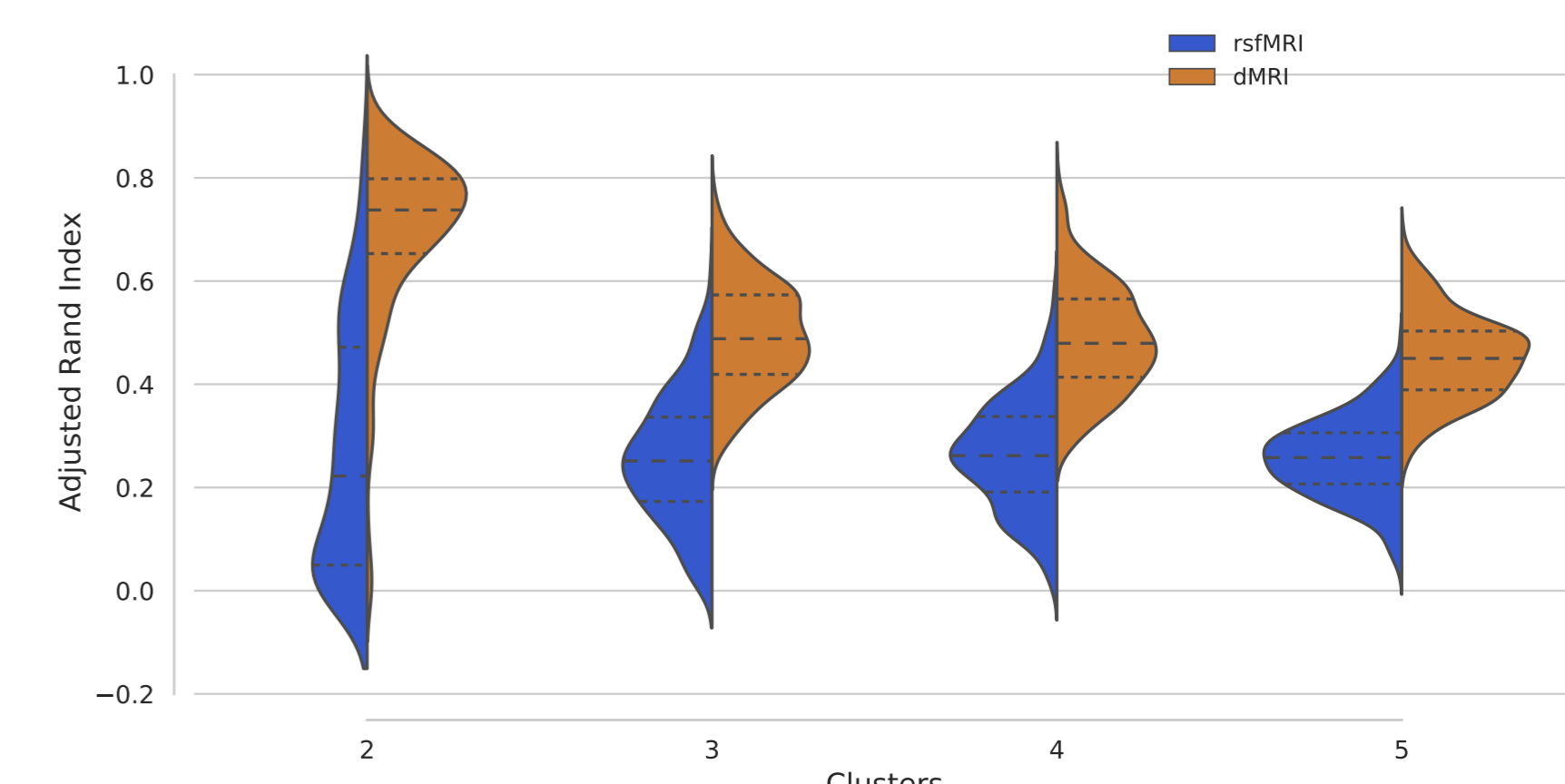**Fig. 8** Adjusted rand index scores between subject clusterings and group-clusterings for rsfMRI and dMRI at k = 2, 3, 4, and 5

- **Example data**: 300 unrelated healthy subjects from the Human connectome Project [3, 4]
- **Example region**: combined preSMA-SMA mask from the Juelich Cytoarchitectonic Atlas [2]
- **Cluster quality metrics** computed for each subject- and group-clustering at all values of clustering granularity k
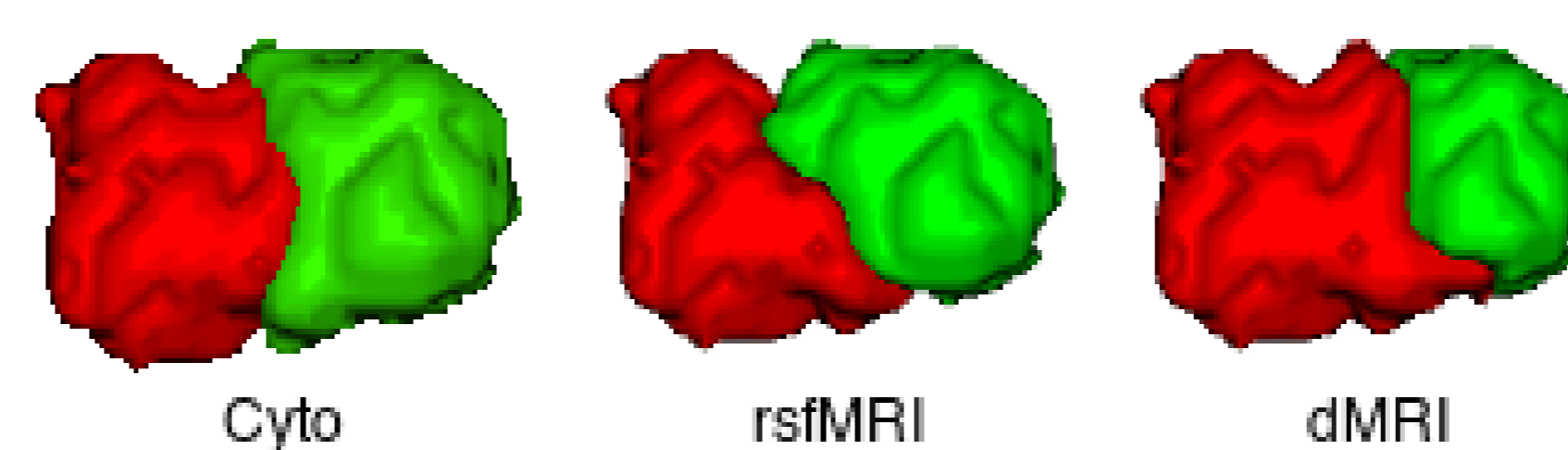
6. NIfTI Images

Fig. 9 3D representation [5] of volumetric *2-cluster combined preSMA and SMA* region-of-interest solution

- Group-clustering labels are mapped on seed image
- 2-cluster solution shown to best fit input data (Fig. 7)
- ARI between cytoarchitectonically defined subdivision and both dMRI and rsfMRI > .7

7. Conclusion

- Here we have outlined and demonstrated the effectiveness of **CBPtools** to procure rsfMRI and dMRI connectivity-derived parcellations.
- By unifying methodological choices behind the rCBP procedure, we offer a fast and stable means to effortlessly and efficiently conduct **reproducible and data-driven parcellation analysis**.
- Computational demands highlighted by complex algorithms and large datasets are mitigated by **efficient parallel execution** of the procedure using the Snakemake [6] software.

Get *CBPtools* @ https://github.com/inm7/cbptools

References:
[1] Eickhoff SB, Thirion B, Varoquaux G, Bzdok D (2015) Connectivity-based parcellation: Critique and implications. Hum Brain Mapp 36:4771–4792. doi: 10.1002/hbm.22933;
[2] Ruan J, Bludau S, Palomero-Gallagher N, et al (2018) Cytoarchitecture, probability maps, and functions of the human supplementary and pre-supplementary motor areas. Brain Struct Funct 223:4169–4186. doi: 10.1007/s00429-018-1738-6;
[3] Glasser MF, Sotiropoulos SN, Wilson JA, et al (2013) The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80:105–124. doi: 10.1016/j.neuroimage.2013.04.127;
[4] Van Essen DC, Smith SM, Barch DM, et al (2013) The WU-Minn Human Connectome Project: an overview. Neuroimage 80:62–79. doi: 10.1016/j.neuroimage.2013.05.041
[5] Generated with Mango (Multi-image analysis GUI; http://ric.uthscsa.edu/mango/)
[6] Köster J, Rahmann S (2012) Snakemake--a scalable bioinformatics workflow engine. Bioinformatics 28:2520–2522. doi: 10.1093/bioinformatics/bts480