

Title: Brain hubs defined in the group do not overlap with regions of high inter-individual variability

Authors: Smith, D.M.^{1,2}, Kraus, B. T.¹, Dworetzky, A.^{1,3}, Gordon, E. M.⁴, Gratton, C.^{1,3,5}

Affiliations:

Department of Psychology¹, Northwestern University, Evanston, IL

Department of Neurology², Division of Cognitive Neurology/Neuropsychology, The Johns Hopkins University School of Medicine, Baltimore, MD

Department of Psychology³, Florida State University, Tallahassee, FL

Department of Radiology⁴, Washington University School of Medicine, St. Louis, MO

Department of Neurology⁵, Northwestern University, Evanston, IL

Corresponding authors:

Derek Smith (dsmit358@jhmi.edu)

Caterina Gratton (cgratton@psy.fsu.edu)

Acknowledgments: This work was supported by an NSF CAREER 2048066 (CG) and NIH grants R01MH118370 (CG), T32NS047987 (DMS, BTK). This research was supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology. DMS would also like to thank The Therapeutic Cognitive Neuroscience Fund.

Competing interests: The authors have no competing interests

Abstract

Connector 'hubs' are brain regions with links to multiple networks. These regions are hypothesized to play a critical role in brain function. While hubs are often identified based on group-average functional magnetic resonance imaging (fMRI) data, there is considerable inter-subject variation in the functional connectivity profiles of the brain, especially in association regions where hubs tend to be located. Here we investigated how group hubs are related to locations of inter-individual variability. To answer this question, we examined inter-individual variation at group-level hubs in both the Midnight Scan Club and Human Connectome Project datasets. The top group hubs defined based on the participation coefficient did not overlap strongly with the most prominent regions of inter-individual variation (termed 'variants' in prior work). These hubs have relatively strong similarity across participants and consistent cross-network profiles, similar to what was seen for many other areas of cortex. Consistency across participants was further improved when these hubs were allowed to shift slightly in local position. Thus, our results demonstrate that the top group hubs defined with the participation coefficient are generally consistent across people, suggesting they may represent conserved cross-network bridges. More caution is warranted with alternative hub measures, such as community density (which are based on spatial proximity to network borders) and intermediate hub regions which show higher correspondence to locations of individual variability.

1. Introduction

In the past two decades there has been a steady increase in the application of network science methods to cognitive neuroscience, particularly in the analysis of functional networks measured with fMRI. Functional brain networks are sets of brain regions with inter-correlated fMRI Blood Oxygen Level Dependent (BOLD) signals. These networks are present when subjects are engaged in a task or at rest (Gratton et al., 2018a). Different functional brain networks have been implicated in distinct psychological functions including sensory, motor, memory, self-referential processing, and cognitive control (Biswal et al., 1995; Dosenbach et al., 2007; Dosenbach et al., 2006; Greicius et al., 2003; Seeley et al., 2007; Thomas Yeo et al., 2011). However, many complex tasks require integration of distinct functional systems, requiring an understanding of the interactions between brain networks (Bullmore and Sporns, 2009; Gratton et al., 2018b; Sporns, 2010). Network science methods provide an opportunity to quantitatively assess the distributed interactions within and across these diverse networks, as well as the role of specific regions within this network structure.

Connector hubs (from this point forward simply referred to as hubs¹) are specialized nodes within a complex system that have connections distributed across networks (Guimerà and Nunes Amaral, 2005; Power et al., 2013). Hubs appear to have an important role in brain networks, just as in many other complex systems (Bullmore and Sporns, 2009; Sporns, 2010; Van Den Heuvel and Sporns, 2011). Their position between different networks suggests that hubs may play an integrative role in brain function, perhaps associated with linking the distinct processes associated with different networks (Bertolero et al., 2018; Bertolero et al., 2017; Gratton et al., 2018b). Evidence in favor of this view comes from studies showing that hub activity is linked to a variety of tasks and cognitive processes (Bertolero et al., 2015; Cole et al., 2013) and their functional connectivity varies across task contexts (Cole et al., 2013; Gratton et al., 2016). Literature on brain lesions also suggests that hubs play a critical role in network organization and brain function. Lesions to hubs lead to wide-spread cognitive deficits, relative to lesions to non-hub regions (Warren et al., 2014) and damage to hubs is has been associated with decreased brain network segregation (Gratton et al., 2012). Jointly, these findings suggest that hubs may play a central role in facilitating inter-network communication that enables various complex behaviors.

However, most studies of functional networks and hubs have been conducted on group data, averaged across participants. Hub measures are noisy and require substantial amounts of data to reach reasonable levels of individual reliability (Gordon et al., 2017). So, the best option open to many researchers is to use group hub measures to overcome noisy data on an individual basis, an approach used by a number of groups (Fransson and Thompson, 2020; Gratton et al., 2016; Liao et al., 2013; Seitzman et al., 2019). However, the last few years have seen a substantial growth in studies showing that there is a considerable amount of inter-subject variation in functional connectivity (Bijsterbosch et al., 2018; Finn et al., 2015; Gordon et al., 2017a; Gordon et al., 2017b; Gratton et al., 2018a; Kong et al., 2019; Miranda-Dominguez et al.,

¹ Note that there are many ways to define hubs within complex systems, ranging from classic degree-based measures, to measures of centrality, to connector hubs (Bullmore, & Sporns, 2012; Van Den Heuvel, & Sporns, 2011). For the purposes of this work, we focus on connector hubs: because of their links to different networks, connector hubs are well situated for transferring information across functional systems and have been the focus of several past studies on the importance of hubs in the human brain (Bertolero et al., 2015; Cole et al., 2013; Gratton et al., 2012; Gratton et al., 2016; Power et al., 2013). Past work has demonstrated that degree-based metrics are less useful for describing correlation networks, such as those estimated from functional connectivity. In these cases, degree metrics primarily index network size (i.e., larger networks have higher degree due to their clustering and high number of ROIs; Power et al., 2013).

2014; Mueller et al., 2013; Seitzman et al., 2019). Reliance on group average data may obscure important individual differences (Smith et al., 2021), including variation in hub regions (Gordon et al., 2018).

When a large amount of high-quality fMRI data is collected, reliable individual-level network maps can be obtained and used to map differences in brain networks across people (Gordon et al., 2017b; Laumann et al., 2015). Locations of high inter-individual variability are most prominent in higher-level association regions (Finn et al., 2015; Gratton et al., 2018a; Kong et al., 2019; Mueller et al., 2013), especially in the lateral frontal cortex and near the temporoparietal junction (Seitzman et al., 2019). The locations that are most different between an individual and the group average (termed here as network ‘variants’ (Seitzman et al., 2019)) have been shown to be mostly stable over time and tasks (Kraus et al., 2021; Seitzman et al., 2019) suggesting that they may be trait-like features of brain organization. While many network variants appear close to network boundaries, a substantial subset also appear at a distance from their typical network location (Dworetsky et al., 2021; Seitzman et al., 2019).

Critically hubs, like ‘variant’ individual difference locations, are typically found in association regions, especially the frontoparietal and cinguloopercular “control” networks (Cole et al., 2013; Power et al., 2013). Thus, an important question is how hubs relate to locations of individual variation in functional connectivity. Hubs, as previously measured in large groups (“typical” hub locations), and inter-individual variability in functional connectivity could be related in at least three different ways. One possibility is suggested by the vital role that hubs seem to play in tasks (Bertolero et al., 2018; Bertolero et al., 2015; Gratton et al., 2016; Warren et al., 2014), and the negative impact of damage to these regions (Gratton et al., 2012; Warren et al., 2014). This view would support the idea that group-level hubs are critical brain locations exhibiting connectivity profiles that are conserved across individuals, where major variations would cause a significant negative impact on brain function and cognition, in the same way that nearly all humans are born with two functioning lungs. If so, we would predict that, despite the high concentration of hubs in association regions, hubs will not overlap with locations of strong inter-individual variation.

A second, contrasting, hypothesis is that, as hubs are locations with variable functional links across networks (and task contexts (Bertolero et al., 2018; Cole et al., 2013; Gratton et al., 2016)), hubs may be locations with generally malleable connectivity profiles, including profiles that can differ strongly across subjects. This would predict a correspondence between hub locations and locations of inter-individual variability. In this view, group hubs would still show connectivity across multiple networks in individual people, but the networks bridged by a given hub location would be variable across individuals.

Finally, it is possible that typical hubs observed in group average data are artifactual, representing locations of high network variability *across* people rather than a hub (an area with network connectivity evenly distributed across multiple networks) *within* a person. That is, a group-level hub could represent an area that is coupled with a single network within each individual but vary in which network is present across individuals, yielding an average connectivity profile that has connectivity evenly distributed across multiple networks. This profile would be mistaken for a hub if researchers focus on analyzing a group-level connectivity map. These scenarios are illustrated jointly in Figure 1. Note that it is also possible that different scenarios apply to different hubs.

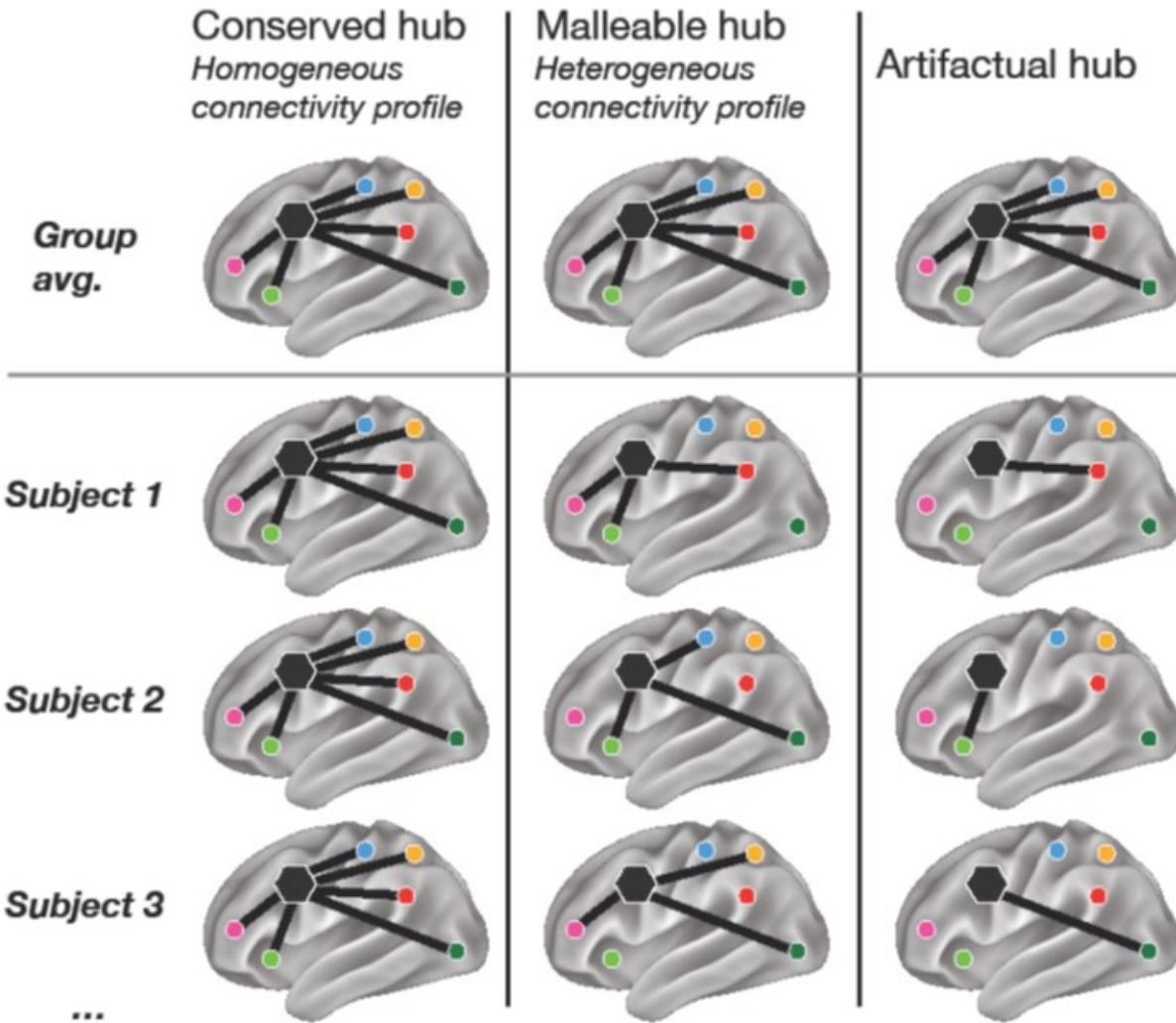


Figure 1: Schematic showing how group hubs may be organized within single individuals. The first column represents a “conserved hub”, a hub that has similar cross-network connectivity in each individual and in group-average data. The second column shows a “malleable hub”, a hub consistently present at the same location across individuals, but with distinct profiles in each individual (i.e., connecting to different networks). The third column represents an “artifactual hub”: within each individual, this location is a non-hub with connections to just one single network, but there is a great of inter-subject variability in the identity of this network. When data is averaged across subjects it appears that connectivity is evenly distributed across networks and makes this location appear to be a hub.

The primary goal of this study is to determine if hubs defined at the group level have a strong tendency to overlap with areas of inter-subject connectivity profile variability. A secondary goal is to better determine the relative likelihood of these different possibilities by determining how group hub locations vary in their connectivity profiles. If researchers wish to continue analyzing hubs defined at the group-level (in order to avoid the challenges of analyzing hubs defined at the individual-level) a comprehensive assessment of their variability is needed. We examined the relationship between hubs and measures of individual variability in both a deep “precision” fMRI dataset (Midnight Scan Club; MSC, N =9 participants with ~5 hrs of resting-

state fMRI) and the large Human Connectome Project dataset (HCP; N = 752, with 1 hr of resting-state fMRI).

2. Methods

2.1 Overview and datasets

Our goal in this project was to determine the extent to which group-average hubs show variation in their functional network profiles. We defined connector hubs using the common used participation coefficient metric (Bullmore and Sporns, 2009; Guimerà and Nunes Amaral, 2005) (see below), a measure of how distributed a region's connections are across different networks. We choose this metric since we define hubs as bridges linking functional brain networks and the participation coefficient captures how evenly distributed a node's edges are between networks. In addition, we conduct secondary analyses using an alternative metric of connector hubs termed "community density" (Power et al., 2013), a measure of how spatially proximal a region is to diverse networks. For both participation coefficient and community density hubs, we used previously published measures of group-average hubs (Power et al., 2013) based on a large 120 person dataset of healthy adults. For simplicity and maximal comparison with the prior literature, we focused on the "top" group hubs in each case – e.g., the 10 regions with the highest participation coefficient values across participants.

We then compared these top group hubs locations to locations of inter-individual variability. To examine patterns of inter-individual variability, we analyzed data from 9 highly sampled subjects from the Midnight Scan Club (Gordon et al., 2017b) and 752 subjects from the Human Connectome Project dataset (Van Essen et al., 2012). The HCP subjects are an expanded sample from that used in Seitzman et al., (2019) including all low motion individuals regardless of familial relationship (see Seitzman et al., 2019 for additional details on the composition and exclusion criteria). In each of these individuals, we identified locations of variation relative to the typical pattern, focusing first on 'variant' locations most different from the group-average, and then on continuous measures of similarity to the group and relationship to specific networks. A dataset of 120 healthy adults was used as a group average reference ((Power et al., 2011; Power et al., 2013), the same as was used to define our group-average hubs). This dataset has been described in detail in Power et al. (2013). Additional analyses also explored whether hubs could be improved with spotlight-based methods. Data collection protocols for all three datasets were approved by Washington University's institutional review board and informed consent was obtained from all participants.

2.2 Data acquisition

HCP data was acquired on a custom Siemens 3T Skyra with a custom 32-channel head coil (Van Essen et al., 2012). The HCP scanning protocol included a pair of T1-weighted (256 slices, 0.7 mm^3 isotropic resolution, TE = 2.14ms, TR = 2400ms, TI = 1000ms, flip angle = 8 degrees) and a pair of T2-weighted (256 slices, 0.7 mm^3 isotropic resolution, TR = 3200ms, TE = 565ms) images (Glasser et al., 2013). Functional scans were collected using a multi-band sequence with MB factor 8, isotropic 2 mm^3 voxels, TE of 33ms, and TR of 720ms (Glasser et al., 2013; Van Essen et al., 2012). One hour of resting state data was acquired per subject in 15 min. intervals over two separate sessions (Van Essen et al., 2012).

For the MSC, high-resolution T1-weighted (224 slices, 0.8 mm^3 isotropic resolution, TE = 3.74ms, TR = 2400ms, TI = 1000ms, flip angle = 8 degrees), T2-weighted (224 slices, 0.8 mm^3 isotropic resolution, TE = 479ms, TR = 3200ms) both with 0.8 mm^3 isotropic resolution, and resting state BOLD data were collected on a Siemens 3T Magnetom Tim Trio with a 12-channel head coil (Gordon et al., 2017b). Functional scans were collected with a gradient-echo EPI sequence, isotropic 4 mm^3 voxels, TE of 27ms, and TR of 2200ms (Gordon et al., 2017b). The MSC dataset

acquired 5 hours of resting state data per subject in 30 min. blocks over 10 separate sessions (Gordon et al., 2017b).

The WashU-120 dataset was collected on a Siemens MAGNETOM Tim Trio, 3T scanner with a Siemens 12 channel Head Matrix Coil. Both T1-weighted (127 slices, 1 mm³ isotropic resolution, TE = 3.06ms, TR = 2400ms, TI = 1000ms, flip angle = 8 degrees) and T2-weighted (32 slices, 2 × 1 × 4 mm³ resolution, TE = 84ms, TR = 6800ms) scans were collected (Power et al., 2013; Power et al., 2014). The amount of resting state data collected per subject ranged from 7.7 to 16.5 min (TE = 27ms, isotropic 4mm³ voxels; TR = 2500ms, flip angle = 90 degrees).

2.3 Preprocessing

2.3.1 General Preprocessing

For each of the three datasets the T1-weighted images were processed via automatic segmentation of the gray matter, white matter, and ventricles in Freesurfer 5.3 (Fischl et al., 2002). The default recon-all command in Freesurfer was then applied to produce the anatomical surface for each subject (Dale, 1999). In the MSC dataset, these surfaces were manually edited to improve the quality of the registration. The surfaces were registered to the fs_LR_32k surface space via the procedure outlined in Glasser et al. (2013).

For the HCP dataset the volumetric BOLD time series from each run were concatenated. Slice timing correction was not performed for the HCP dataset in accordance with the minimal preprocessing pipeline guidelines (Glasser et al., 2013). Field inhomogeneity distortion correction was conducted using the mean field map. Motion correction was conducted using rigid body transforms aligning to the first frame of the first run. After this step whole-brain intensity values across each BOLD run were normalized to achieve a mode value of 1000 (Miezin et al., 2000). This dataset was processed in MNI atlas space with 2 mm isotropic voxels.

The preprocessing pipelines used for the MSC and WashU-120 datasets were almost identical to the HCP with some minor exceptions. Field inhomogeneity distortion correction using the mean field map was applied to all sessions for the MSC dataset but not for the WashU 120 given that field maps were not collected for this dataset (Gordon et al., 2017b; Laumann et al., 2015). Slice timing correction was performed in both the MSC and WashU-120 datasets using sinc interpolation to account for temporal misalignment in slice acquisition time. This was followed by motion correction which was performed within and across BOLD runs (aligned to the first frame of the first run) via a rigid body transformation. Then whole-brain intensity values across each BOLD run were normalized to achieve a mode value of 1000 (Miezin et al., 2000). For the WashU-120 functional BOLD data was then registered directly to a high resolution T1-weighted structural image from each participant. For the MSC functional BOLD data was first registered to a T2-weighted image and then to the T1. An affine transformation was used for registration in both datasets. The T1-weighted image was aligned to a template atlas (Lancaster et al., 2000) conforming to Talairach stereotactic atlas space using an affine transformation. All computed transformations and re-sampling to 3 mm isotropic voxels were simultaneously applied at the end of these steps.

2.3.2 Resting State Connectivity Pipeline

Steps were taken to mitigate the influence of artifacts on resting state BOLD time series. The impact of nuisance signals was attenuated via regression of average signal from the white matter, ventricles, global signal, motion parameters, as well as their derivatives and expansion terms (Friston et al., 1998; Power et al., 2014). We acknowledge that there is controversy surrounding the application of global signal regression (GSR); however a large number of

studies comparing processing strategies have demonstrated that GSR is consistently one of the best approaches to remove artifacts, especially from movement and respiration (Burgess et al., 2016; Ciric et al., 2017; Power et al., 2012; Power et al., 2014; Power et al., 2015; Satterthwaite et al., 2013). Lack of GSR allows spurious differences in motion and breathing to distort correlations. Given that motion and respiration can vary across individuals and contaminate identification of selective individual differences in brain networks (Siegel et al., 2017).

The impact of motion was further mitigated via the removal of frames with framewise displacement > 0.2 mm, in addition to sequences containing less than 5 contiguous low motion frames, the first 30 seconds of each run, and runs with < 50 low motion frames (Power et al., 2014). In the HCP dataset, before censoring high-motion frames, motion parameters were low-pass filtered at 0.1 Hz to reduce the effects of respiratory artifacts on motion estimates stemming from the short-TR multi-band acquisition (Fair et al., 2020; Gratton et al., 2020; Siegel et al., 2017). Then a filtered FD threshold of 0.1 mm was applied to censor frames (mean filtered FD = 0.0203 mm (± 0.0052); mean number of frames retained = 4335 (± 330)). The same filtering procedure was also applied to two MSC subjects (MSC03 and MSC10) with respiratory contamination in their motion parameters. In all cases, flagged head motion frames were removed and the time points were replaced with interpolated data using a power-spectral matched approach (Power et al., 2014), after which a bandpass filter (0.009 Hz-0.08 Hz) was applied to the data.

As previous results have indicated that ~45 min. of low motion data are necessary to achieve high reliability of network variant locations (Kraus et al., 2021; Seitzman et al., 2019), we then removed any participant with less than 75%, or 45 min., of data in the HCP. In the 1200-HCP release, this resulted in 752/1206 final participants. In the MSC dataset 9/10 participants were retained (Gordon et al., 2017b). The excluded MSC participant was removed due to high motion and drowsiness (Gordon et al., 2017b; Laumann et al., 2017).

For all datasets the processed BOLD data were mapped to each individual's native midthickness surface via the ribbon-constrained sampling procedure (Marcus et al., 2013). Then, the mapped data were registered to the fsaverage surface in one step using the deformation map generated from the ribbon-constrained sampling procedure described in Glasser et al., (2013). Next, smoothing was conducted via a geodesic Gaussian smoothing kernel to the surface registered data (FWHM = 6 mm, sigma = 2.55) (Gordon et al., 2016; Marcus et al., 2011). Temporally interpolated frames were then removed prior to functional connectivity analysis. Functional connectivity was calculated as the Pearson correlation coefficient between different cortical locations, based on time series averaged across regions.

2.4 Regions of interest and functional brain networks

A set of 264 spherical (10 mm diameter) regions of interest from (Power et al., 2011) were used as a basis to define group-average brain hubs (Power et al., 2013). These regions divide into networks largely overlapping with the 14 canonical networks defined in (Gordon et al., 2017a): the default mode (DMN), visual, fronto-parietal (FP), dorsal attention (DAN), language (Lang), salience, cingulo-opercular (CO), somatomotor dorsal (SMd), somatomotor lateral (SMl), auditory, temporal pole (Tpole), medial temporal lobe (MTL), parietal medial (PMN), and parieto-occipital (PON). Group hubs were defined from the 264 spherical regions, based on participation coefficient estimates previously published in Power et al. (2013) which were calculated from the group average of the WashU-120 dataset (see Section 2.5).

The 14 canonical networks are also defined at the cortical surface vertex level in the WashU-120 group average (Laumann et al., 2015). The cortical surface networks were used for the network profile analyses (see section 2.6.3). These networks were defined with the Infomap clustering algorithm (Rosvall and Bergstrom, 2008) which yielded data-driven functional network definitions for a range of edge density thresholds from 0.3%-5% (Gordon et al., 2017b). A group

average network consensus map was derived by collapsing network definitions across thresholds. This was done through a consensus procedure used in numerous published studies (Gordon et al., 2017b; Laumann et al., 2015) to collapse network labels across Infomap edge density thresholds. Each node was given a network assignment based on the sparsest threshold at which it was successfully assigned. Following this step node assignments were adjusted by the removal of small networks that were only detected at one threshold. The approach aims to integrate information dense thresholds in which more nodes were successfully assigned and more sparse thresholds which tend to produce smaller networks.

Measures of inter-individual variation (spatial correlations between an individual and the group average and identification of network variants) were calculated at the single cortical vertex, rather than regional, level. These are described in Section 2.6.

2.5 Hub Definition

Hubs were defined by two approaches: participation coefficient and community density. Primary analyses focused on the locations of brain hubs defined in group average data relative to locations of individual differences. These were derived from the WashU-120 group-average dataset based on published values from Power et al., (2013).

2.5.1 Participation coefficient hubs:

The participation coefficient is a graph theoretic measure that captures how evenly distributed a node's connections are across networks (Guimerà and Nunes Amaral, 2005); see Figure 2A for schematic). The participation coefficient for node i is defined as $PC_i = 1 - \frac{\sum_{s=1}^{N_M} (\frac{k_{is}}{k_i})^2}{N_M}$, where k_i is the degree (the number of edges/connections to nodes in the given node's module/network) of node i , k_{is} is the number of edges of node i to nodes in module/network s , and N_M is the total number of networks/modules in the graph. In the original work by Guimerà and Amaral (Guimerà and Nunes Amaral, 2005) connector hubs were defined through a joint criteria of having both high participation coefficient and high within-module degree (proportion of connections within their own system). However, recent work has demonstrated that in functional brain networks within-module degree does not show a sufficiently broad distribution to enable direct application of the Guimerà & Amaral joint criteria for connector hubs (Power et al., 2013), and has instead focused more exclusively on hubs defined with the participation coefficient.

Participation coefficient hubs in the group were identified based on a previously published analysis of the large WashU-120 group-average dataset (Power et al., 2013). In brief, in that work the participation coefficient was calculated for each of 264 regions (nodes) for a range of sparsity thresholds from 2-10% edge density in 1% steps (network definitions were derived at each threshold for these calculations as well using the Infomap community detection algorithm [(Rosvall and Bergstrom, 2008)]). Participation coefficient values were then normalized and summed across thresholds to result in a final value for each region. This normalization was conducted per threshold of analysis (not conducted across subjects) to make the range of participation coefficient values comparable across thresholds. The ten regions with the highest participation coefficient values out of a set of 264 regions of interest were selected for analysis. The MNI coordinates of the centers of each of these ten spherical regions of interest were projected to a vertex on the Conte69 midthickness surface and dilated to a 5mm radius. In addition, these locations were checked to determine if more than 30% of their vertices overlapped with a low signal mask (mean BOLD signal less than 750 computed as in (Ojemann et al., 1997)). None of the top 10 hub vertices overlapped with the low signal mask. In additional

analyses, we examined continuous values of participation coefficient across all regions of interest, after removing those with > 30% of vertices in low signal regions.

2.5.2 Community density hubs:

Hubs defined by community density were based on the top 10 community density peaks listed in Power et al., (2013). In brief, these were defined through the following procedure. Cortical voxels were assigned to networks using the Infomap clustering algorithm (Rosvall and Bergstrom, 2008) at a range of density thresholds (0.5-2.5% at 0.5% intervals). Community density was defined as the number of networks appearing within a radius of a given voxel (with radii ranging from 5-10mm in increments of 1mm). Values were summed across thresholds and radii after normalizing the values within each analysis, resulting in one final community density value per voxel. The MNI coordinates of the top 10 community density peaks were then projected to the cortical surface and dilated 5mm. As before, regions were checked for overlap with a low signal mask; none of these regions exceeded the overlap threshold.

2.5.3 Locations of inter-subject variation in functional connectivity

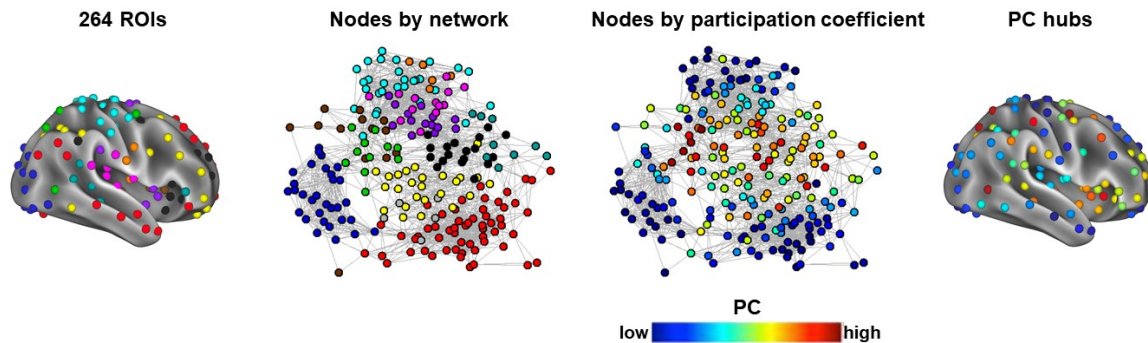
The similarity of an individual's connectivity profile and the group average connectivity profile was gauged via a spatial correlation following previously published methods (Laumann et al., 2015; Seitzman et al., 2019); see Figure 2b for schematic). For each vertex on the cortical surface, its BOLD time series was correlated with the time series for every other vertex to form a seed correlation map. Each seed correlation map (connectivity profile for a given location) was Fisher Z transformed, vectorized, and then correlated (Pearson correlation) with the Fisher Z transformed correlation vector of the corresponding cortical vertex in the group average data map, resulting in a single similarity value for that location. Across all locations we use these values to form a map of correlations between the individual level connectivity profile and the group average connectivity profile, which we refer to in this work as a "*similarity map*".

Areas of extreme idiosyncratic functional connectivity, which we call "*network variants*" were then defined from this map using a recently developed procedure (Seitzman et al., 2019). This procedure was designed to find locations within a subject that exhibit highly deviant connectivity profiles that do not fit with the connectivity pattern of its group average network of typical assignment hence the term network variant. We began with the similarity map and then identified the locations that were most dissimilar (bottom 10%) between an individual and the group. The procedure has been shown to yield very similar results across different thresholds indicating that threshold is not a major determinant of the results (Kraus et al., 2021; Seitzman et al., 2019). Regions were required to be composed of at least 50 contiguous vertices, fall outside areas with a low signal (mean BOLD signal less than 750 computed as in (Ojemann et al., 1997)), and to not overlap with the network the area's vertices were originally assigned (Seitzman et al., 2019). In order to determine where variants are most frequently found, we created an overlap map of the network variants across participants. The frequency of variants was defined as the percentage of subjects with a variant at a given cortical location (vertex). Variant frequency maps were produced for both the HCP and MSC datasets (Seitzman et al., 2019). Notably, network variants do not appear to exhibit a relationship with areas of anatomical variability. Seitzman and colleagues (Seitzman et al., 2019) found very low overlap between variants and deformations due to surface registration (mean dice overlap = 0.0001). Similar findings have been reported with other approaches to mapping individual differences in functional networks (Gordon et al., 2017b), showing that these locations do not relate strongly to areal distortion, sulcal depth mis-alignment, and curvature mis-alignment metrics (Gordon et

al., 2017b). Thus, variation in functional connectivity does not seem to be strongly associated to individual variations in gross anatomy.

One potential pitfall of this procedure is its dependence on a group average reference dataset. The findings could be specific to the reference. Yet, an analysis of the MSC dataset showed that when the average of all MSC subjects other than a given subject was used as a reference the resulting similarity maps were highly correlated (Range: .81-.87) with WashU-120 references (see Figure S5).

A



B

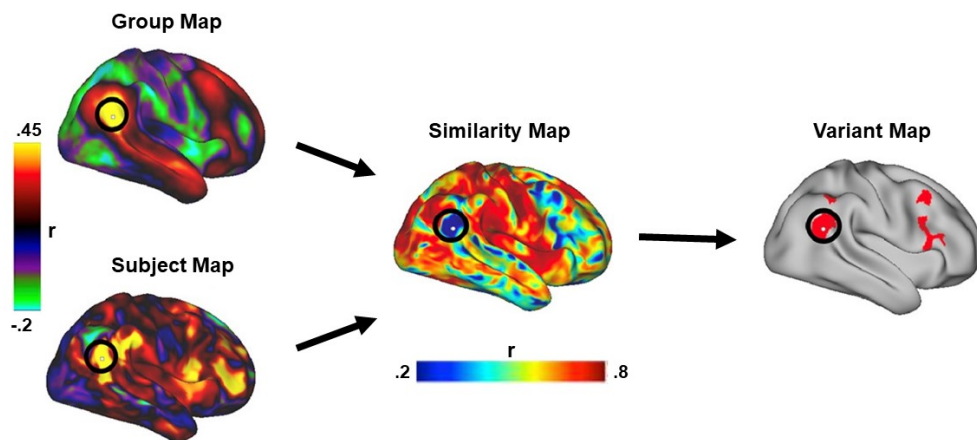


Figure 2: A schematic of the identification of connector hubs and locations of inter-individual variability. (A) Connector hubs are locations with connections to multiple networks. Brain network regions, or ‘nodes’ can be depicted on the brain based on their location (left image) or as a spring-embedded network (middle left), where nodes with more connections are placed closer together (colors = networks). In this depiction, it becomes clear that some nodes lie intermediate to multiple networks, a characteristic that can be quantified with the participation coefficient (Guimera & Amaral 2005; middle right). Here, we ask how connector hubs relate to locations of inter-individual variability. (B) The identification of locations of inter-individual variability starts by comparing the connectivity profile for each vertex in an individual level seed connectivity map (subject map) to the seed connectivity map for the corresponding vertex in the group average reference dataset (group map). For each vertex the spatial correlation between these two maps is calculated to produce a similarity map that represents how similar an individual is to the group at each vertex. Variants are defined as sets of at least 50 contiguous vertices, not falling in low SNR locations, that are all in the bottom similarity decile (lowest 10% of locations).

2.6 Relationship between hubs and locations of individual variability

The relationship between hubs and locations of individual variability was analyzed in several ways. First, we examined whether hubs collectively overlap with network variant locations (locations of particularly strong variability). We also examined to extent to which single hub locations overlapped with variants. Second, we used the continuous individual-to-group similarity map to examine hub locations variability with a finer resolution. Third, we examined

the network profile of connector hubs within single individuals and employed a local spotlight procedure to determine if it was possible to slightly adjust the position of hubs to improve inter-subject correspondence.

2.6.1 Quantification of overlap between group-average hubs and variants

We first examined the overlap of variants with group hub locations. The top 10 group hubs were defined as described above, using both participation coefficient and community density. A variant frequency map was then produced for each dataset, capturing the frequency of variants at each cortical location. We then measured the overlap between these maps.

To determine if the frequency of variants at hub locations is greater than what would be expected for a random set of cortical locations of the same size with a similar spatial configuration, a null distribution was generated by randomly rotating a set of hubs across the cortical surface (Gordon et al., 2016). For each set of hubs (either participation coefficient or community density based) 1000 rotations were randomly generated and performed within each hemisphere. If any of the hubs intersected with the medial wall the rotation was recalculated until none of the hub locations overlapped with the medial wall. Given that the variant definition procedure ignores low signal zones, rotated hubs that intersected with a low signal mask were ignored. For each rotation, we calculated the average frequency of variants at rotated hubs. Across 1000 rotations, this produced a null distribution of the expected variant frequency at rotated hubs. We then compared the actual variant frequency for hubs (averaged across hub vertices) with that of the rotated distribution. This distribution was used to derive percentiles and 95% confidence intervals for the variant frequency of hubs. A similar analysis was conducted for each single hub as well, in this case comparing variation at that hub to a null made based on rotating only that hub location.

2.6.2 Quantification of overlap between group-average hubs and subtle variation

The aforementioned analyses are geared towards determining if hubs overlap with areas of extreme inter-subject deviation in functional connectivity. We also conducted a secondary analysis to investigate whether top hubs overlapped with more subtle forms of inter-subject variation. For this analysis, we used the continuous (unthresholded) individual-to-group similarity map described in section 2.5.3 for variant definition. For each vertex, a continuous value represents how similar this location is to the group average in a given MSC or HCP subject. We then determined whether this individual-to-group similarity of hubs was lower than expected by chance by comparing the values at hub locations with the values obtained through random rotations of hubs, as described above.

In addition, we carried out a supplemental analysis to examine how regions across the full participation coefficient spectrum varied across individuals. All of the Power 264 nodes with the exception of cerebellar, subcortical, and nodes that were not assigned to a network (45 nodes excluded) were projected to cortical surface. Nodes with 30% of their vertices overlapping with the low signal mask (7 nodes) were removed from analysis. For each of the remaining nodes the correlation between the individual level connectivity profile and the group average was calculated (similarity) and Fisher Z transformed in each MSC participant. The Fisher Z transformed similarity values were correlated with the sum of the participation coefficient across edge density thresholds.

2.6.3 Quantification of the network profile of hubs across individuals

Next, we examined the network profile of hub locations within individual participants of the MSC, to determine whether they exhibited high connectivity to multiple (similar) networks. For

each participation coefficient hub, the average correlation of the hub to each of the 14 canonical networks (defined on the cortical surface; see section 2.4) was calculated. This resulted in a 14x1 network correlation vector for each participant for each hub that we call a ‘network profile’.

2.6.4 Local adjustments of group hub locations

Launching from the results of the network profile procedure above (2.6.3), we asked whether group hub locations could be slightly spatially shifted within individuals to improve the correspondence of hubs across people. A spotlight procedure was applied to the MSC dataset in an effort to improve the correspondence of hubs across people. A spotlight was formed by dilating 10mm around the central vertex of a group hub. The hub center was then moved throughout this spotlight (with hub extent still defined based on a 5mm dilation). Potential hub locations with less than 70% of its vertices within a single network were removed from consideration. These potential hub locations were then trimmed to encompass a single network (using the vertex-wise canonical network maps described in section 2.4).

At each potential hub location, we then examined how hubs were related to each of the 14 canonical networks, creating a network profile vector as described in section 2.6.3. The resulting network profiles were compared (via Euclidean distance) to a group-average reference profile based on the WashU-120 dataset. The best fitting (lowest Euclidean distance to the reference) potential hub location was selected as the final adjusted hub location for a given individual. Euclidean distance was chosen since it takes both magnitude differences and relative standing consistency into account. The sensitivity of this analysis to the choice of distance metric was tested by comparing findings obtained with different distance metrics (Pearson and Spearman correlation based distance metrics). Both alternative distance metrics yielded results that were extremely similar to those found with Euclidean distance (see Supplemental Figure 6).

Initial results of this analysis were examined qualitatively. The robustness of this procedure was then tested quantitatively by splitting each MSC subject’s data in half (odd and even sessions). For each subject the spotlight procedure was applied to each set (odd and even). We then examined the improved similarity to the WashU-120 reference set in the opposing set.

2.7 Data and code availability

All of the data have previously been made publicly available (HCP: <https://www.humanconnectome.org/>; MSC: <https://openneuro.org/datasets/ds000224/versions/00002>; WashU 120: <https://legacy.openfmri.org/dataset/ds000243/>). Code for analysis related to network variants in MATLAB is available at: <https://github.com/GrattonLab/SeitzmanGratton-2019-PNAS>; other code related to MSC processing can be found at: <https://github.com/MidnightScanClub>. Code related specifically to the analyses in this article will be located at this link upon publication: <https://github.com/GrattonLab/>.

3. Results

3.1 Overview

The aim of our investigation was to determine how group-defined hubs relate to areas of variation in functional connectivity across people. We hypothesized that group-level hubs are critical regions with connectivity profiles that should be conserved across individuals; thus they should exhibit relatively little variability in these profiles across subjects. A second contrasting hypothesis is that these regions are malleable in their functional connectivity, exhibiting a high

degree of variability in their connectivity profiles across participants (but still remaining a hub across participants). A final alternative is that group-level hubs arise from averaging multiple different non-hub regions each associated with a different single network across individuals, creating an artifactual hub representation (see Figure 1).

To test these alternatives, we examined how group-level hubs relate to locations of individual variation in functional connectivity, focusing first on ‘variant’ locations with especially strong variations across people, and then at continuous measures of inter-individual variability. We then characterized group hubs in more detail, by examining their network profiles and determining whether small local adjustments in those profiles improved correspondence across individuals.

3.2 Group hubs defined by the participation coefficient do not overlap with network variants

Do group-level hubs correspond with locations that vary strongly across people? We began by examining the locations of the strongest group-level hubs, defined as the top 10 participation coefficient regions estimated from a group-average of 120 healthy young adults (*WashU-120*) in previous work (Power et al., 2013). The participation coefficient measure defines hubs as regions with distributed functional connectivity across networks (see section 2.5.1). In parallel, we identified locations of high inter-subject variability (locations we term “network variants”; Seitzman et al 2019). We asked how group hub locations corresponded with network variants across people.

As can be seen in Figure 3, network variant locations are especially frequent in the temporoparietal junction, lateral frontal cortex, and the dorsal posterior cingulate. Participation coefficient group hubs are also found in association systems, but more prominently in the anterior insula, superior parietal cortex, and dorsolateral and medial frontal cortex (Figure 3, Power et al., 2013). Thus, there qualitatively appears to be low overlap between participation coefficient hubs and locations that frequently vary across people.

Confirming this qualitative description, participation coefficient hubs, as a whole, occurred over variants at a low rate in the HCP dataset, within the bounds of what would be expected by chance relative to 1000 random rotations of the hub set (Fig. 3B; 1.69% of people had variants at hub locations, at the 7th percentile of random rotations, 95% CI [1.21%, 8.90%]). This pattern replicated in the precision MSC dataset (Supp. Fig. 1A; including in a second set of analyses omitting global signal regression, Supp. Figs. 7 & 8). Similar results were also seen when unthresholded versions of connector hub measures were used to identify the top group level hubs (Supp. Figs. 9 & 10). Thus, group hubs defined by the participation coefficient do not show significant correspondence to areas of strong inter-subject variability.

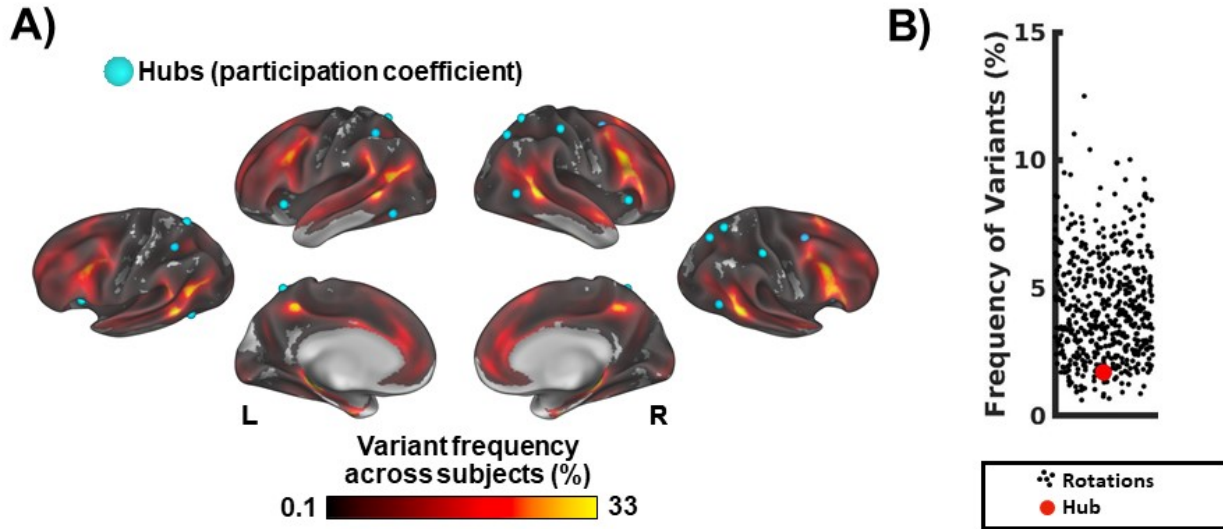


Figure 3: Comparison of top group participation coefficient hubs to locations of inter-individual variation. A) The locations of the top 10 group hubs defined using the participation coefficient are represented as light blue foci on the cortical surface. The heat map displayed on the cortical surface captures the frequency of variants (percentage of subjects with a variant at a location) based on the HCP dataset, with warmer colors indicating greater variant frequency. B) Comparison of the true frequency of variants at participation coefficient hubs (red dot) relative to random rotations of the hub set (black dots). Group hubs defined by the participation coefficient do not frequently overlap with areas of idiosyncratic functional connectivity.

We also examined how each single hub varied (Fig. 4). The 10 top participation coefficient hubs were separately compared to a distribution of variant frequency in the HCP. This distribution was contrasted with a null overlap distribution produced from 1000 random rotations of a region of the same size as the hubs. None of these top participation coefficient hubs deviated from what would be expected from their null distribution (frequency of network variants at single hubs: 1.59% $\pm$ 1.63%; frequency of network variants with random rotations: left hemisphere = 4.40% $\pm$ 6.15%; right hemisphere: 4.15% $\pm$ 5.21%). The hub with the highest overlap with network variants was in the right superior caudal portion of the frontal lobe; this location had variants in 5.47% of people, still within the bounds of what would be expected by chance (70th percentile rotation). All other hubs were within a standard deviation of the null distribution's mean.

These findings show that group participation coefficient hubs do not frequently overlap with variants. This result is in line with the conserved hub hypothesis which states that the connectivity profiles of hubs are similar across individuals. In contrast, the malleable and artifactual hub hypotheses predict that group hubs should correspond with areas of high idiosyncratic functional connectivity.

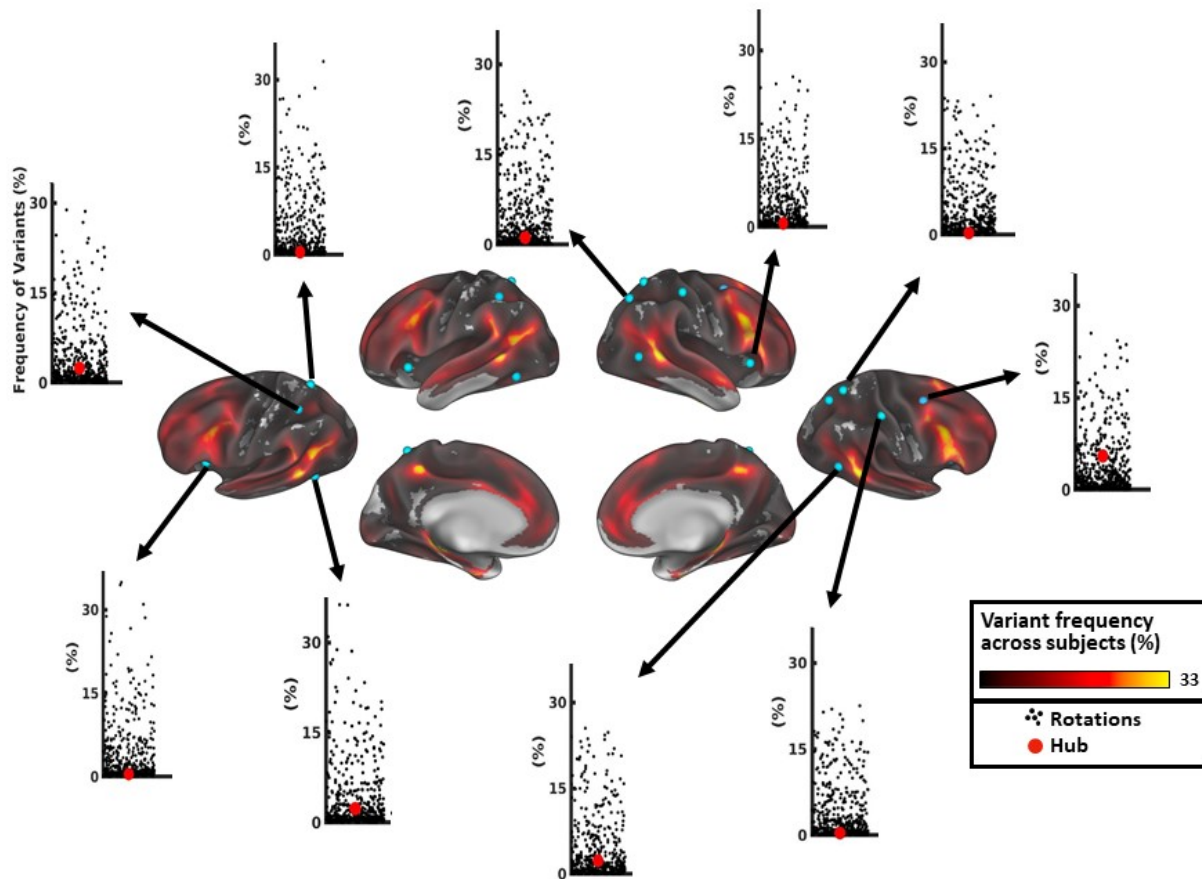


Figure 4: Network variant frequency at single participation coefficient hub locations. Each foci on the cortical surface represents a participation coefficient hub based on group-average data. The scatter plots capture the relationship between the true frequency of network variants at a given hub region (red dot) and the amount expected by random rotations (black dots). None of the hubs significantly differed from what would be expected from random rotations and all but one of the hubs (in the superior frontal cortex) were below the mean of the null distribution.

3.3 Community density hubs do overlap with locations of variability

In past work, community density has been proposed as an alternative measure of connector hubs (Power et al., 2013). This measure defines hubs based on their proximity to multiple different networks, under the assumption that regions at the intersection between networks are well situated to mediate cross-network interactions. However, as many locations of individual differences occur near the borders between networks (Dworetzky et al., 2021; Kraus et al., 2021; Seitzman et al., 2019) it is possible that this measure will show a greater correspondence with locations of inter-individual variability.

As before, we compared the locations of the top 10 group hubs, in this case defined based on community density in the same large sample of healthy young adults used in previous work (Power et al., 2013), with the map of the frequency of variants in the HCP dataset (Seitzman et al., 2019). As depicted in Figure 5A, the top community density hubs are found in somewhat

similar locations to participation coefficient hubs, but more ventrally in the anterior insula, along the superior frontal cortex, and near the temporoparietal junction. As a whole, community density hubs overlapped with network variant locations significantly more frequently than what would be expected by chance, as assessed with random rotations (Fig. 5B; 10.23% of people had variants at community density hubs, at the 99th percentile of random rotations, 95% CI [1.27%, 8.86%]). This result was replicated in the MSC dataset (Supp. Fig. 1B; including in a second set of analyses omitting global signal regression, Supp. Figs. 7 & 8). These findings suggest that community density hubs from group-average data often overlap with areas of idiosyncratic functional connectivity, suggesting they may be identifying malleable (Fig. 1B) or artifactual (Fig. 1C) hubs.

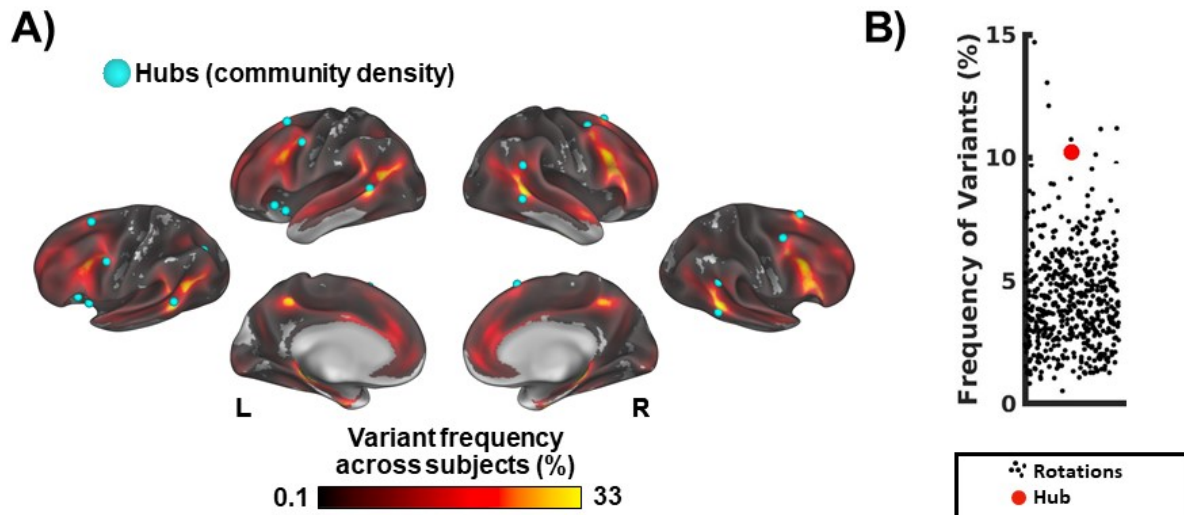


Figure 5: Comparison between group-average community density hubs and locations of strong inter-individual variability. A) The locations of group hubs defined using the community density metric are represented as light blue foci on the cortical surface. The heat map displayed on the cortical surface captures the frequency of network variants across people (percentage of subjects with a variant in at a location) based on the HCP dataset, with warmer colors indicating greater variant density. B) A scatter plot comparing the true frequency of variants (red dot) for the community density hub set to random rotations of the hub set (black dots). The high degree of correspondence between variants and community density hubs suggests that these hubs might be malleable or even artifactual hubs.

As before, we examined these results in more detail by quantifying the overlap of specific community density hubs with network variants in the HCP dataset (Supp Fig. 2). On average, 9.61% of people had variants over a connector hub (standard deviation: 6.79%; range: 0.58%-20.44%). Two hubs, one in the right superior frontal cortex and one near the left temporoparietal junction stood out as having numerically the highest variant overlap, but did not reach significance individually. Hubs in the left lateral frontal cortex, right temporoparietal junction, and the right lateral frontal cortex also exhibited relatively high frequency of variants. Thus, community density hubs have a high general tendency to overlap with network variants.

3.4 Locations of participation coefficient hubs in individuals exhibit similar connectivity to the group-average

The previous analyses demonstrate that group-average hubs defined with the participation coefficient do not overlap strongly with network variants, areas of particularly strong individual deviation in functional connectivity. However, it is possible that more subtle forms of variation would be present at these hubs that are not captured by network variants. To

investigate this question, we examined continuous measures of similarity at participation coefficient hubs (see Section 2.6.2).

We measured the individual-to-group similarity of whole-brain functional connectivity for the top 10 participation coefficient hubs (Fig. 6). Participation coefficient hub locations generally had good spatial correlations with the group-average connectivity profile in both HCP ($r = .59 \pm .04$) and MSC participants ($r = .64 \pm .04$). Although the pattern of hub connectivity is generally consistent across subjects, there are some deviant cases like MSC09 hub 3 (which will be explored further in sections 3.5 and 3.6). Comparisons with random rotations confirmed that hub regions show similar correspondence to the group average as other regions of cortex (see Supp. Fig. 3 and Supp. Table 1). This suggests that participation coefficient hubs do not differ substantially across individuals, even in more subtle forms of variation.

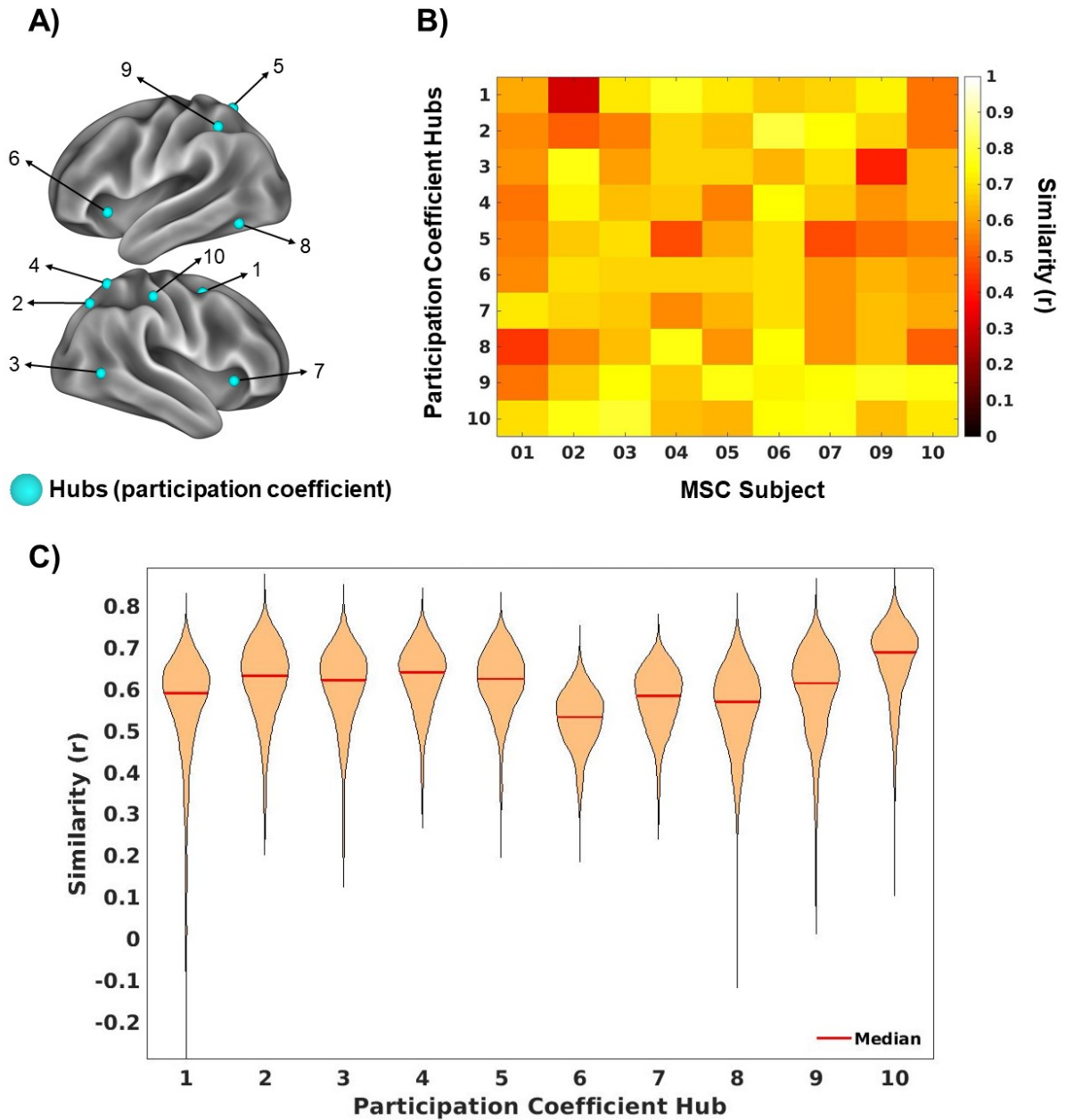


Figure 6: Continuous measures of individual-to-group similarity for participation coefficient hub locations. A) The location of group-average participation coefficient hubs are represented by the light blue circles, numbered for comparison to panels B and C. B) Each column of the x axis represents one of the analyzed MSC subjects and each row corresponds to one of the top participation coefficient hubs (corresponding number). The color scale represents the similarity of FC for a hub in each MSC individual relative to the group average, with warmer colors representing greater similarity. For the most part hubs are similar to the group average connectivity profile. (C) The same measures were calculated for the 752 HCP participants and are represented in a violin plot. The median similarity is marked by a red line.

3.5 Hubs with high participant coefficients also show good correspondence to the group using continuous measures

The main aim of this paper was to assess the inter-subject consistency of the top hub regions (the upper echelon of regions in terms of participation coefficient) defined at the group level. Regions with high participation coefficient are frequently treated as a special category in past work (Bertolero et al., 2015; Cole et al., 2013; Gratton et al., 2016; Power et al., 2013; Warren et al., 2014), emphasizing the utility of this approach for readers.

We have shown that the top 10 participation coefficient hubs do not exhibit low levels of similarity with the group average connectivity profile. Yet, this does not tell us about the relationship between the participation coefficient and similarity across the entire participation coefficient distribution. To this end, we conducted an additional analysis investigating the relationship between continuous participation coefficient measures and similarity to the group average connectivity profile. Specifically, the relationship between participation coefficient and similarity to the group average connectivity profile was examined for the cortical regions in Power et al., (2011) 264 ROIs in each MSC participant. The participation coefficient exhibited a weak negative correlation with similarity (Mean $r = -.22 \pm .08$; range $-.33$ to $-.09$) for most of the MSC subjects (see Figure 7).

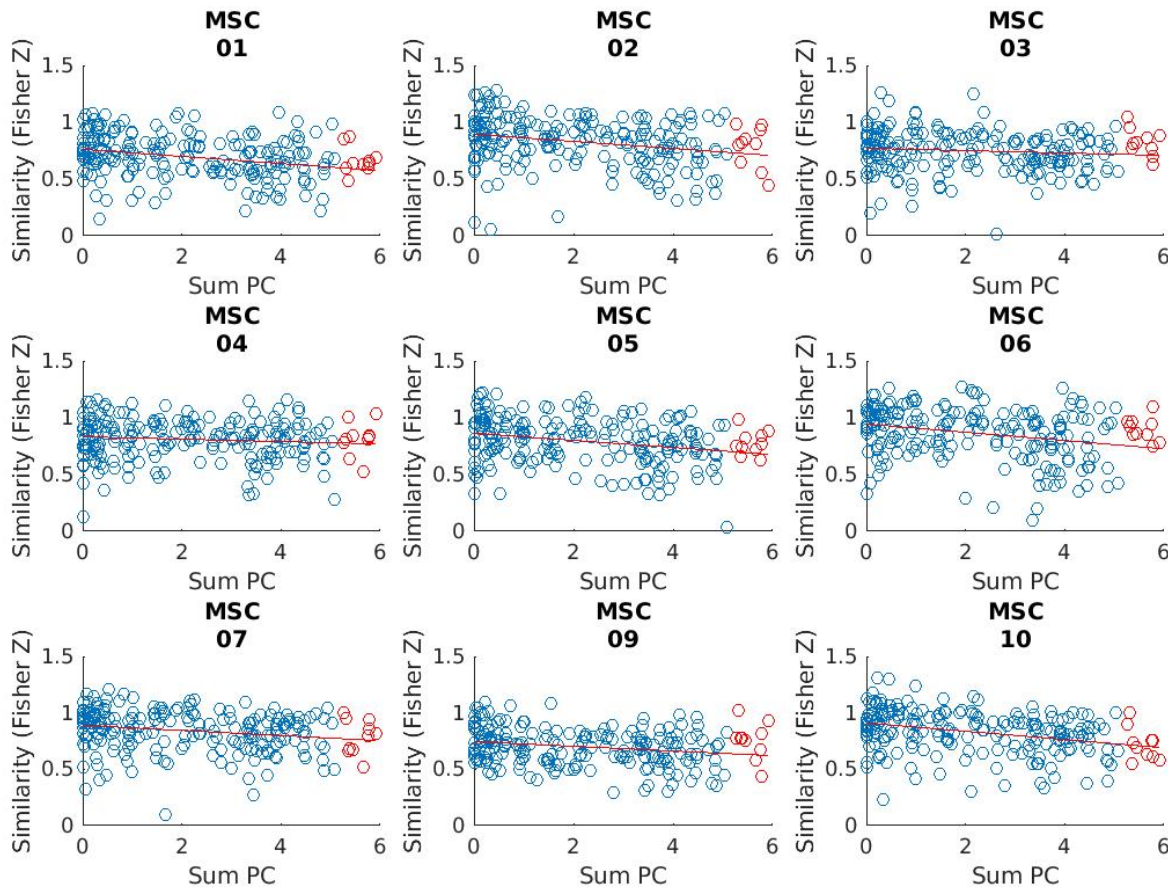


Figure 7: Continuous relationships between participation coefficient and group average similarity. For each MSC subject the participation coefficient (summed across thresholds) is on the x axis and the y axis represents the similarity (Fisher Z transformed correlation) of functional connectivity profile for a given region to the group average (similarity). The red line is the line of best fit and the red circles mark the top 10 nodes in terms of the participation

coefficient (hubs). All other nodes are represented by blue circles. As can be seen, relationships are generally slightly negative. However, the top 10 nodes in terms of the participation coefficient show good similarity to the group, in range of other regions throughout the brain. We examine the top 25% of participation coefficient regions in more detail in Fig. 8 to display those with relatively higher and lower similarity to the group average.

Visual inspection of the figure above suggests that the extremely high participation coefficient nodes (the targets of our primary analyses) have typical similarity to the group average reference (represented by the red circles; Mean $r = .64 \pm .04$). Figure 8 displays the top 25% of nodes in terms of participation coefficient and nodes are colored based on similarity to the group average connectivity profile. The lateral frontal cortex seems to disproportionately contain high participation coefficient nodes with relatively low similarity to the group, and should therefore be considered with caution when identified based on group-average maps. Nevertheless, the top 10 hubs demonstrate good consistency across participants.

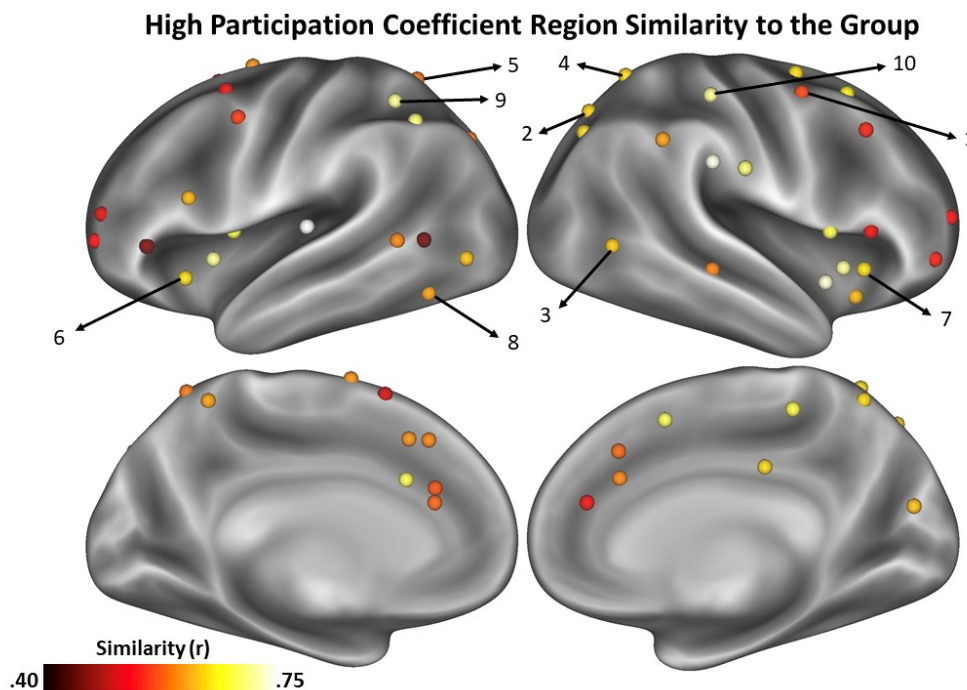


Figure 8: The top 25% of nodes in terms of sum participation coefficient. The color scale represents the cross MSC subject average similarity (cool to hot) as index by the correlation to the group average connectivity profile. The numbers denote hubs (top 10 nodes in terms of summed participation coefficient) nodes. Lateral frontal cortex nodes tended to have weaker similarity but most nodes exhibit robust similarity to the group average connectivity profile.

3.6 Characterizing hubs based on their cross-network profiles

The current findings suggest that hubs are largely similar in whole-brain functional connectivity across individuals, in the bounds of what would be expected for other regions of cortex. Next we examined which networks each hub was connected within individuals, to gain insights into the viability of the conserved vs. malleable hub hypotheses posed in the introduction. For each of the top 10 participation coefficient hubs in each MSC participant, we measured its connectivity to 14 canonical networks, creating a network profile for that region.

These profiles are shown for the top 10 participation coefficient hubs in Figure 9. Many hubs showed strong connectivity between 2 networks (1, 2, 5, 8), while others appeared to connect with a broader set (e.g., hub 3, 10). The network profiles were generally consistent across the

MSC participants and the group average, in accordance with the conserved hub hypothesis. However, some exceptions were present (e.g., MSC02 for Hub 1). This strong correspondence is reflected in the intraclass class correlation coefficients (ICC) for the hubs (cross-hub mean ICC .80 +/- .06) which can be found in S Table 3 (Shrout and Fleiss, 1979).

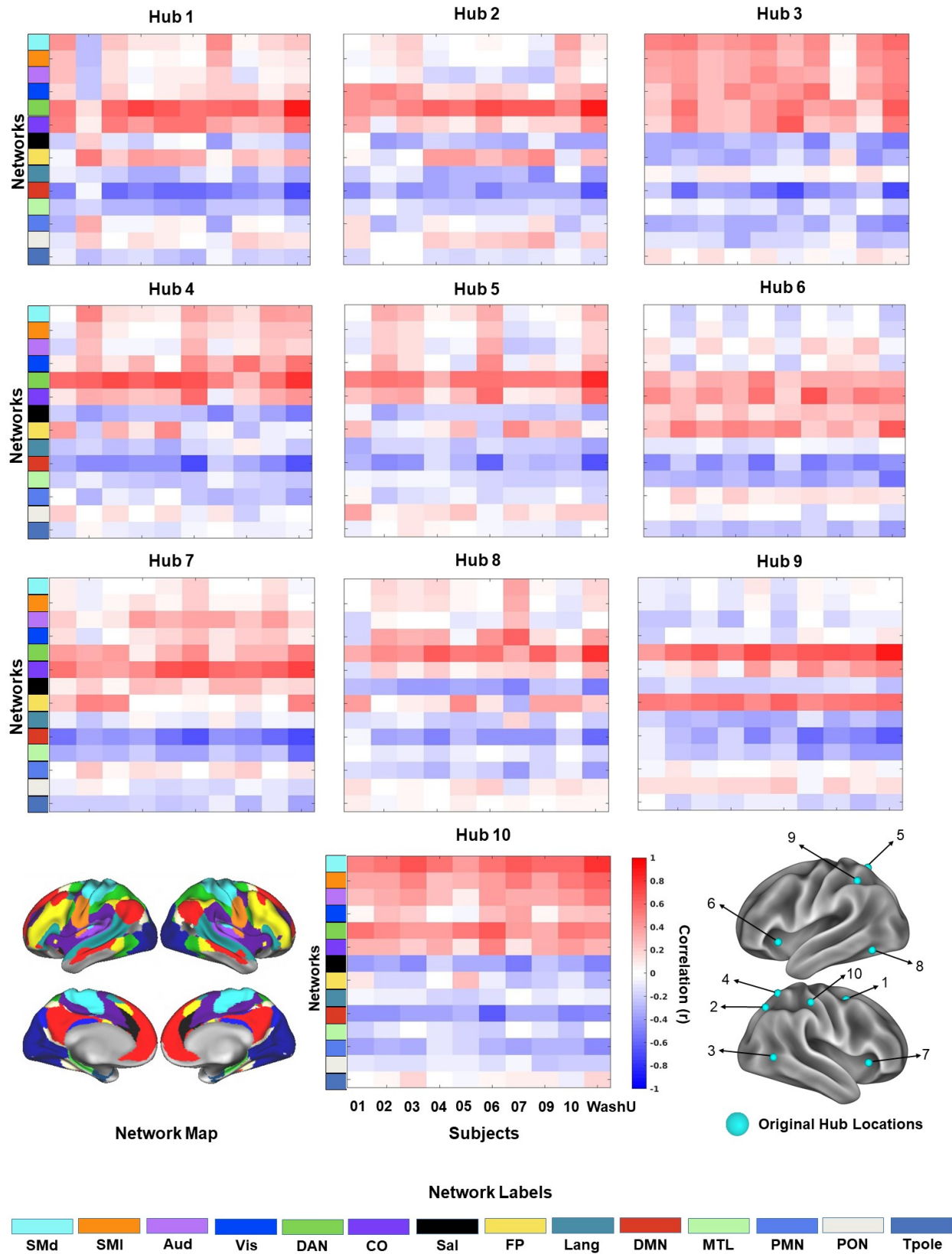


Figure 9: Network profiles for group-average participation coefficient hubs. For each hub (color map sub-plots), we show the network profile for the 9 MSC participants (columns) and WashU-120 group average (final column) to 14 canonical networks (rows; see *Methods*). The hub locations are shown in the bottom right corner and the canonical network maps are shown in the bottom left corner (each network is represented by a color).

3.7 Hub locations improve in correspondence if adjusted slightly in position

We observed some exceptions to the conservation in network profiles (Fig. 9) and similarity (Fig. 6) across participants (e.g., MSC02 for Hub 1, MSC09 for Hub 3). We next asked whether these exceptions could be ameliorated if hub locations were allowed to shift slightly in location between individuals. In order to investigate this possibility, we developed a spotlight procedure, in which hubs were adjusted slightly in position within each individual within a 10mm radius to find the location with the best matching network connectivity profile (see *Methods*; Fig. 10).

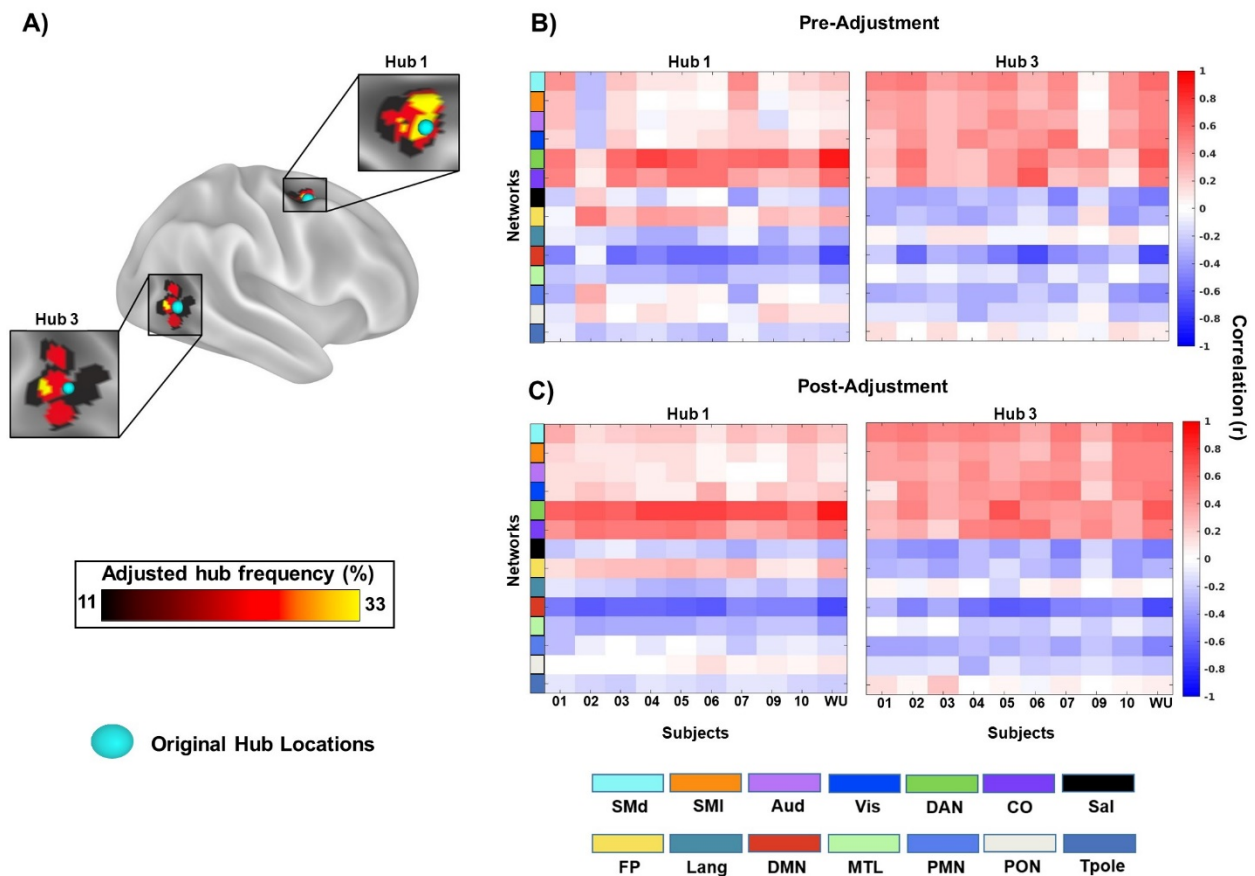


Figure 10: Hub local adjustment procedure. Results of a spotlight search procedure for two representative example hubs (Hub 1 and Hub 3) show that for most subjects, hub like areas can be found within a tight zone around a group hub. A) The blue spheres represent the original locations for group hubs 1 and 3. The underlying color map depicts the final adjusted hub locations across the 9 individuals in the MSC (see Supp. Fig 3 for adjusted hub locations for all hubs). (B) The original network profiles are shown for hub 1 (left) and hub 3 (right). (C) The adjusted hub network profiles are shown for the same two hubs. Adjusting hub locations improved correspondence across participants, especially in exception cases (e.g., MSC02 for hub 1, MSC09 for hub 3).

In most cases regions can be found in the spotlight that have a fair degree of resemblance to the group-average network profile (see two representative example hubs in

Figure 10). In particular, note that the subjects that were previously exceptions to the general pattern (MSC02 for hub 1, MSC09 for hub 3) show higher correspondence with the group average network profile at a nearby location. The improvement in network profile correspondence was reflected in the ICCs (Original cross-hub mean ICC $.80 \pm .06$ vs. Adjusted cross-hub mean ICC $.94 \pm .01$) which are displayed in S Table 3 (original hubs) and S Table 4 (adjusted hubs).

The adjusted hub network profiles are shown for all hubs below in Figure 11. In comparing Figure 9 to Figure 11, note the enhanced consistency in network profile for hubs across individual participants. This enhanced correspondence argues for group hubs showing conserved connectivity profiles across individuals, if you allow for a relatively minor local adjustment in location.

With this enhanced correspondence, it is also clearer that adjusted hubs seem to fall into one of two patterns. The first pattern is marked by the hub exhibiting a high degree of similarity with two to three networks. This pattern characterizes hubs 1, 2, 5, 8, and 9 (see Fig 10C & Fig 11). For example, hub 1 primarily bridges the dorsal attention and cinguloopercular networks, with weaker links to the frontoparietal and somatomotor systems. The other pattern is a high degree of similarity with a wide set of networks (see Fig 10C & Fig 11). For example, Hub 3 shows strong links to a range of sensorimotor and control networks.

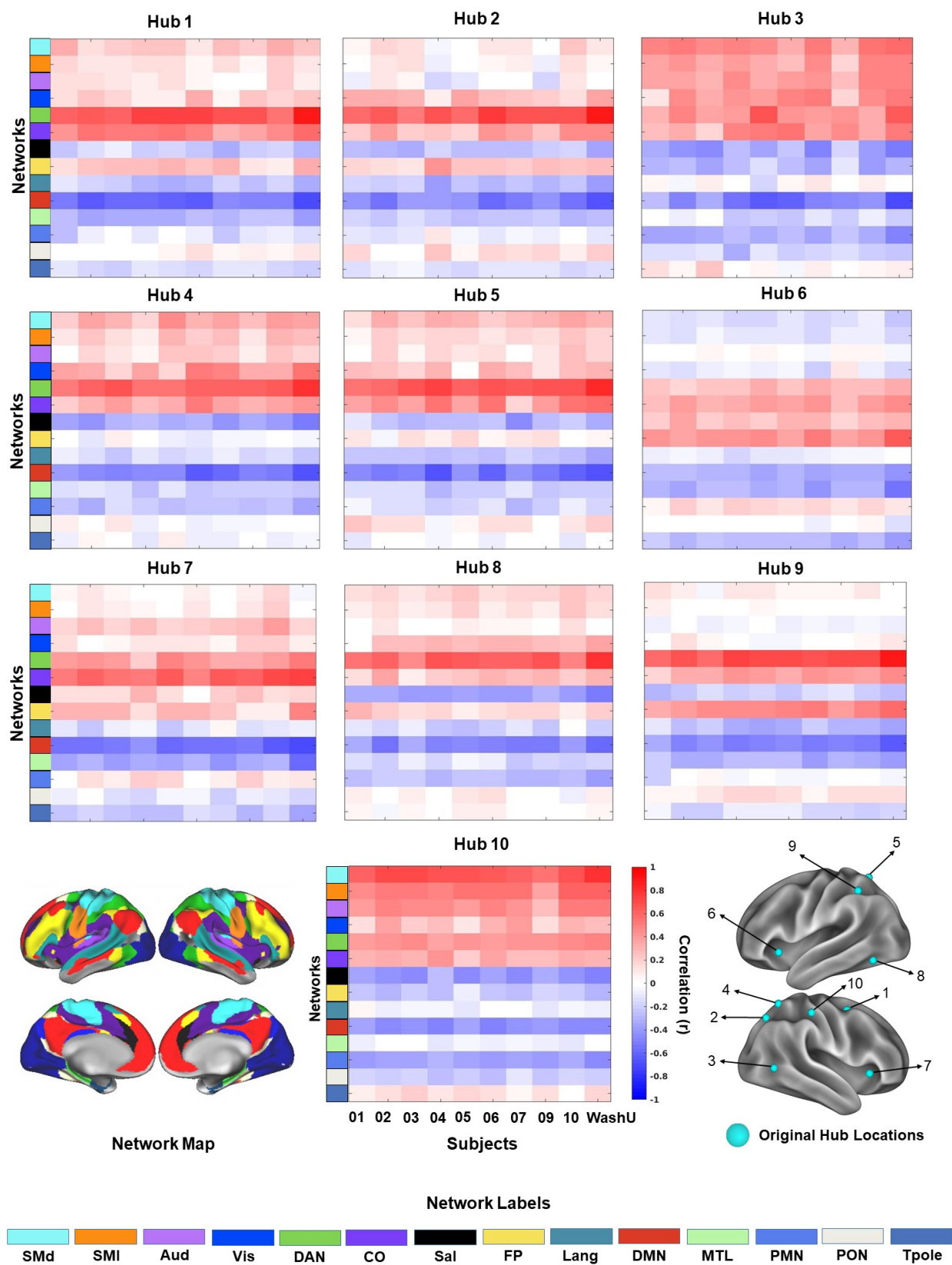


Figure 11: Network profiles for group hubs after local adjustment. This figure is similar to Figure 9, but shows the network profiles for each of the 10 top participation coefficient hubs after those hubs have been slightly adjusted in location using our spotlight analysis procedure (see *Methods* and Supp. Fig. 3 for final adjusted locations). For each hub, we show the network profile for the 9 MSC participants (columns) and WashU-120 group average (final column) to 14 canonical networks (rows). The original hub locations are reproduced in the inset on the lower left, and canonical networks are shown in the lower right.

3.8 Hub location adjustment procedure yields consistent results

As would be expected, the spotlight search procedure finds locations with profiles that better resemble the group average hub profile. We next sought to provide a quantitative assessment of the robustness of this procedure using independent samples of data. To this end, we split each MSC subject's data into two independent halves (odd and even sessions). The spotlight procedure was applied to one half (e.g., odd sessions). Then, the similarity of the spotlight location to the WashU-120 reference set was tested in the left out independent sessions (e.g., even set). The procedure was then repeated in the opposing direction. Supplemental Table 2 shows the outcomes. Applying the spotlight procedure improved the similarity of hub locations to the group-average pattern (i.e. decreased the dissimilarity as measured by a Euclidean distance measure), even when this approach was applied to independent data to get an unbiased estimate of improvements (Odd Original vs. Outset Dissimilarity Mean = .74 +/- .08 vs. Mean = .52 +/- .07; Even Original vs. Outset Dissimilarity Mean = .73 +/- .06 vs. Mean = .51 +/- .07). Indeed, the dissimilarity of the out-of-set and in-set minimums were highly correlated (Odd Set $r = .985$; Even Set $r = .998$). These results suggest that the spotlight procedure leads to a consistent improvement in correspondence of hubs to the group average.

4: Discussion

Our goal in this work was to determine how the top group-defined hubs vary across individuals. We hypothesized that hubs, which are thought to be regions critical to many important functions, should have connectivity profiles that are relatively conserved across individuals. However, we also considered the alternatives that (1) group hubs, as regions with diverse connectivity across networks may also show malleability across subjects or (2) that group hubs may be artifactual, driven by variation in functional networks within single individuals both of which would be associated with high correspondence between inter-subject connectivity profile variation and connector hub location. We demonstrated that the top group hubs, defined with the participation coefficient, do not overlap with locations of idiosyncratic functional connectivity, show relatively good correspondence in functional connectivity and network profiles across participants. Further improvements in network profile correspondence to the group is observed if, for each individual, these hubs are adjusted slightly in location. These findings are most consistent with the idea that the top group hubs have connectivity profiles that are relatively conserved across individuals, to a similar degree seen for many other cortical locations. More caution is warranted with alternative hub measures, such as the proximity-based community density metric, which tended to overlap with areas of idiosyncratic functional connectivity. Caution is also warranted for intermediate participation coefficient hubs, which both show a higher correspondence to regions of more inter-individual variability. This link between community density hubs and locations of inter-individual variability may stem from "border shifts" in functional brain networks between individuals (see (Dworetzky et al., 2021)). Overall, the findings lend support to the idea that group hubs (defined with the participation coefficient) have relatively consistent functional connectivity characteristics across people, and this suggests that these are likely not sites of particularly high malleability or artifacts. We close by discussing the

impact of this work on future studies with group hubs, and the potential for new findings in investigations focusing on individual-level hubs.

4.1 The connectivity profiles of group hubs are relatively similar across people

The participation coefficient is a measure of connection diversity, measuring the extent to which a given node shows connections across multiple distributed networks. Many investigators (Bertolero et al., 2018; Bertolero et al., 2015; Cole et al., 2013; Gordon et al., 2018; Gratton et al., 2016; Gratton et al., 2012; Power et al., 2013; Warren et al., 2014) have identified brain connector hubs using this metric, often based on group-average data (Power et al., 2013). Although these regions are defined by their diverse connection profile and are located in association regions of the brain (which show high inter-subject variation; Kong et al., 2019; Mueller et al., 2013; Seitzman et al., 2019), here we show that they do not have a strong correspondence to diversity *across people*: that is, diversity in functional connectivity across networks does not track with diversity in connectivity across individuals. Group-average hubs defined with the participation coefficient do not frequently overlap with extreme variants (Fig 3), are relatively similar to the group average (Fig 6), and have a fairly consistent network profile (Fig. 9; especially if these hubs are allowed to move slightly in position across people; see Fig 10-11). The lack of correspondence between group hubs and regions of high inter-subject variability in connectivity was observed in two datasets, and was true for both extreme deviations from the group average connectivity profile (variants) and more subtle deviations (similarity to the group).

This relative lack of variation is consistent with previous literature, which has shown that group-average estimates of hubs serve as good priors for locations that have outsized impacts on network structure and behavior if damaged in lesion patients (Gratton et al., 2012; Warren et al., 2014). We and others have also proposed that connector hubs (defined by the participation coefficient) play a critical role during task execution (Bertolero et al., 2018; Cole et al., 2013; Gratton et al., 2016). Hub regions are activated across a range of cognitive processes (Bertolero et al., 2015) and hub connectivity is modulated by task context (Cole et al., 2013; Gratton et al., 2016). Moreover, some evidence shows that participation coefficient hubs promote network modularity during task performance by tuning the connectivity of their neighbors (Bertolero et al., 2018). This body of work suggests that hubs may help coordinate activity between networks as needed in task control. This is a relatively essential set of functions to everyday life, likely depending on the ability to enact specific patterns of cross-network connectivity, and thus requiring some degree of uniformity in connector hub profiles across subjects.

However, although not *more* variable than would be expected by chance, participation coefficient hubs were also not generally below the expected variation for the cortex. The lack of a significant difference was observed across two datasets (HCP and MSC) and across multiple investigative approaches focusing on both extreme deviations and continuous gauges of similarity. Indeed, although recent work has highlighted strong variation between individuals (Bijsterbosch et al., 2018; Finn et al., 2015; Gordon et al., 2017a; Gratton et al., 2018a; Kong et al., 2019; Miranda-Dominguez et al., 2014; Mueller et al., 2013), much of this variation is relatively punctate and restricted to particular locations in a particular person (Seitzman et al., 2019), leading many places in the cortex to show good correspondence to the group average. Thus, the connectivity profiles of group hubs may be relatively conserved, but not more so than many other regions of the brain.

When the entire participation coefficient distribution is examined, a weak negative relationship between participation coefficient and similarity is observed. This relationship appears largely driven by a set of regions with intermediate participation coefficient value and low similarity to the group average. Regions in the top 25% of participation coefficient values

typically exhibiting robust similarity to the group average connectivity profile. Thus, more caution is necessitated when evaluating these more intermediate regions.

In conclusion, while these results argue that the top group (participation coefficient) hubs are not especially malleable or driven by artifacts, it is likely not correct to interpret them as particularly more conserved in their connectivity profiles than other cortical regions. Further, caution is still warranted in making strong conclusions about the conserved nature of putative hubs outside of the top 10 deeply investigated in this work. Our spotlight analysis hints that improved correspondence between hubs may be possible with methods that respect and/or address individual variability (e.g., see section 3.6 – Fig 9 and section 3.7 – Fig 11).

4.2 Limitations and future directions

We have shown that the top hubs defined with group-average data have connectivity profiles that are relatively conserved. Yet, it is still not clear if these top group-average hubs identify all (or the best) hubs within single individuals. The results of our spotlight analysis (Fig. 9-11) suggest that, at a minimum, researchers should consider exploring the areas surrounding group-average hub locations to ameliorate individual differences.

A limitation of the current study is that we focus on variability in previously identified group-average hubs, rather than hubs defined in individuals. Even when group hubs are reproducible across individuals, they could miss important features of functional neuroanatomy that can only be observed when looking at individually-defined hubs. For example individual level hubs might be needed to identify hub sub-types, which may help uncover more refined properties associated with cognition (Gordon et al., 2018).

Rigorous testing of the three scenarios outlined in the introduction will require an investigation of individual-level hubs. It might be the case that group hubs (for the most part) detect hub locations that are conserved across individuals, but that other hubs exist within each person that are more malleable, fitting scenario 2, and difficult to capture in group maps. If these idiosyncratic hubs are common, then the degree of cross-subject overlap in the locations of individual-hubs could be significantly dampened. However, measuring individual-level hubs robustly is challenging, as the participation coefficient calculated on individual subject data exhibits modest reliability even with 40 min. of data (Gordon et al., 2017a; Gordon et al., 2017b). Improving on these methods to identify reliable individual-level hubs and understand their characteristics and correspondences across individuals should be a priority. This avenue of research will allow us to address several of the outstanding questions left with the current work. Charting the correspondence between individual and group hubs will be a challenging but important area for future research.

For example, procedures that adopt individual-level connectivity maps, like the spotlight analysis we employed, could be further refined to help answer these questions. The high reliability of precision fMRI data makes it especially well-suited for ensuring robust applications of these procedures (Gordon et al., 2017b). In addition, high resolution imaging made possible by 7T scanners would allow for greater confidence in the accuracy of individual-level maps through improved signal and removal of artifactual hubs caused by spatial blurring or close proximity (Braga et al., 2019; Viessmann and Polimeni, 2021). If techniques such as these can improve hub identification in individuals they may help to further chart correspondence between the location of hubs defined on individual subject data and hubs defined at the group level, which do not always show very tight correspondence (Gordon et al., 2018).

Regardless of this study's limitations, we demonstrate correspondence in individual connectivity profiles across group-level hub locations (which is further improved at slightly adjusted locations, based on the spotlight procedure). The results of this analysis imply that hubs defined at the group level are not particularly malleable regions and are probably best

characterized as hubs at the individual level. Future research should elucidate the many questions regarding the correspondence between individual level and group level hubs.

4.3 Variants are not strong group-level hubs

Thus far, we have framed our results and discussion in terms of what our findings reveal for hubs. However, our findings also provide insights into variant locations. As locations of idiosyncratic functional connectivity, it is possible that variants are manifestations of flexible bridges between networks, which drives their inconsistency in their network membership across subjects (Vázquez-Rodríguez et al., 2019; Zhang et al., 2016). Yet, we did not find a clear correspondence between variants and participation coefficient hubs across several analyses in two datasets. While there is a correspondence between variants and community density hubs, this may be artifactual, due to averaging across people, rather than a manifestation of variants arising from a hub-like nature. These results suggest that, as a general rule, variants do not overlap with sites of group level connector hubs, implying that variants do not have a functional role as chameleon-like integrative regions. However, it is still possible that specific variants may show overlap with individual-level hubs that are not captured with this group approach. Future work will be needed to further investigate the properties of variants and their correspondence to individual-level hubs.

4.4 Community density is a hub metric that should be used with caution in group-level analyses

Unlike hubs defined by the participation coefficient, community density hubs defined in group-average data had a relatively strong correspondence with idiosyncratic variant regions. Community density hubs are defined based on their proximity to multiple different networks (Power et al., 2013), following the assumption that hubs should be positioned spatially intermediate to the systems that they bridge. However, although distributed somewhat similarly to participation coefficient hubs (Power et al., 2013), community density hubs in the group average are found more ventrally in the anterior insula, along the superior frontal cortex, and near the temporoparietal junction – locations in closer accord with locations of inter-subject variation (Fig 5).

This correspondence with regions of inter-subject variation suggests that group hubs defined with community density represent either malleable regions or artifactual hubs caused by mixed signals across individuals. Individual differences in brain networks often occur near the boundaries between systems (Dworetzky et al., 2021; Seitzman et al., 2019). As community density hubs sit at a nexus of multiple networks, then they may be more vulnerable to contamination from these shifting boundaries and the appearance of artifactual hubs. Future work using community density hubs should be cautious of this potential correspondence, especially if the hubs are identified with group average data. A solution to this issue may be to identify community density hubs within individual subject data instead, where hub inflation caused by inter-subject variation can be ruled out. In those cases, it may be possible to identify better representations of community density hubs that help to mediate interactions between networks based on their proximity.

4.5 Conclusion

Our findings from two independent datasets suggest that group-average brain hubs are relatively consistent in their functional topography across individuals. These findings suggest that hub locations may be a relatively conserved property of brain networks, perhaps due to the

critical role these regions have been proposed to have in cognition and brain function. Alternative metrics for hubs based on spatial proximity, such as community density, can overlap more strongly with locations of individual variability and should be used with more caution in the group average.

References

- Bertolero, M.A., Yeo, B.T., Bassett, D.S., D'Esposito, M., 2018. A mechanistic model of connector hubs, modularity and cognition. *Nature human behaviour* 2, 765-777.
- Bertolero, M.A., Yeo, B.T., D'Esposito, M., 2015. The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences* 112, E6798-E6807.
- Bertolero, M.A., Yeo, B.T., D'Esposito, M., 2017. The diverse club. *Nature communications* 8, 1-11.
- Bijsterbosch, J.D., Woolrich, M.W., Glasser, M.F., Robinson, E.C., Beckmann, C.F., Van Essen, D.C., Harrison, S.J., Smith, S.M., 2018. The relationship between spatial configuration and functional connectivity of brain regions. *Elife* 7, e32992.
- Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine* 34, 537-541.
- Braga, R.M., Van Dijk, K.R., Polimeni, J.R., Eldaief, M.C., Buckner, R.L., 2019. Parallel distributed networks resolved at high resolution reveal close juxtaposition of distinct regions. *Journal of neurophysiology* 121, 1513-1534.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10, 186-198.
- Burgess, G.C., Kandala, S., Nolan, D., Laumann, T.O., Power, J.D., Adeyemo, B., Harms, M.P., Petersen, S.E., Barch, D.M., 2016. Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the human connectome project. *Brain connectivity* 6, 669-680.
- Ciric, R., Wolf, D.H., Power, J.D., Roalf, D.R., Baum, G.L., Ruparel, K., Shinohara, R.T., Elliott, M.A., Eickhoff, S.B., Davatzikos, C., 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage* 154, 174-187.
- Cole, M.W., Reynolds, J.R., Power, J.D., Repovs, G., Anticevic, A., Braver, T.S., 2013. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat Neurosci* 16, 1348-1355.
- Dale, A.M., 1999. Optimal experimental design for event-related fMRI. *Human brain mapping* 8, 109-114.
- Dosenbach, N.U., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., 2007. Distinct brain networks

for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences* 104, 11073-11078.

Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen, S.E., 2006. A core system for the implementation of task sets. *Neuron* 50, 799-812.

Dworetsky, A., Seitzman, B.A., Adeyemo, B., Smith, D.M., Petersen, S.E., Gratton, C., 2021. Two common and distinct forms of variation in human functional brain networks. *bioRxiv*, 2021.2009.2017.460799.

Fair, D.A., Miranda-Dominguez, O., Snyder, A.Z., Perrone, A., Earl, E.A., Van, A.N., Koller, J.M., Feczko, E., Tisdall, M.D., van der Kouwe, A., 2020. Correction of respiratory artifacts in MRI head motion estimates. *Neuroimage* 208, 116400.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience* 18, 1664-1671.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341-355.

Fransson, P., Thompson, W.H., 2020. Temporal flow of hubs and connectivity in the human brain. *Neuroimage* 223, 117348.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. *Neuroimage* 7, 30-40.

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* 80, 105-124.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Gilmore, A.W., Nelson, S.M., Dosenbach, N.U., Petersen, S.E., 2017a. Individual-specific features of brain systems identified with resting state functional correlations. *Neuroimage* 146, 918-939.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2016. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral cortex* 26, 288-303.

Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., 2017b. Precision functional mapping of individual human brains. *Neuron* 95, 791-807. e797.

Gordon, E.M., Lynch, C.J., Gratton, C., Laumann, T.O., Gilmore, A.W., Greene, D.J., Ortega, M., Nguyen, A.L., Schlaggar, B.L., Petersen, S.E., Dosenbach, N.U.F., Nelson, S.M., 2018. Three Distinct Sets of Connector Hubs Integrate Human Brain Function. *Cell Rep* 24, 1687-1695.e1684.

Gratton, C., Dworetsky, A., Coalson, R.S., Adeyemo, B., Laumann, T.O., Wig, G.S., Kong, T.S., Gratton, G., Fabiani, M., Barch, D.M., 2020. Removal of high frequency contamination from motion estimates in single-band fMRI saves data without biasing functional connectivity. *Neuroimage* 217, 116866.

Gratton, C., Laumann, T.O., Gordon, E.M., Adeyemo, B., Petersen, S.E., 2016. Evidence for Two Independent Factors that Modify Brain Networks to Meet Task Goals. *Cell Rep* 17, 1276-1288.

Gratton, C., Laumann, T.O., Nielsen, A.N., Greene, D.J., Gordon, E.M., Gilmore, A.W., Nelson, S.M., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., 2018a. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron* 98, 439-452. e435.

Gratton, C., Nomura, E.M., Pérez, F., D'Esposito, M., 2012. Focal brain lesions to critical locations cause widespread disruption of the modular organization of the brain. *J Cogn Neurosci* 24, 1275-1285.

Gratton, C., Sun, H., Petersen, S.E., 2018b. Control networks and hubs. *Psychophysiology* 55.

Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences* 100, 253-258.

Guimerà, R., Nunes Amaral, L.A., 2005. Functional cartography of complex metabolic networks. *Nature* 433, 895-900.

Kong, R., Li, J., Orban, C., Sabuncu, M.R., Liu, H., Schaefer, A., Sun, N., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., 2019. Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cerebral cortex* 29, 2533-2551.

Kraus, B.T., Perez, D., Ladwig, Z., Seitzman, B.A., Dworetsky, A., Petersen, S.E., Gratton, C., 2021. Network variants are similar between task and rest states. *Neuroimage* 229, 117743.

Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. *Human brain mapping* 10, 120-131.

Laumann, T.O., Gordon, E.M., Adeyemo, B., Snyder, A.Z., Joo, S.J., Chen, M.Y., Gilmore, A.W., McDermott, K.B., Nelson, S.M., Dosenbach, N.U., Schlaggar, B.L.,

Mumford, J.A., Poldrack, R.A., Petersen, S.E., 2015. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron* 87, 657-670.

Laumann, T.O., Snyder, A.Z., Mitra, A., Gordon, E.M., Gratton, C., Adeyemo, B., Gilmore, A.W., Nelson, S.M., Berg, J.J., Greene, D.J., McCarthy, J.E., Tagliazucchi, E., Laufs, H., Schlaggar, B.L., Dosenbach, N.U.F., Petersen, S.E., 2017. On the Stability of BOLD fMRI Correlations. *Cereb Cortex* 27, 4719-4732.

Liao, X.-H., Xia, M.-R., Xu, T., Dai, Z.-J., Cao, X.-Y., Niu, H.-J., Zuo, X.-N., Zang, Y.-F., He, Y., 2013. Functional brain hubs and their test–retest reliability: a multiband resting-state functional MRI study. *Neuroimage* 83, 969-982.

Marcus, D., Harwell, J., Olsen, T., Hodge, M., Glasser, M., Prior, F., Jenkinson, M., Laumann, T., Curtiss, S., Van Essen, D., 2011. Informatics and data mining tools and strategies for the human connectome project. *Frontiers in neuroinformatics* 5, 4.

Marcus, D.S., Harms, M.P., Snyder, A.Z., Jenkinson, M., Wilson, J.A., Glasser, M.F., Barch, D.M., Archie, K.A., Burgess, G.C., Ramaratnam, M., 2013. Human Connectome Project informatics: quality control, database services, and data visualization. *Neuroimage* 80, 202-219.

Miezin, F.M., Maccotta, L., Ollinger, J., Petersen, S., Buckner, R., 2000. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11, 735-759.

Miranda-Dominguez, O., Mills, B.D., Carpenter, S.D., Grant, K.A., Kroenke, C.D., Nigg, J.T., Fair, D.A., 2014. Connectotyping: model based fingerprinting of the functional connectome. *PloS one* 9, e111048.

Mueller, S., Wang, D., Fox, M.D., Yeo, B.T., Sepulcre, J., Sabuncu, M.R., Shafee, R., Lu, J., Liu, H., 2013. Individual variability in functional connectivity architecture of the human brain. *Neuron* 77, 586-595.

Ojemann, J.G., Akbudak, E., Snyder, A.Z., McKinstry, R.C., Raichle, M.E., Conturo, T.E., 1997. Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage* 6, 156-167.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142-2154.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., 2011. Functional network organization of the human brain. *Neuron* 72, 665-678.

Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* 84, 320-341.

- Power, J.D., Schlaggar, B.L., Lessov-Schlaggar, C.N., Petersen, S.E., 2013. Evidence for hubs in human functional brain networks. *Neuron* 79, 798-813.
- Power, J.D., Schlaggar, B.L., Petersen, S.E., 2015. Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage* 105, 536-551.
- Rosvall, M., Bergstrom, C.T., 2008. Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci U S A* 105, 1118-1123.
- Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* 64, 240-256.
- Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience* 27, 2349-2356.
- Seitzman, B.A., Gratton, C., Laumann, T.O., Gordon, E.M., Adeyemo, B., Dworetsky, A., Kraus, B.T., Gilmore, A.W., Berg, J.J., Ortega, M., Nguyen, A., Greene, D.J., McDermott, K.B., Nelson, S.M., Lessov-Schlaggar, C.N., Schlaggar, B.L., Dosenbach, N.U.F., Petersen, S.E., 2019. Trait-like variants in human functional brain networks. *Proc Natl Acad Sci U S A* 116, 22851-22861.
- Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin* 86, 420.
- Siegel, J.S., Mitra, A., Laumann, T.O., Seitzman, B.A., Raichle, M., Corbetta, M., Snyder, A.Z., 2017. Data quality influences observed links between functional connectivity and behavior. *Cerebral cortex* 27, 4492-4502.
- Smith, D.M., Perez, D.C., Porter, A., Dworetsky, A., Gratton, C., 2021. Light through the fog: using precision fMRI data to disentangle the neural substrates of cognitive control. *Current Opinion in Behavioral Sciences* 40, 19-26.
- Sporns, O., 2010. *Networks of the Brain*. MIT press.
- Thomas Yeo, B., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* 106, 1125-1165.
- Van Den Heuvel, M.P., Sporns, O., 2011. Rich-club organization of the human connectome. *Journal of Neuroscience* 31, 15775-15786.
- Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T.E., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., Della Penna, S., Feinberg, D., Glasser, M.F., Harel, N., Heath, A.C., Larson-Prior, L., Marcus, D., Michalareas, G.,

Moeller, S., Oostenveld, R., Petersen, S.E., Prior, F., Schlaggar, B.L., Smith, S.M., Snyder, A.Z., Xu, J., Yacoub, E., 2012. The Human Connectome Project: a data acquisition perspective. *Neuroimage* 62, 2222-2231.

Vázquez-Rodríguez, B., Suárez, L.E., Markello, R.D., Shafiei, G., Paquola, C., Hagmann, P., Van Den Heuvel, M.P., Bernhardt, B.C., Spreng, R.N., Misic, B., 2019. Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences* 116, 21219-21227.

Viessmann, O., Polimeni, J.R., 2021. High-resolution fMRI at 7 Tesla: challenges, promises and recent developments for individual-focused fMRI studies. *Current Opinion in Behavioral Sciences* 40, 96-104.

Warren, D.E., Power, J.D., Bruss, J., Denburg, N.L., Waldron, E.J., Sun, H., Petersen, S.E., Tranel, D., 2014. Network measures predict neuropsychological outcome after brain injury. *Proc Natl Acad Sci U S A* 111, 14247-14252.

Zhang, J., Cheng, W., Liu, Z., Zhang, K., Lei, X., Yao, Y., Becker, B., Liu, Y., Kendrick, K.M., Lu, G., 2016. Neural, electrophysiological and anatomical basis of brain-network variability and its characteristic changes in mental disorders. *Brain* 139, 2307-2321.

SUPPLEMENTAL MATERIAL

Supplemental Tables

Table S1: Similarity to the Group for MSC Subjects

MSC subject	Mean R to group	95% CI
MSC01	.57	[0.48, 0.66]
MSC02	.63	[0.57, 0.74]
MSC03	.67	[0.53, 0.69]
MSC04	.66	[0.56, 0.72]
MSC05	.65	[0.56, 0.72]
MSC06	.71	[0.57, 0.76]
MSC07	.65	[0.56, 0.73]
MSC09	.62	[0.53, 0.67]
MSC10	.61	[0.57, 0.73]

Supp. Table 1: For each MSC subject the mean correlation (across vertices included in the hub) is shown in the middle column. The third column contains the corresponding 95% confidence intervals.

Table S2: Split-Half Spotlight Analysis

Hub #	Original Dissimilarity (O)	Original Dissimilarity (E)	Inset Minimum Dissimilarity (O)	Inset Minimum Dissimilarity (E)	Outset Dissimilarity (O)	Outset Dissimilarity (E)
1	.75(.37)	.72(.33)	.47(.12)	.43(.12)	.49(.12)	.45(.12)
2	.77(.23)	.80(.19)	.52(.11)	.52(.14)	.54(.12)	.54(.14)
3	.78(.26)	.79(.30)	.62(.18)	.62(.18)	.66(.16)	.64(.18)
4	.82(.26)	.76(.25)	.53(.11)	.50(.13)	.55(.11)	.53(.12)
5	.89(.22)	.82(.23)	.45(.11)	.46(.13)	.50(.13)	.49(.13)
6	.65(.11)	.65(.11)	.51(.09)	.50(.09)	.55(.12)	.52(.11)
7	.75(.16)	.71(.13)	.55(.08)	.57(.14)	.58(.06)	.60(.14)
8	.74(.27)	.71(.25)	.44(.15)	.41(.11)	.46(.16)	.44(.12)
9	.66(.22)	.63(.25)	.44(.13)	.43(.14)	.46(.13)	.45(.14)
10	.61(.24)	.66(.20)	.35(.09)	.38(.09)	.40(.11)	.41(.09)

Supp. Table 2: Across subject mean (standard deviation) for key spotlight analysis metrics for odd and even sessions. The first two columns present the distance of the original hub location to the WashU-120 reference for each data half. The third and fourth columns present the dissimilarity (Euclidean distance) minimum in the spotlight. The fifth and sixth columns present the dissimilarity of outset's minimum location to the WashU-120 reference based on the inset's data (O=Odd, E=Even).

Table S3: Original Hub Intraclass Correlation Coefficients

Hub #	ICC	95% CI
1	.76	[0.60, 0.89]
2	.75	[0.59, 0.89]
3	.84	[0.72, 0.93]
4	.77	[0.62, 0.90]
5	.72	[0.55, 0.87]
6	.86	[0.74, 0.94]
7	.85	[0.73, 0.94]
8	.72	[0.55, 0.87]
9	.88	[0.78, 0.95]
10	.90	[0.81, 0.96]

Supp. Table 3: For each participation coefficient hub the intraclass correlation coefficient is shown in the middle column. The third column contains the corresponding 95% confidence intervals.

Table S4: Adjusted Hub Intraclass Correlation Coefficients

Hub #	ICC	95% CI
1	.95	[0.91, 0.98]
2	.93	[0.86, 0.97]
3	.92	[0.85, 0.97]
4	.94	[0.89, 0.98]
5	.94	[0.88, 0.98]
6	.94	[0.89, 0.98]
7	.94	[0.88, 0.97]
8	.94	[0.89, 0.98]
9	.96	[0.92, 0.98]
10	.97	[0.95, 0.99]

Supp. Table 4: For each participation coefficient hub the intraclass correlation coefficient is shown in the middle column. The third column contains the corresponding 95% confidence intervals.

Supplemental Figures

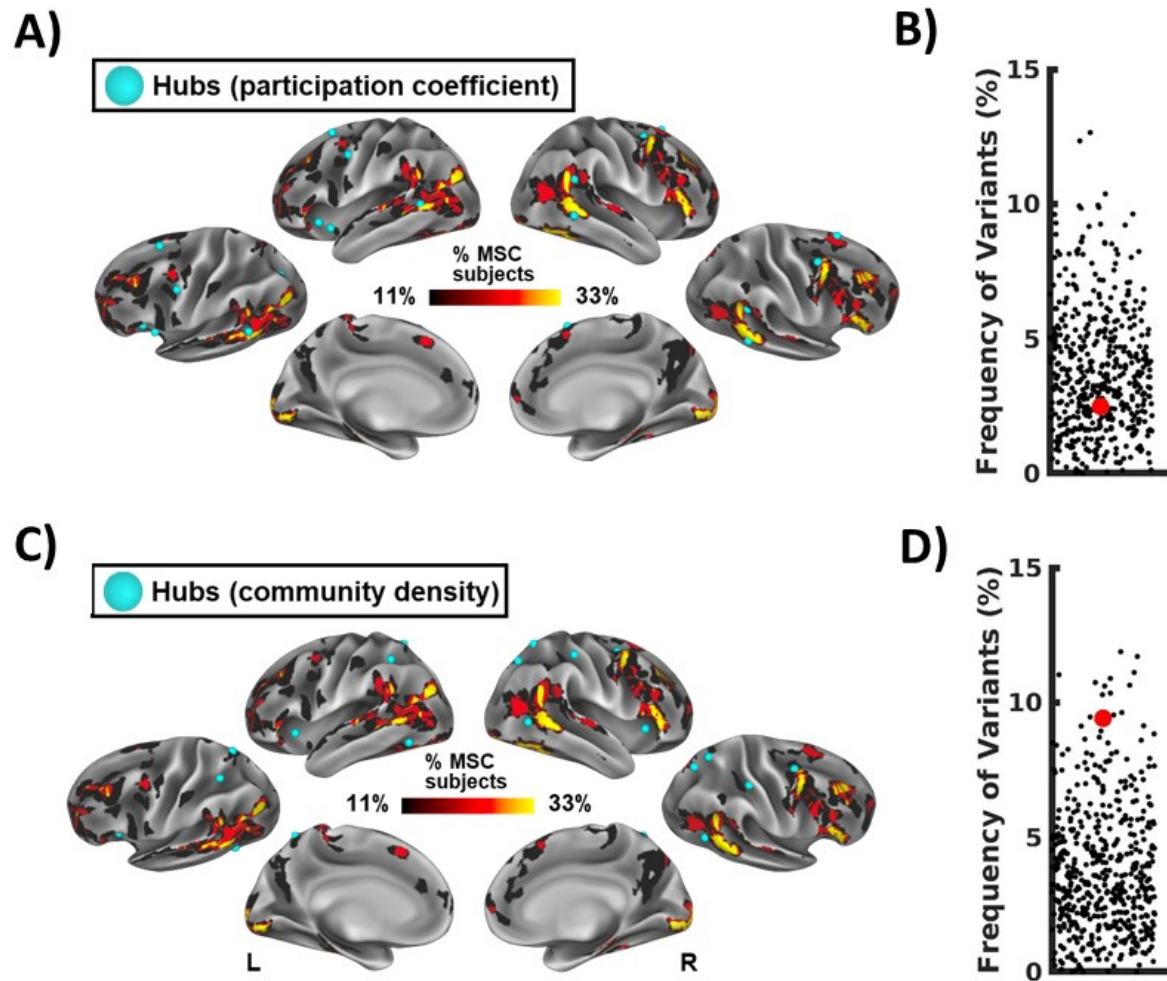


Figure S1: Locations of group hubs relative to the frequency of network variants in the Midnight Scan Club (MSC) dataset. A) This figure plots the locations of the participation coefficient hubs (light blue foci) overlapped on frequency of network variants in the MSC dataset. B) This figure displays the frequency of networks variants at participation coefficient hubs (red dot) relative to random rotations (black dots). In the MSC, 2.49% of people had network variants at participation coefficient hubs (33rd percentile of rotations, 95% CI [0.40%, 9.21%]). C) This figure is similar to A, but shows locations of community density hubs overlapped on the frequency of network variants in the MSC dataset. D) As in B, this figure shows the frequency of network variants at community density hubs (red dot) relative to random rotations (black dots). In the MSC, 9.40% of people had network variants at community density hub locations, corresponding to the 98th percentile of random rotations (95% CI [0.33%, 9.13%]).

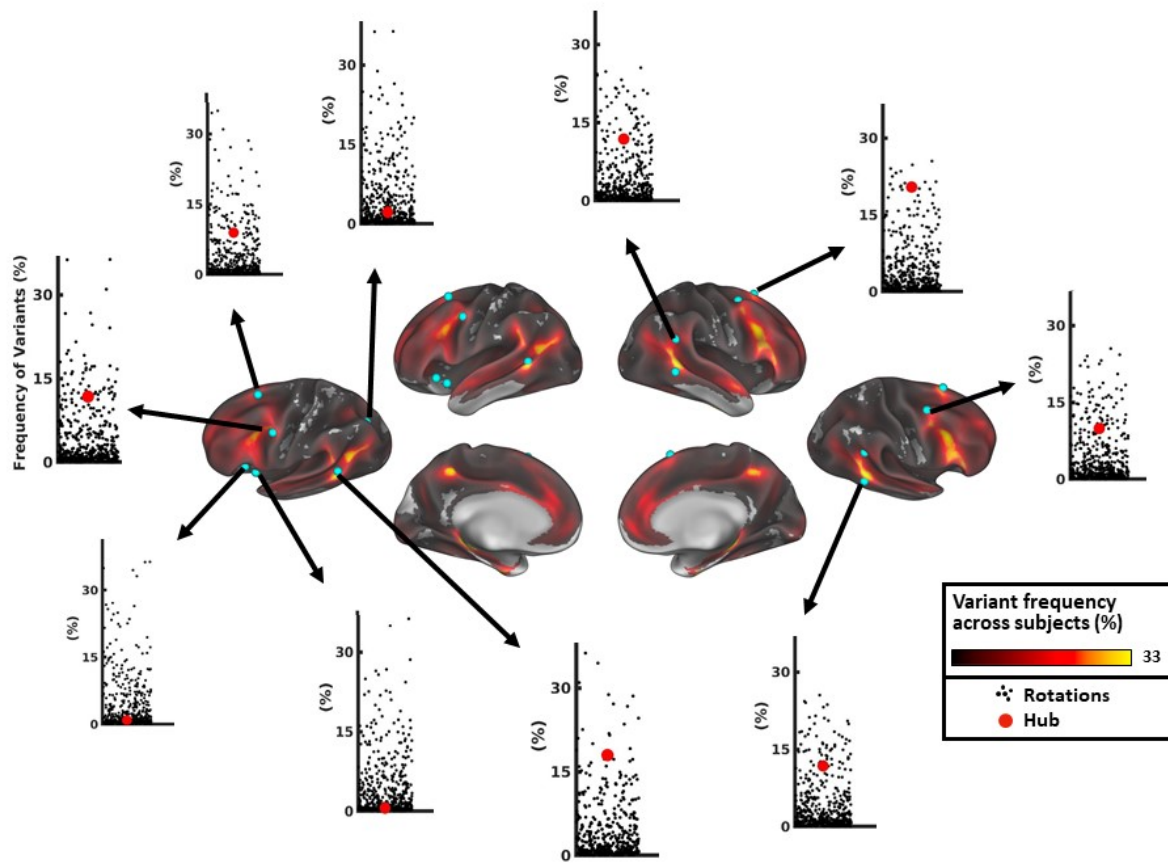


Figure S2: Network variant frequency at single community density hub locations. Single community density hub locations tend to exhibit a high degree of variant overlap. Each foci on the cortical surface represents a community density based hub. The scatter plots display the relationship between true frequency of variants in a hub region (red dot) and the amount expected by random rotations (black dots). All but three hubs fell above the mean of their respective hemisphere's null distribution. None of the hubs exhibited a statistically significant difference from the random locations.

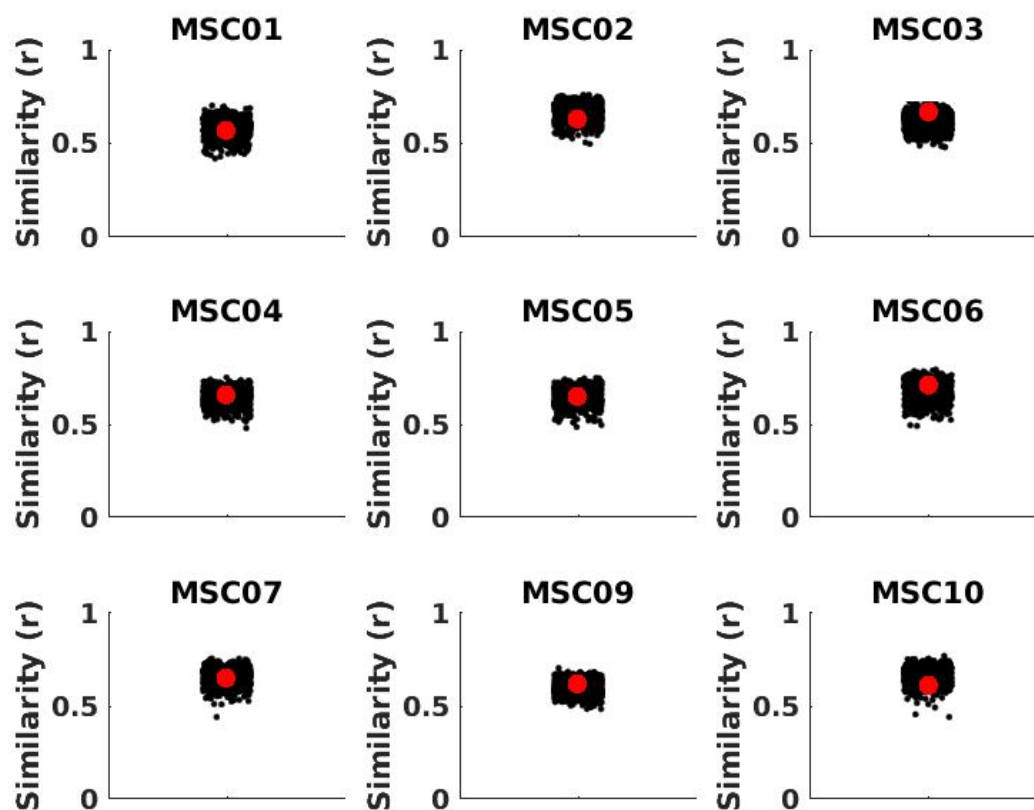


Figure S3: Similarity to the group average connectivity profile distributions for each MSC subject. For each MSC participant the similarity of hub locations to the group average connectivity profile is represented by the red dot and the smaller black dots represent the similarity for random rotations of the hub set. None of the participants significantly deviated from what would be expected by the null distribution (see Supp. Table 1 for confidence intervals). The average observed similarity to the group connectivity profile (average of the red dots) across participants was .64 with a standard deviation of .04.

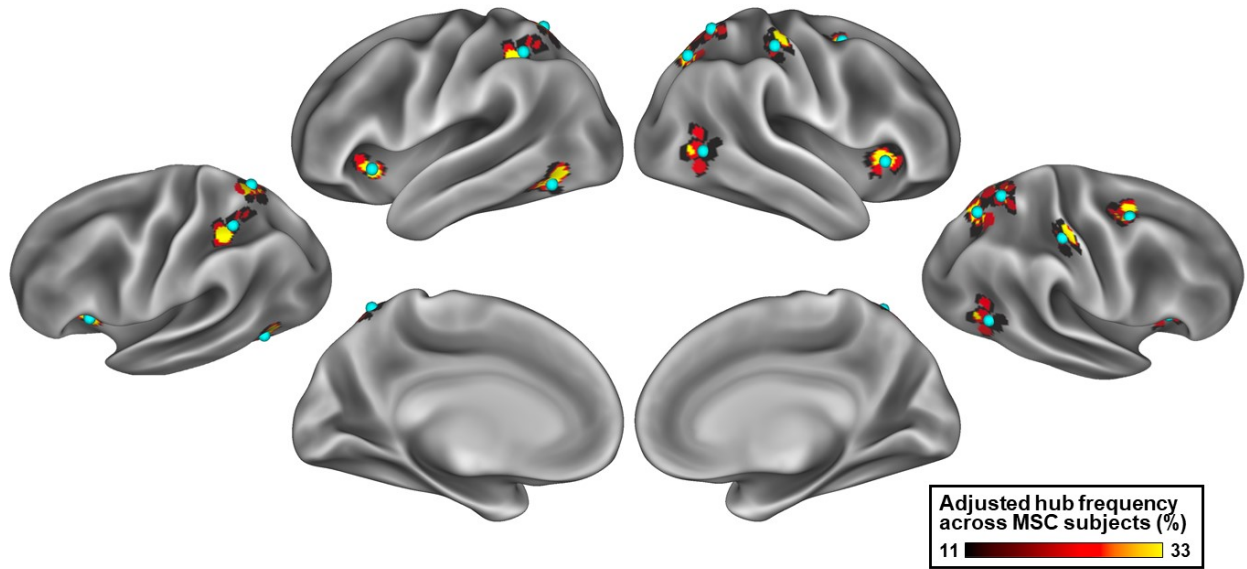


Figure S4: For each group-average participation coefficient hub (blue spheres) the frequency of adjusted hub locations is marked. The frequency of adjusted hub locations is defined as the percentage of MSC subjects with an adjusted hub overlapping with the given cortical location.

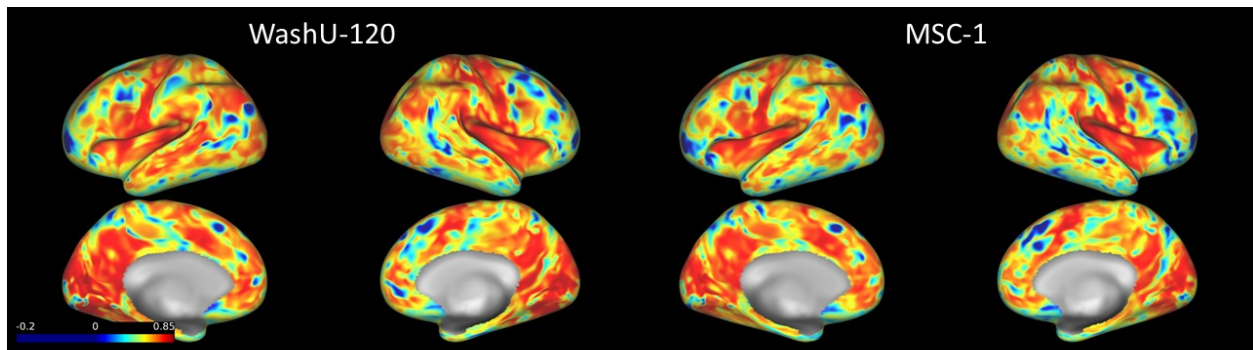
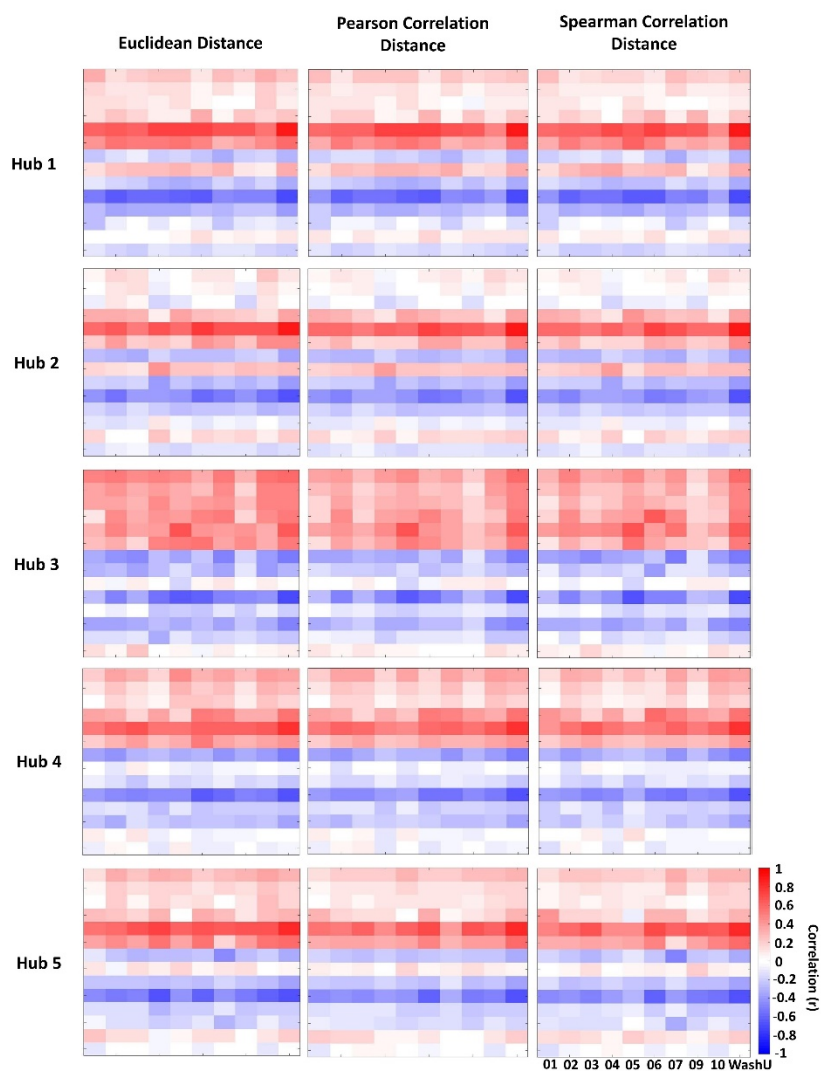


Figure S5: Reference dataset comparison. To gain insights into the influence of the reference dataset on an individual's similarity map, a reference dataset was produced for each of the MSC subjects (the average of all MSC subjects other than the given subject). A similarity map was produced by correlating the hold out MSC subject's connectivity profile with the MSC-1 average profile. Then the resulting similarity map was correlated with the given subject's Wash U 120 based similarity map. Across participants, correlations were high (Mean $r = .84$; SD = .02; Range .81-.87), regardless of reference set. A representative subject's similarity maps are presented above, showing MSC01's similarity map to the WashU-120 reference (left) and to the MSC participants' average, omitting that subject (MSC-1 reference = average of MSC02-MSC10; right). The spatial distributions of network variants are very similar across the reference sets. These findings suggest that test/reference dataset does not have a strong influence on the identification of network variants.

A



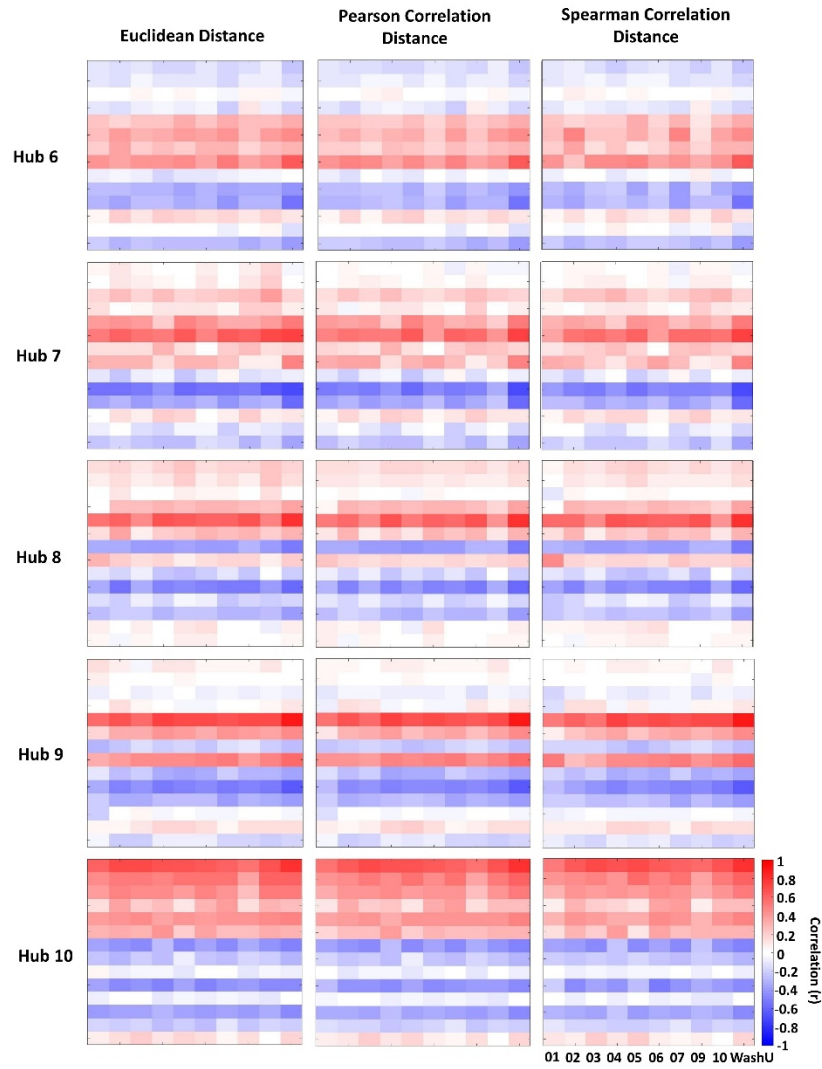
B

Figure S6: A comparison of the spotlight adjusted hub network connectivity profiles for the MSC dataset for the 10 group level hubs (A 1-5; B 6-10). The three distance metrics include Euclidean distance, 1-Pearson correlation, and 1-Spearman correlation.

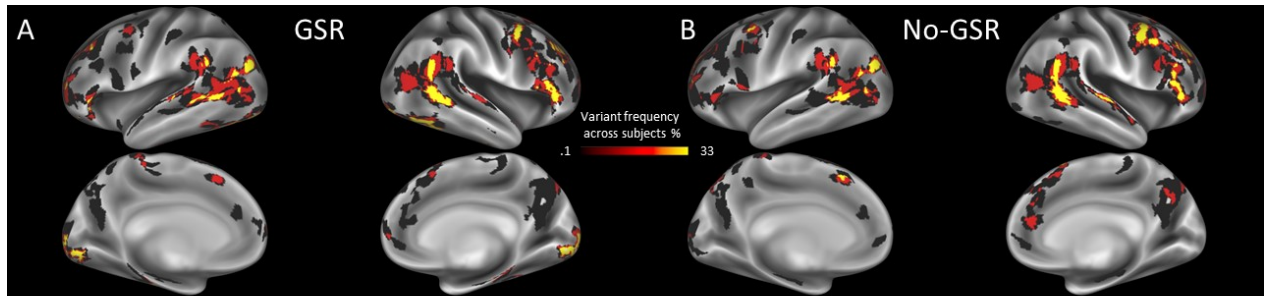


Figure S7: Network variants computed with and without global signal regression. The heat map displayed on the cortical surface captures the frequency of variants (percentage of subjects with a variant at a location) based on the MSC dataset, with warmer colors indicating greater variant frequency. The map of variant frequency on the left (A) is based on the output of our typical (GSR+) pipeline. On the right (B) is the variant density map for a pipeline that did not include GSR. These maps are similar ($r=.73$) indicating that GSR does not have a major effect on the variant landscape.

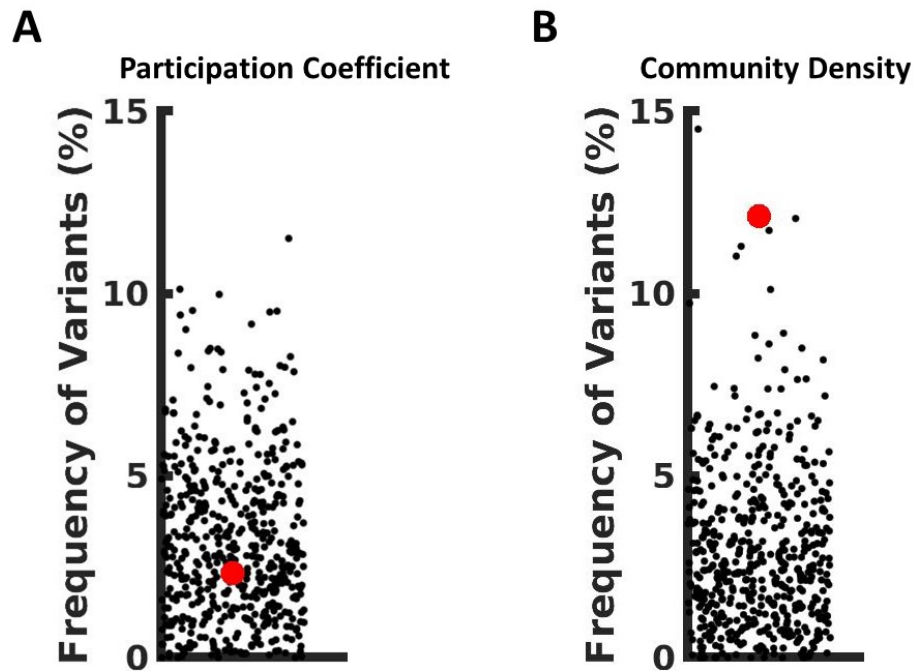


Figure S8: Relationship between variants and hubs when GSR was not included in processing. The results of the rotation test when applied to the No-GSR data indicated that the participation coefficient hubs do not differ in terms of variant overlap relative to random locations (2.31%, 95% CI [0.14%, 8.40%]) while community density hubs overlapped a greater degree than what would be expected (12.12%, 95% [0.05%, 7.86%]). This figure on the left (A) displays the frequency of networks variants at participation coefficient hubs (red dot) relative to random rotations (black dots). The figure on right (B) is figure is similar to A but shows the frequency of network variants at community density hubs (red dot) relative to random rotations (black dots). These results are comparable to the findings when GSR was included as a processing step.

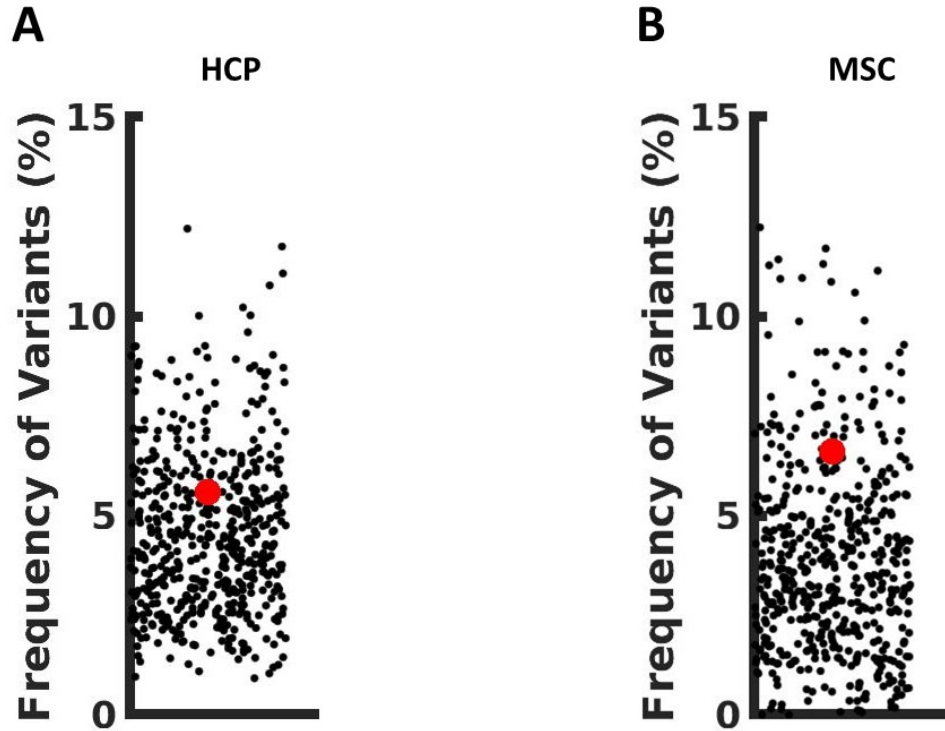


Figure S9: Consistency of top hubs defined with unthresholded connector hub measures. The results of the rotation test when applied to the diversity coefficient hubs (which use unthresholded positive connections; Rubinov & Sporns, 2011) indicated that these hubs, like participation coefficient hubs, do not differ in terms of variant overlap relative to random locations (HCP 5.59%, 95% CI [1.57%, 8.98%]; MSC (6.61%, 95% CI [0.37%, 9.10%])). This figure on the left (A) displays the frequency of networks variants at diversity coefficient based hubs (red dot) relative to random rotations (black dots) in the HCP dataset. The figure on right (B) displays the same information for the MSC dataset.

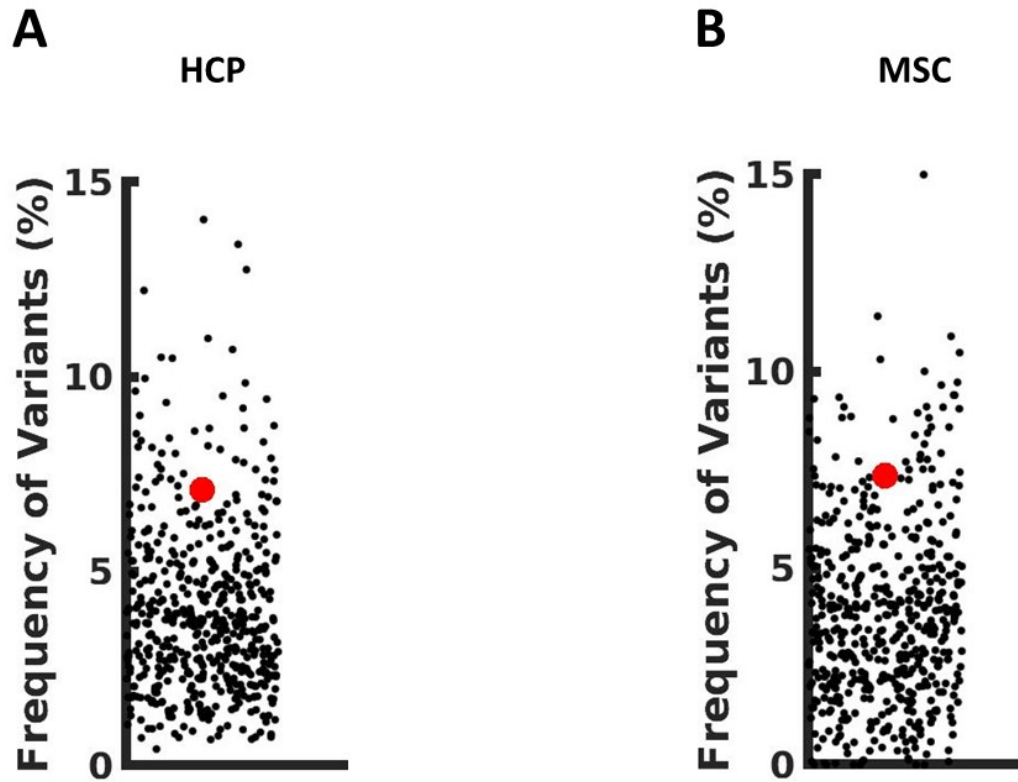


Figure S10: Consistency of top hubs defined with an unthresholded hub measure that incorporates negative edges. The results of the rotation test when applied to the positive/negative diversity coefficient hubs indicated that these hubs, like participation coefficient hubs, do not differ in terms of variant overlap relative to random locations (HCP 7.10%, 95% CI [0.95%, 9.56%]; MSC 7.34%, 95% CI [0.24%, 9.37%]). This figure on the left (A) displays the frequency of networks variants at diversity coefficient based hubs (red dot) relative to random rotations (black dots) in the HCP dataset. The figure on right (B) displays the same information for the MSC dataset.