



Network Level Analysis Toolbox 2020

Table of Contents

Preface	3
1. NLA Overview.....	4
1.1 The connectome and network structure	4
1.2 Why use this toolbox?	4
1.3 Introduction to NLA and enrichment	5
1.4 NLA Alternatives / Comparison to other analysis methods	6
2. NLA Methodology.....	8
2.1 Inputs / Brain Network Map Selection	8
2.2 General Linear Model / Edge-wise Statistical Model Selection.....	8
2.3 Connectivity Matrices	8
2.4 The NLA Method	8
2.5 How should the test statistic threshold be chosen?.....	9
2.6 How should the networks be chosen?	9
3. Getting Started	11
3.1 Pipelines Folder.....	11
3.2 Supportfiles Folder.....	11
3.3 Functional Connectivity Matrices	11
3.4 IM.mat File	11
3.5 Setting Parameters	12
4. Running Analyses.....	14
4.1 Network Level Analysis Functions	14
4.2 Chi-square and Hypergeometric Statistics	14
4.3 Permutation and Null Hypothesis Testing	14
4.4 dataOut Structure and Subfields.....	14
5. Plotting Results	16
5.1 Visualizing Significant Network Results.....	16
5.2 pairsOut Structure and Subfields.....	16
5.3 Figure Outputs	17
5.4 Re-visualizing Data.....	27
6. Generating Results Tables.....	28
6.1 Extracting Chi-square and McNemar test p-values.....	28
References	29

Preface

This is the reference manual for the Network Level Analysis (NLA) Toolbox. NLA is an extensible MATLAB-based software package for the analysis of behavioral associations with brain connectivity data. NLA utilizes a statistical approach known variously as ‘pathway analysis’, ‘over representation analysis’, or ‘enrichment analysis’, which was first used to describe behavioral or clinical associations in genome-wide association studies.^{1–4}

Enrichment is a model-based data reduction approach to elucidate statistically significant network-features. The suite developed here includes data-driven permutation-based false-positive-rate procedures that manage multiple comparisons corrections for one or two independent groups.

Software and Hardware Requirements:

NLA has been tested on MATLAB 2015a-2020. NLA requires the Statistics and Machine Learning Toolbox. Best performance will be achieved on a server setup with multiple cores to support parallel processing (particularly for the permutation testing portion of the toolbox).

NLA Toolbox v1.0 has been used in the following manuscripts:

1. **Wheelock, M.D.***, Lean, R.E.*, Bora, S., Austin, N.C., Melzer, T.R., Eggebrecht, A.T., Smyser, C.D., & Woodward, L.J. (2021). Altered brain-behavior relationships underlie attention impairments in very preterm children. *Cerebral Cortex*. 31, 1383-1394
2. **Wheelock, M.D.**, Hect, J.L., Hernandez-Andrade, E., Hassan, S.S., Romero, R., Eggebrecht, A.T., & Thomason, M.E. (2019). Sex differences in functional connectivity during fetal brain development. *Developmental Cognitive Neuroscience*. 36, 100632
3. Thomason, M.E., Hect, J.L., Rauh, V.A., Trentacosta, C., Arora, M., **Wheelock, M.D.**, Eggebrecht, A.T., Espinoza-Heredia, C., & Burt, S.A. (2019). Prenatal lead exposure impacts cross-hemispheric and long-range connectivity in the human fetal brain. *NeuroImage*. 191, 186-192.
4. **Wheelock, M.D.**, Austin, N.C., Bora, S., Eggebrecht, A.T., Melzer, T.R., Woodward, L.J., & Smyser, C.D. (2018). Altered functional network connectivity relates to motor development in children born very preterm. *NeuroImage*. 183, 574-583.

1. NLA Overview

1.1 The connectome and network structure

The term *connectome* essentially describes any network description of whole brain connectivity, from the microscale of single neurons and synapses up to the macroscale of entire brain regions and pathways.⁵ Connectomics is an ever-advancing field, and large-scale scientific endeavors such as the NIH's Human Connectome Project have made significant progress in mapping, analyzing, and understanding the human connectome. Contemporary connectome research views the brain as an extensive, complex network of non-adjacent, yet functionally and structurally connected brain regions.^{6,7} The connectome can be utilized to assess whole-brain associations between behavior and spatially distinct neural networks.

MRI has traditionally been viewed as the gold standard for mapping the connectome and has been used to demonstrate consistencies between the spatial topology of task-based activation studies and the brain networks derived from task-free functional connectivity.^{7,8} Contemporary cluster correction approaches do not utilize the spatial topology of brain networks when estimating cluster size significance.^{9–11} Therefore, there is an urgent need for standardized tools that address the robust hierarchical network structure of the brain and the limitations of contemporary neuroimaging analysis approaches by utilizing this biologically informed network structure to increase reproducibility and biological interpretation of neuroscience results.

1.2 Why use this toolbox?

The NLA toolbox is designed to address the multiple comparisons problem that occurs within connectome research, wherein studies use hundreds of regions of interest (ROI) to create connectomes with thousands of potential connections, yet they lack the tools to establish statistical significance when analyzing associations between connectome and behavior. For example, previous research failed to find any significant differences in brain connectivity that passed a connectome-wise false discovery rate (FDR) correction between individuals with a neurological disorder and healthy controls—a finding which contradicts the recognized role of the brain in neurological functioning.¹² Other studies have found connectome-behavior associations that pass the FDR correction, but lack the statistical tools necessary to definitively establish these observations.¹³ NLA, therefore, serves as a valuable tool for the statistical quantification of network-level associations with behavior. The toolbox relies on cross-disciplinary biostatistical approaches to evaluate brain-behavior relationships within the connectome and allows for control of FDR at the network level. In this way, NLA diverges from most contemporary tools with a focus on single connection associations, in that it is not

dependent on edgewise false positive rate (FPR) or spatially contiguous brain regions. By organizing connectivity-behavior associations according to underlying neurobiology (i.e., networks), NLA leverages the structure of the human connectome and provides a framework for rational interpretation and replication of findings across research methodologies. Finally, the integration of connectome analysis and visualization techniques within a single, extensible MATLAB-based pipeline makes NLA an expedient tool for statistical testing and production of publication quality images all in one package.

1.3 Introduction to NLA and enrichment

Network Level Analysis uses enrichment to evaluate whether functional network pairs demonstrate significant clustering of strong brain-behavior correlations. Enrichment applies common statistical tests to measure the clustering of associations within a given network pair and reduces the number of comparisons to those performed at the network level.⁴ Network level statistics such as the Chi-Square test, Hypergeometric test, and Kolmogorov Smirnov test have been used in numerous network-level investigations including joint attention and motor function in infants and toddlers, maternal inflammation during gestation, motor and attention development in very preterm children, sex differences during fetal brain development, and autism in adults.^{14–24}

First, correlations are calculated between behavioral scores and Fisher z-transformed functional connectivity correlation measures for each pair of ROI. For behavioral scores that are normally distributed, Pearson r correlations are used to calculate the associations. Non-parametric Spearman rank correlations are used to assess the relationship between functional connectivity and behavioral scores that are not normally distributed. Network pairs are then tested for enrichment of strong correlation values, defined as only those values that remain after being nominally thresholded. An uncorrected p -threshold (e.g., 0.05 or 0.01) is applied and the remaining correlations are binarized.

Within Group Network Level Statistic. The 1-degree-of-freedom χ^2 test is used to compare the observed number of strong (thresholded and binarized) brain-behavior correlations within one pair of functional networks to the number of strong brain-behavior correlations that would be expected if strong correlations were uniformly distributed across all possible network pairs. A large resulting test statistic can indicate that the number of strong correlations within a specific network pair is enriched. The hypergeometric test aims to assess the likelihood of observing a given number of strong correlations within a pair of networks, given 1) the total number of strong correlations observed over the entire connectome and 2) the total number of possible hits for that network pair (i.e., the total number of ROI-pairs within a given network pair).

Significance for both the hypergeometric and χ^2 tests is determined using permutation testing. Behavioral labels are randomly permuted and correlated with the connectome data (typically 10k times) to create null brain-behavior correlation matrices. Hypergeometric and χ^2 tests are calculated on these permuted brain-behavior correlation matrices generating a null distribution of network level statistics. The measured (real) Hypergeometric and χ^2 are compared to this null distribution to establish network-level significance.

Between Group Network Level Statistic. Finally, a McNemar χ^2 test can be used to investigate network pair-specific differences in brain-behavior relationships between groups. For each of these tests—the χ^2 enrichment, the hypergeometric, and the McNemar tests—empirical significance levels are determined using randomization. Networks are then considered to differ significantly between groups if they are both significantly different in their association with behavior between groups (McNemar $\chi^2 p < 0.05$) and are significantly enriched for brain-behavior (hypergeometric p and $\chi^2 p < 0.05$).

1.4 NLA Alternatives / Comparison to other analysis methods

The NLA toolbox's use of a novel enrichment approach makes it a transformative tool in connectome-wide association studies, given that all current enrichment analysis methods are built for use with genome data and NLA is the first enrichment tool designed to analyze the connectome. Many alternative methods for connectome analysis rely on spatial extent cluster correction in order to control voxel-wise whole brain connectome FPR.^{25,26} Despite mounting evidence that spatially non-contiguous brain regions are strongly correlated and often co-activate to the same stimuli, cluster extent correction is often regarded as the ideal thresholding approach in human connectome literature. By basing statistical significance on contiguous voxels, however, cluster extent correction methods fail to account for this covariance structure. Therefore, brain regions that are known to be highly correlated and part of the same network—such as the anterior cingulate and posterior cingulate—may be thresholded separately, resulting in one or both separate regions not meeting statistical thresholds.²⁷ NLA is distinguished from the cluster extent correction methodology in that it groups highly correlated, non-contiguous brain regions based on pre-defined network modules prior to estimating network-level significance.

Given this deviation from the popular extent cluster correction thresholding method, the most conceptually similar existing connectome analysis approach to NLA is the *Network Based Statistic* (NBS) toolbox.²⁸ NBS was the first tool control the edgewise FPR by leveraging graph-based estimates of modularity. Still, several crucial differences exist between NLA and NBS: (a) the results from NBS focus on edgewise significance as opposed to network-level significance, (b) NBS does not

have a built-in visualization functionality, and (c) NBS allows for different module sizes, number of network modules, and configurations of edges assigned to network modules across various clinical populations, but draws no conclusions regarding the biological relevance of identified networks. The NLA pipeline addresses this issue by presenting a vast array of analysis and visualization options that utilize biologically informed hierarchical organization models of the brain.

Graph Theoretical Toolboxes are another comparable approach to NLA, offering an analysis methodology to quantify network characteristics such as integration, segregation, resilience, and relative contribution of individual network nodes to overall information flow within the network.²⁹ Various other toolboxes have been created to address network thresholding, graph metric calculation, and graph visualization—such as GRETNA, GEPHI, and BrainNet Viewer. Additional methodologies aim to determine network topology differences by leveraging generalized estimating equations and generalized linear and nonlinear mixed models.^{30–32} Each of these tools has helped to advance the application of graph theory approaches to connectome analysis. The NLA toolbox estimates statistical associations edgewise, rather than on network topology features, thereby providing a crucial and complementary approach to the existing collection of brain network analysis tools.

Statistical Inference and the use of liberal primary thresholds. NLA establishes statistical significance in the weak sense similar to traditional voxelwise cluster-level inference.³³ In voxelwise cluster correction, a liberal primary threshold is employed in addition to a cluster-extent threshold (determined by e.g., random field theory or Monte Carlo simulations). The resulting clusters are significant but inferences cannot be made about any particular sub-regions or voxels within a cluster. Similarly, NLA with Chi-square or Hypergeometric tests employs a liberal primary threshold in order to calculate the network-level statistic and significance is established with permutation testing, but claims cannot be made about the significance of any given ROI-pair within the network. One could apply an FDR correction within each network pair similar to the statistics outlined in the Network Based Statistics toolbox though this would still only control the false positive rate in the weak sense. The motivation of all of these approaches (cluster-level inference, network-level enrichment, network-based statistic) is to control the false positive rate when a massive number of tests are performed. Controlling the false positive rate in the strong sense with several thousand functional connections (e.g., 30k) will often result in no single ROI-pair surviving OR a few scattered ROI-pairs surviving with no clear biological pattern.¹²

2. NLA Methodology

2.1 Inputs / Brain Network Map Selection

The first step of NLA involves specifying the network map that will be used to depict the known architecture of the human connectome, which is crucial given that the network map selection affects both statistical significance testing and interpretation.³⁴ The current pipeline uses network maps that are generated with Infomap, due to its greater congruence with networks derived from task-activation and seed-based connectivity studies than alternative modularity algorithms.^{7,35} Network maps can be generated using one's preferred algorithm or one of several published ROI and corresponding network map options that will be included in the NLA toolbox.^{6,7,36–39} The use of standardized ROI and network maps creates a common, reproducible framework for testing brain-behavior associations across connectome research.

2.2 General Linear Model / Edge-wise Statistical Model Selection

After selecting a network map, the next step in the pipeline involves specifying the desired statistical model for testing associations between behavioral data and edge-wise—or ROI-pair connectivity—connectome data. The analysis pipeline within the NLA toolbox offers both parametric and non-parametric correlation.

Future versions of the NLA toolbox will potentially include generalized estimating equations (GEE), which allow for covariate adjustment and correlated outcomes, as well as statistical tests for assessing group differences, including nonparametric Wilcoxon rank sum and parametric Welch's t-test. The addition of these statistical options will facilitate the analysis of connectome architecture differences between groups and of correlated and longitudinal data.

2.3 Connectivity Matrices

Other software packages are used to create the connectivity matrices that are provided as input into the NLA toolbox. One useful option for mapping functional connectivity matrices is CONN—a MATLAB-based software with the ability to compute, display, and analyze functional connectivity in fMRI.

2.4 The NLA Method

The first step in the NLA pipeline consists of selecting a network map to characterize the known architecture of the human connectome. Once that is chosen, connectome-wide associations are calculated between ROI-pair connectivity and behavioral data, resulting in a set of standardized regression coefficients that specify the brain-behavior association at each ROI-pair of the

connectome matrix. Next, network level analysis—consisting of transformation of the edge-wise test statistics and enrichment statistic calculation⁴⁰—is done to determine which networks are strongly associated with the behavior of interest.

Both *p*-value and test-statistic binarization are offered in the current NLA pipeline.^{14,15} Prior research has supported the incorporation of a proportional edge density threshold, given that uneven edge density thresholds have been shown to unfairly bias results.⁴¹ Other transformation options include transforming edge-wise connectivity into absolute values prior to NLA or leaving data untransformed.^{17,42}

For enrichment statistic calculation, the NLA pipeline offers chi-square and hypergeometric tests. Prior research has relied on chi-square and Fisher's Exact test, as well as a Kolmogorov-Smirnov (KS) test and non-parametric tests based on ranks, which compare the distribution of test values within a region to other regions.^{15,17,43,44} In addition, KS alternatives such as averaging or minmax have also shown promise in connectome applications.⁴⁵⁻⁴⁸

The next step in the pipeline involves conducting data-driven permutation testing to establish significance. In the NLA toolbox, network level significance is determined by comparing each measured enrichment statistic to permuted enrichment *p*-values which are calculated by randomly shuffling behavior vector labels and computing the enrichment statistic 10,000 times to produce a null distribution for each network. The FPR is controlled at the network level either using Bonferroni or FDR methods. Therefore, NLA is able to retain edge-wise correlations within each network module, but network communities are used to reduce the number of comparisons and control the FPR at the network level. After significance is determined, the pipeline allows users to create publication quality images to visualize network level findings both in connectome format and on the surface of the brain.

Note: While the behavior vector labels are shuffled to conduct permutations in the enrichment pipeline, functional connectivity data are not shuffled in order to preserve the inherent covariant structure of the data across permutations.

2.5 How should the test statistic threshold be chosen?

A nominal threshold is used for the thresholding and binarization step of the chi-square test. The nominal threshold is uncorrected and is typically set at 0.05 or 0.01 using params.Pmax. In contrast, a network-level corrected threshold using the Bonferroni method is set in params.B, where the nominal threshold is divided by the number of tests being done to correct for multiple comparisons.

2.6 How should the networks be chosen?

There are many canonical ROI sets and there are many network definitions. Some of these network definitions include ROI that are not consistently assigned to any network. These ROI are typically removed prior to network level analysis, as is the case in the IM_Seitzman_15nets_288ROI and the IM_Gordon_12nets_286Parcels network structures included in this version of the toolbox. Network structures that are not included in this package may also be used, but they must first be formatted into an IM structure.

3. Getting Started

3.1 Pipelines Folder

After downloading the NLA toolbox, begin by navigating to the pipelines folder. The pipelines folder contains multiple pipeline options for running NLA.

- Enrichment_pipeline.m – Script for running Enrichment analysis between functional connectivity data and behavioral data on a single group.
- Script_for_EnrichmentVisualization.m – Script for visualizing single group Enrichment results.
- Enrichment_Mcnemar_pipeline.m – Script for assessing enrichment within groups using the chi-square and hypergeometric tests and comparing enrichment between groups using the McNemar chi-square test.

3.2 Supportfiles Folder

Within the supportfiles folder are multiple atlas spaces and IM.mat files to choose from. The atlas options are as follows:

- MNI_coord_meshes_32k.mat – MNI atlas in MNI coordinate space. Can be visualized in standard, inflated, or very inflated space.
- Conte69_on_TT_32k.mat – standard MNI atlas, transformed into Talairach space. Can be visualized in standard, inflated, or very inflated space.
- CH2_MNI_withCB.mat – Colin Holmes atlas in MNI with cerebellum. Can only be visualized in standard space.

3.3 Functional Connectivity Matrices

The NLA toolbox requires that functional connectivity matrices use z-values. Matrices can initially be constructed using either r- or z-values, but the FisherR2Z command must be used to transform matrices with r-values into z-values before proceeding with the enrichment pipeline.

3.4 IM.mat File

The IM.mat file outlines the organization of the data for the enrichment pipeline. IM files are located within the supportfiles folder of the NLA toolbox. Included in the toolbox are several canonical IM.mat file options.

- IM_Seitzman_17nets_300ROI – in volume space. Consists of 300 ROI and corresponding 17 networks.⁴⁹ ROI can be found at

<https://wustl.app.box.com/s/twpyb1pfj6vrlxgh3rohyqanxbdpelw>

- IM_Seitzman_15nets_288ROI – in volume space. Consists of 288 ROI and corresponding 15 networks, due to 12 ROI lacking consistent placement in a network.⁴⁹
- IM_Gordon_13nets_333Parcels – in surface space. Consists of 333 parcels and corresponding 13 networks.⁶
- IM_Gordon_12nets_286Parcels – in surface space. Consists of 286 parcels and corresponding 12 networks.⁶

Each IM.mat file includes the following:

- name – specifies name of IM structure
- cMap – specifies network colors
- Nets – specifies network names
- ROIxyz – specifies ROI in corresponding order that they appear in networks
- key – specifies ROI and the networks they fall into (always in final network order)
 - column 1 = total count of ROI
 - column 2 = network that each ROI falls into
- order – a sort function that tells ROI where to be placed in the matrix corresponding to the network order in IM.key

Please note that both the IM_Gordon_13nets_333Parcels and IM_Gordon_12nets_286Parcels files contain surface parcels, but they are being visualized as spherical ROI.

3.5 Setting Parameters

There are numerous parameters that must be specified for the NLA model.

- params.np – number of permutations or iterations of shuffling subject functional connectivity and behavior pairs. 1e4 is default
- params.type – Pearson (normally distributed behavioral data) or Spearman (not-normally distributed behavioral data)
- params.Pmax – amount of correlations let through that are nominally significant
- params.nnPmax – network significance
- params.group – group name
- params.BxName – behavior name
- params.B – Bonferroni level (i.e., how many behaviors are being used?)
- params.fn – file name
- params.ScaleRadius – sets ROI size to reflect degree

- params.roiradius – should be set at 5 for adult data

4. Running Analyses

4.1 Network Level Analysis Functions

There are multiple versions of the network level analysis function, based on which pipeline one uses to run their analyses.

- One group at one timepoint – `dataOut=fcBx_Enrich_1tp(fc,Bx,IM,params);`
- Two groups at one timepoint –
`[dataOut]=fcBx_Chi_McNemar_TwoGroups(dataIN.fc,dataIN.Bx,dataIN.IM,dataIN.params);`

4.2 Chi-square and Hypergeometric Statistics

The NLA pipeline utilizes both chi-square and hypergeometric tests, which are embedded within the NLA functions.

4.3 Permutation and Null Hypothesis Testing

Permutation testing is also embedded within the NLA toolbox. First, the number of desired permutations is set in `params.np`, with the default being 10,000 iterations. In each iteration, the behavioral data are randomized across subjects to make the functional connectivity-behavior pairs all ‘false.’ With that false data, correlations are then calculated between ROI pairs and the behavior. Next, chi-square and hypergeometric values are calculated based on the significant hits within each network pair of the permuted data. This step is done for every iteration, resulting in a set of chi-square and hypergeometric values for the permuted data that serves as the basis for the null distribution. The null distribution is then compared to the true statistics for the observed functional connectivity-behavior pairings.

The permutation-based FPR can be used to assess the empirical probability of calculating the statistical value of enrichment for the actual, observed data at the 5% false positive rate level for the permutations.¹⁴ The control of the FPR for the chi-square and hypergeometric tests is done by testing the ‘true’ statistics for all network pairs against the pooled set of ‘false’ statistics across all pairs and permutations. The alternative approach would be to instead calculate the final *p*-value based on each network pair’s permutations alone, but currently, the toolbox utilizes comparisons at the network-level.

4.4 dataOut Structure and Subfields

The `dataOut` structure contains numerous helpful statistics that may be plotted or published in a table.

- `Np` – number of ROI pairs per network pair
- `rho` – measured brain-behavior correlation, using non-shuffled data for both groups

- pval – measured *p*-value
- thresholdedROIpairs – edges that remain after nominal thresholding
- Chi_stats – chi-square within-group enrichment statistics
- Chi_pval0 – chi-square thresholded *p*-value
- Chi_pval – chi-square measured *p*-value
- Chi_EWpval – chi-square experiment-wide corrected *p*-value
- Chi_pval_gte – chi-square experiment-wide *p*-value that pass nominal threshold
- HGp – hypergeometric measured *p*-value
- HGpp – hypergeometric permuted *p*-value
- HGppEW – hypergeometric experiment-wide corrected *p*-value
- McNemar_b – used in McNemar statistic $(b-c)^2 / (b+c)$, where b and c are discordant values between two groups
- McNemar_c – used in McNemar statistic $(b-c)^2 / (b+c)$, where b and c are discordant values between two groups
- McNemar_P – measured *p*-value for groups
- McNemar_PP – permuted *p*-value from shuffling behavioral labels
- McNemar_PPEW – McNemar experiment-wide corrected *p*-value
- McNemar_stats – measured chi-square statistic for both groups
- TNidx – network pair indexing. Index number for network pairs if tril (triangle lower) is converted to column format and zeroes are removed from matrix.
- Emp_FDR_CS – empirical chi-square permuted distributions that underlie permuted *p*-values
- Emp_FDR_HG – empirical hypergeometric permuted distributions that underlie permuted *p*-values
- Chi_sig – logical index of when chi-square experiment-wide corrected *p*-values pass significance threshold
- HG_sig – logical index of when hypergeometric experiment-wide corrected *p*-values pass significance threshold
- Emp_FDR_McN – McNemar values corresponding to the tests of the permuted between-group chi-square values
- McN_sig – logical index of significant McNemar networks passing user-specified threshold
- ROIpairthreshold – nominal *p*-value threshold, as specified in params.Pmax

5. Plotting Results

5.1 Visualizing Significant Network Results

There are three options for visualizing significant network pairs on the brain:

- View_Clear_Brain_ROI_Corr – used to plot each significant network pair separately on glass brain, with colored lines to indicate whether the correlations are negative or positive as shown in Figure 5.
- View_Clear_Brain_ROI_Corr_Multi – used to plot all significant network pairs together on one brain, with colored lines to indicate whether the correlations are negative or positive as shown in Figure 6.
- View_Clear_Brain_ROI_Corr_fc – used to plot each significant network pair separately on glass brain with colored lines to indicate the direction of functional connectivity, as shown in Figure 7.

5.2 pairsOut Structure and Subfields

The pairsOut structure is generated by the View_Clear_Brain_ROI_Corr_fc function. Each cell within the structure corresponds to one of the significant network pairs that were plotted. The network pair can be determined by looking to the first and second column, which correspond to the index in IM.key where the ROI fell. The pairsOut structure should be in the order that the r,c (row, column) were specified. The columns of data in the pairsOut structure provide information about the significant network pairs:

- Columns 1 & 2 – network indexes for each ROI pair that was nominally significant at the pmax value (typically either 0.05 or 0.01)
- Column 3 – the brain-behavior correlation coefficient (Pearson or Spearman)
- Column 4 – Coloring of lines between ROI pairs. Color can be determined based on the percentage of functional connectivity values that are greater than zero, the mean of the functional connectivity values, or the median of the functional connectivity values
- Column 5 – p-value for the brain-behavior correlation
- Column 6 – minimum functional connectivity across subjects for specified ROI pair
- Column 7 – maximum functional connectivity across subjects for specified ROI pair
- Column 8 – standard error of the mean
- Column 9 – mean
- Column 10 – standard deviation
- Column 11 – median

- Column 12 – bottom quartile value of sorted functional connectivity data
- Column 13 – top quartile value of the sorted functional connectivity data

Columns 3 and 5 can be used to get a broad sense of which pairs are the most significant and what direction (negative or positive) those correlations are. This information is used in the Script_for_EnrichmentVisualization.m to make scatter plots for specific ROI pairs within specific network pairs. First, choose a significant network pair from within the pairsOut structure. The column number of the desired network should be specified as the fnPair. Within a specific network pair, there is information about each individual ROI pair. The row number of an ROI pair of interest should then be indicated as the pairNum. After specifying the network and ROI pairs, a scatter plot can be generated, as shown in Figure 10.

5.3 Figure Outputs

The following are examples of typical NLA figure outputs:

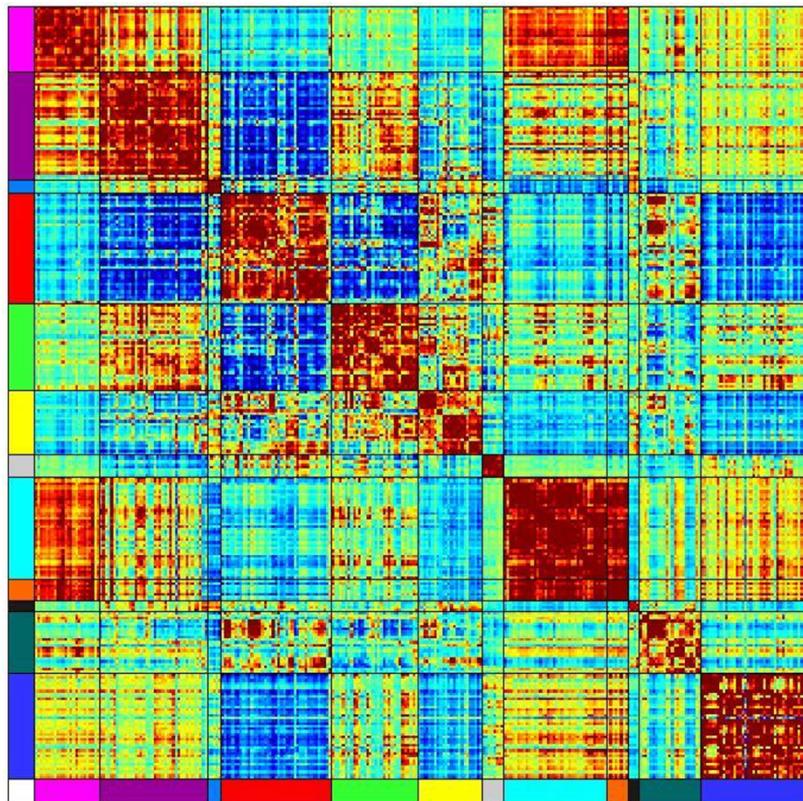


Figure 1. Matrix of r -values as organized by the Gordon 12 Network IM file. Created using Martrix_Org3 function.

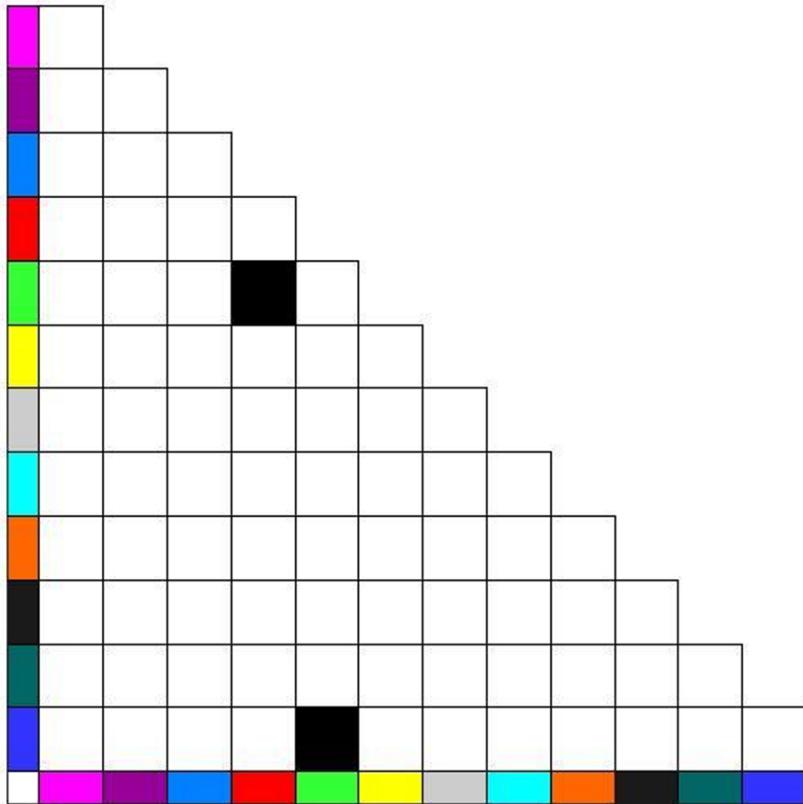


Figure 2. Matrix of network pairs meeting significance threshold. Created using Enrichment_Figs2 function.

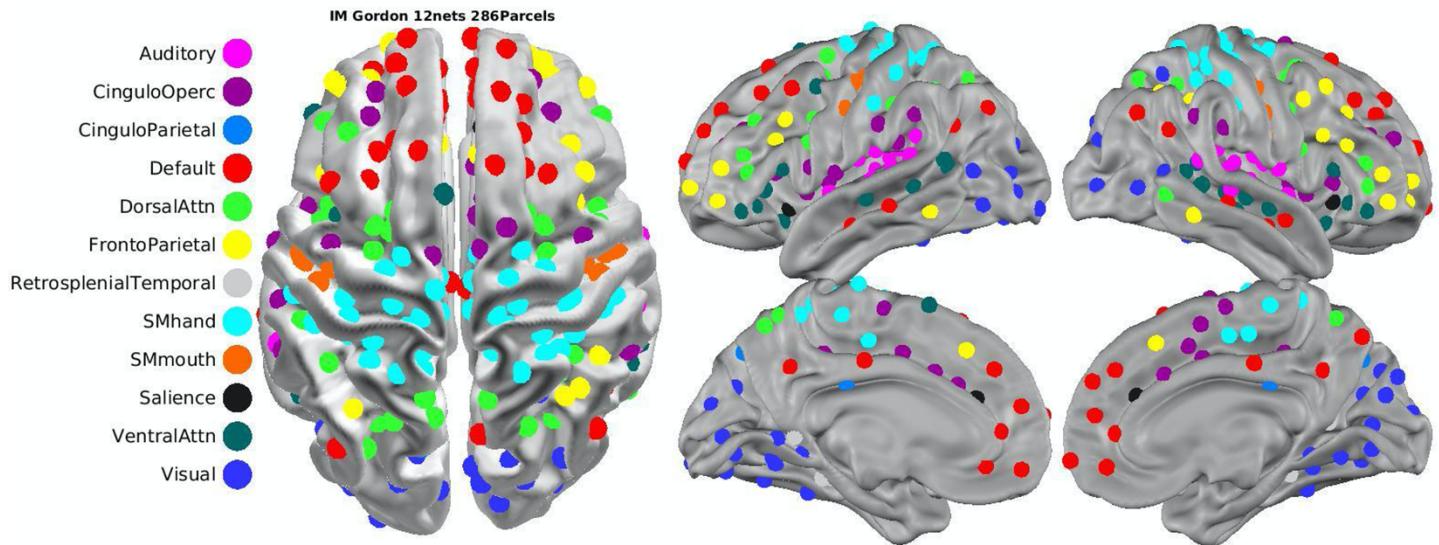


Figure 3. Visualization of the Gordon 12 network IM model on brain surface with network names and corresponding colors. Created using View_ROI_Modules function. Note that the Gordon surface parcels are plotted as a sphere at the center of mass for each parcel.

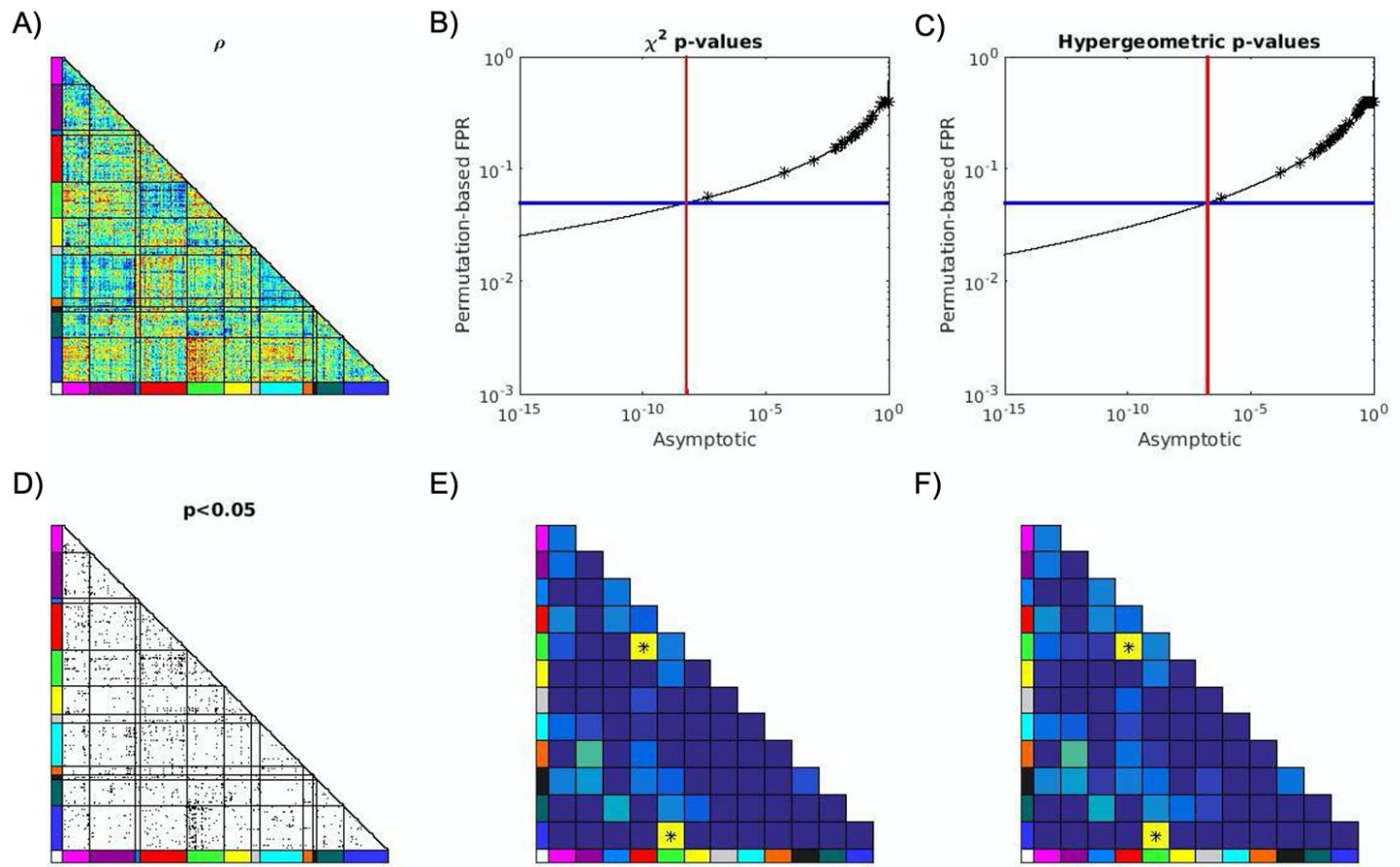


Figure 4. A) Matrix with rho correlation values for each network pair. B) Chi-square and C) hypergeometric p-values. Significance is determined via a permutation-based FPR. The asymptotic p-value reflects the probability of the measured statistic assuming infinitely many samples. Black lines: the brain-wide FPR represents the frequency of a given asymptotic p-value occurring in any network-pair in the randomization. Blue lines: 0.05 significance threshold in the FPR. Red lines: threshold in asymptotic values. Black stars: network-pair specific statistics. D) Matrix with black dots representing correlations significant at $p < 0.05$. E) Matrix showing significant network pairs from chi-square test. F) Matrix showing significant network pairs from hypergeometric test. Created using fcBx_Enrich_1tp function.

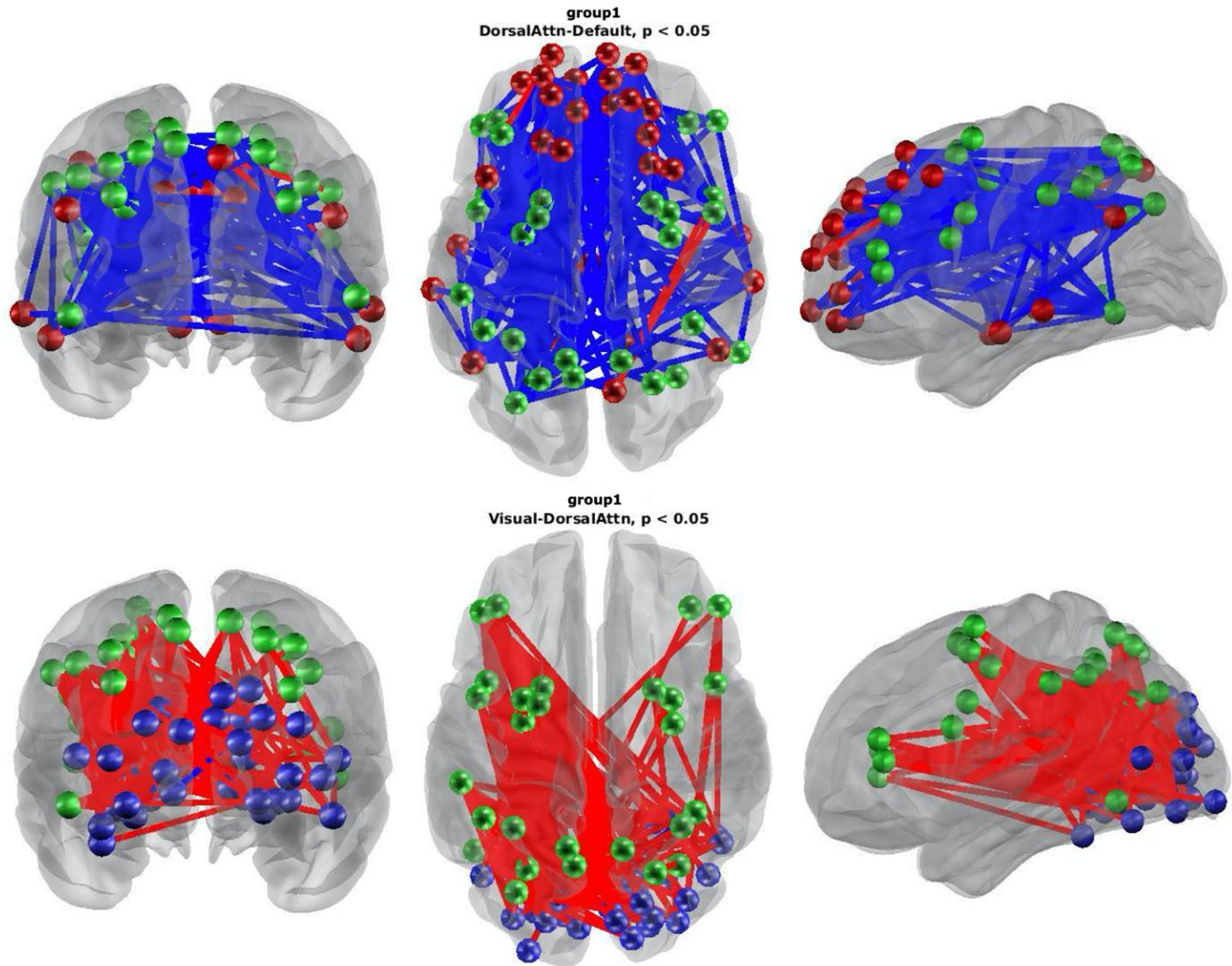


Figure 5. Visualizations of significant network pairs on a glass brain, with blue lines indicating a negative correlation and red lines indicating a positive correlation. Created using View_Clear_Brain_ROI_Corr function.

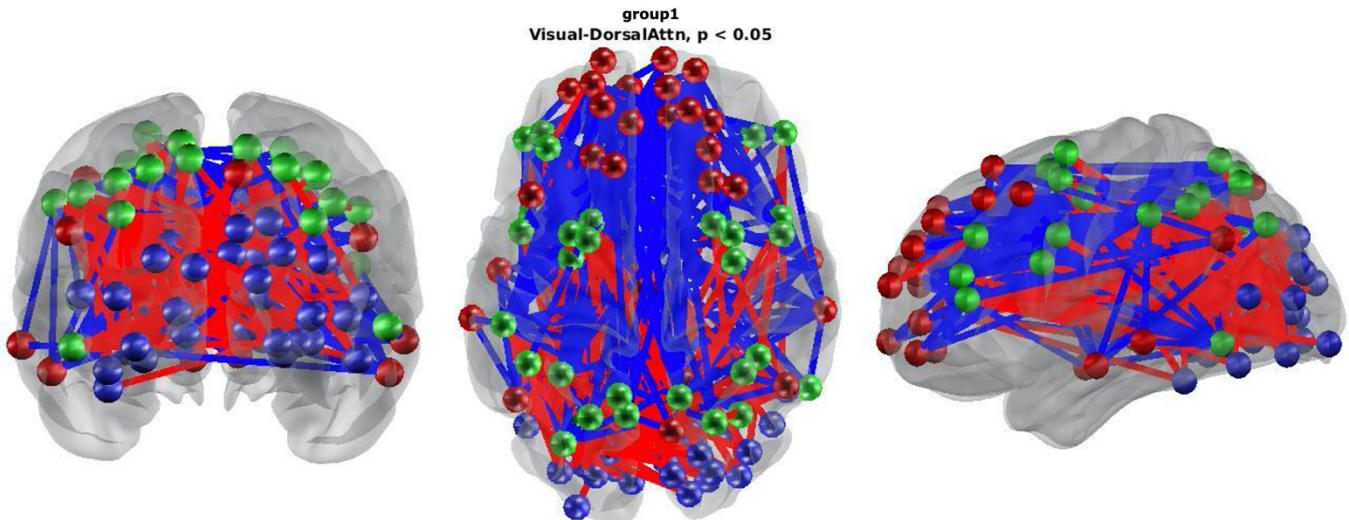


Figure 6. Visualization of all significant network pairs plotted on single glass brain, with blue lines indicating a negative correlation and red lines indicating a positive correlation. Created using View_Clear_Brain_ROI_Corr_Multi function.

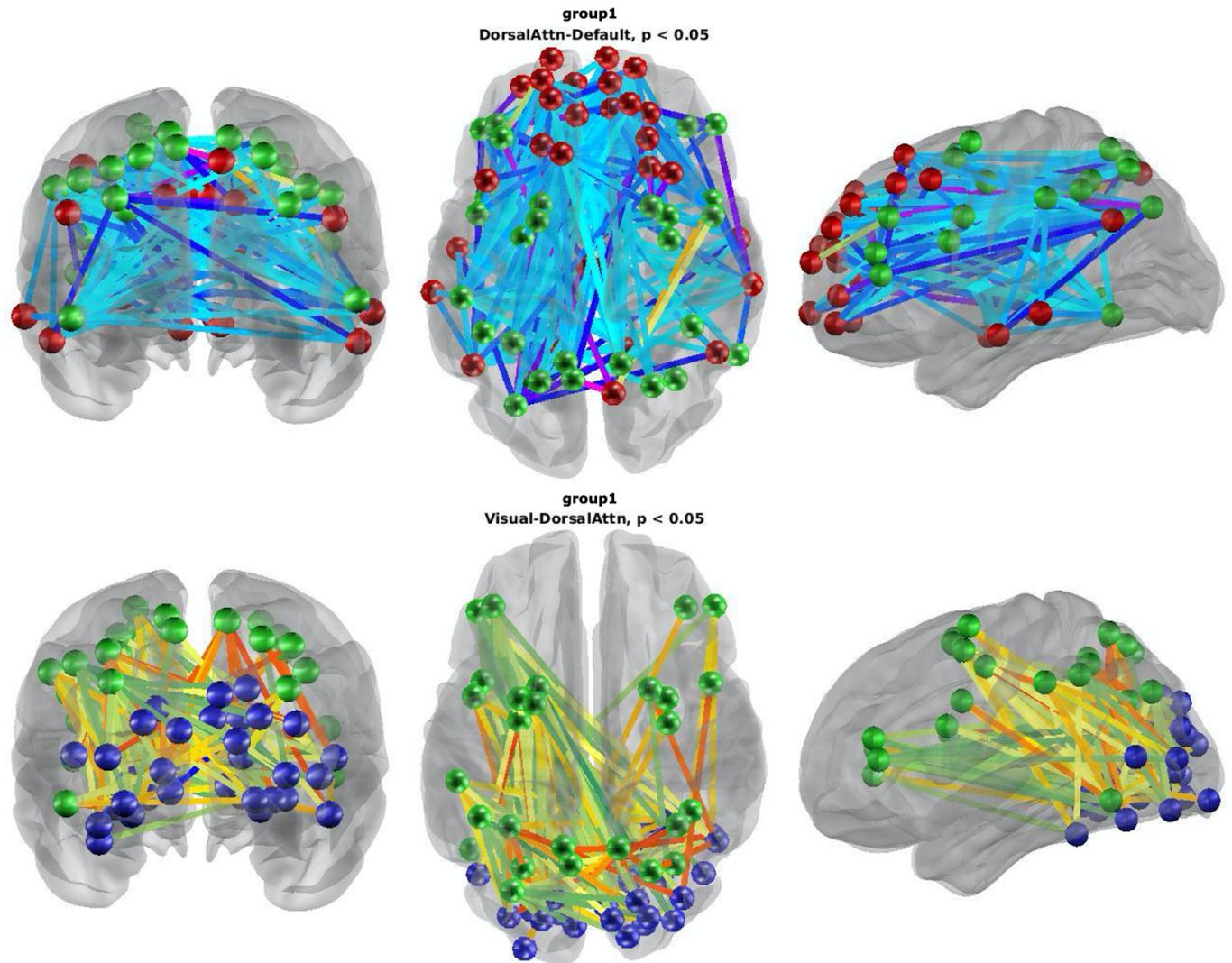


Figure 7. Visualizations of significant network pairs on a glass brain, where the line color indicates the direction of functional connectivity according to the color bar in Figure 8. Created using View_Clear_Brain_ROI_Corr_fc function.

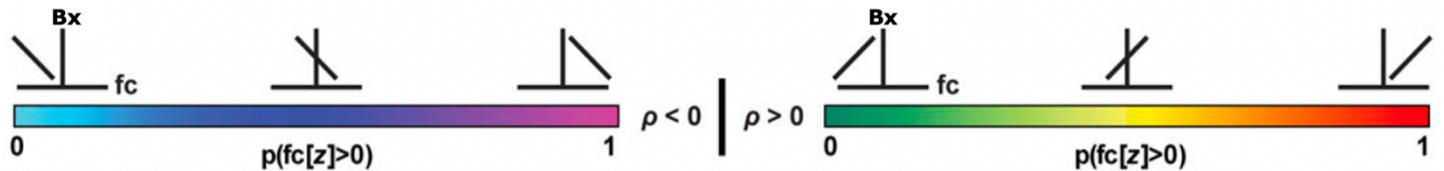


Figure 8. Color bars representing both negative (left) and positive (right) correlations for brain-behavior associations. The color of the lines connecting pairs of ROI (as shown in Figure 7) reflects

the proportion of individual fcMRI values that are above zero. Light blue to dark blue to magenta colors reflect negative brain-behavior correlation and green to yellow to red colors denote positive brain-behavior correlation. Light blue and green denote that the ROI pair contains only negative fcMRI values between the ROI pairs. Magenta and red reflect ROI pairs with only positive fcMRI. Blue and yellow reflect fcMRI values distributed across zero.

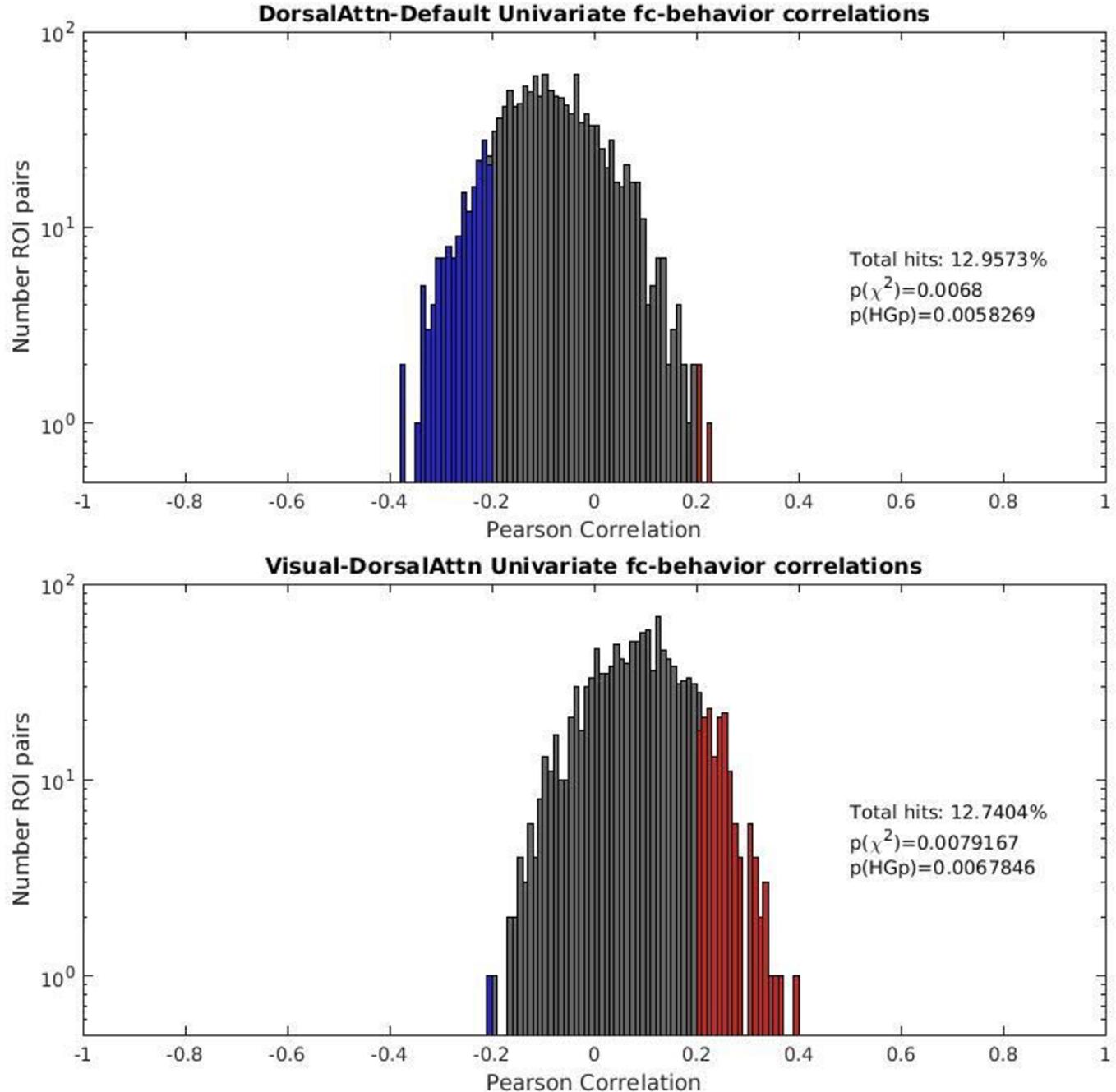


Figure 9. Plots of the distribution of correlations between functional connectivity and behavior. Blue (negative) and red (positive) bars represent ROI pairs that pass the user-specified threshold. In the first example, the majority of the significant ROI pairs passing the nominal threshold are negative, meaning negative correlations are driving the significance of this network pair. In the second

example, the majority of the significant ROI pairs passing the nominal threshold are positive, meaning positive correlations are driving the significance of this network pair. Created using View_Clear_Brain_ROI_Corr function.

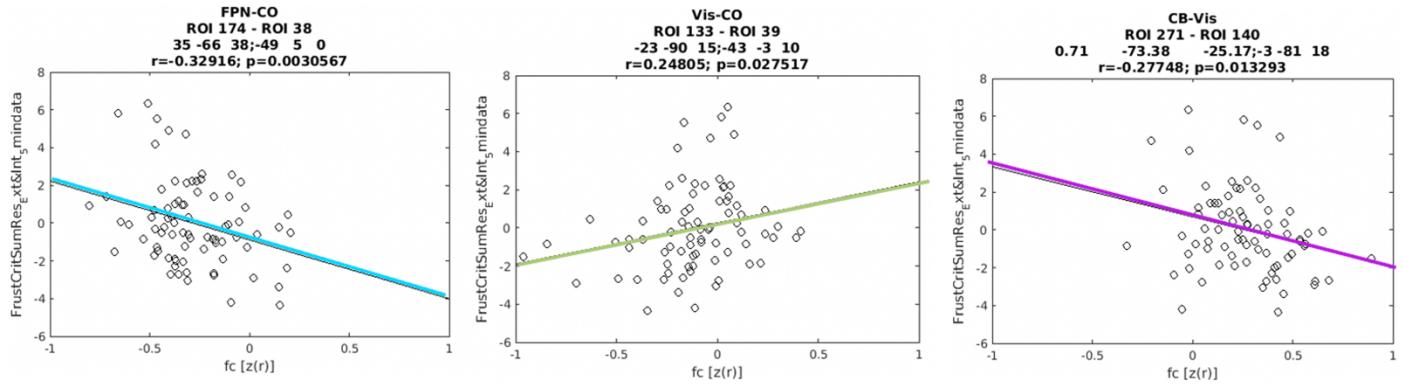


Figure 10. Scatterplot examples for a specific ROI pair within a specific network pair. Plots indicate the network names, ROI index numbers, coordinates, Spearman or Pearson rho values for the ROI pair, and the corresponding *p*-value. Note: colored lines overlaying scatterplots are not included in this toolbox release and are shown here only to demonstrate the use of the color bar in Figure 8.

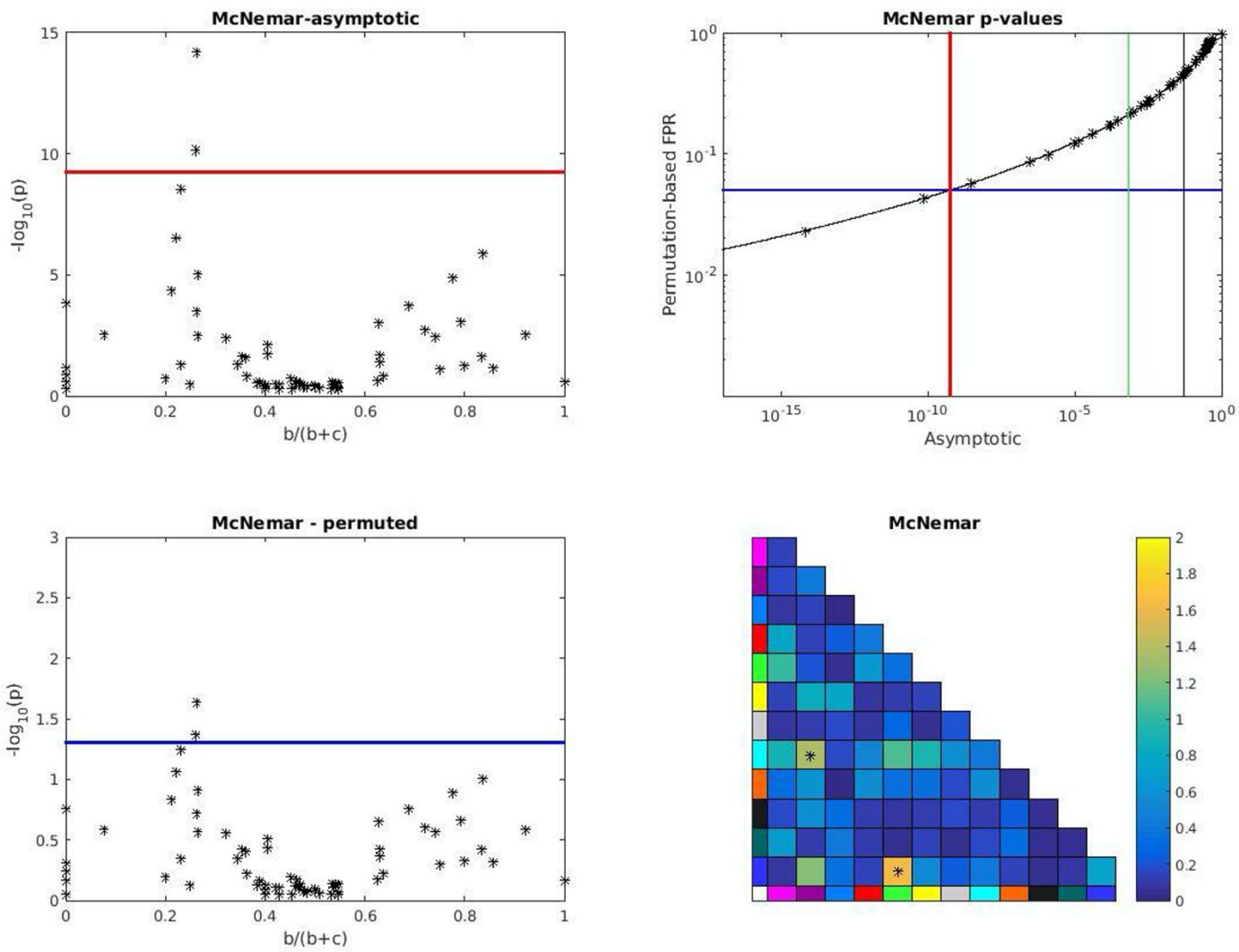


Figure 11. Plots illustrating the McNemar χ^2 test, which is used to determine the differences in brain-behavior relationships between two groups. The McNemar statistic uses the number of discordant tests between groups, and it is typically depicted as $(b - c)^2 / (b + c)$, where b indicates the number of ROI pairs that are true for group 1 but false for group 2 and c indicates the number of ROI pairs that are false for group 1 but true for group 2. Created using `fcBx_ChI_McNemar_TwoGroups` function.

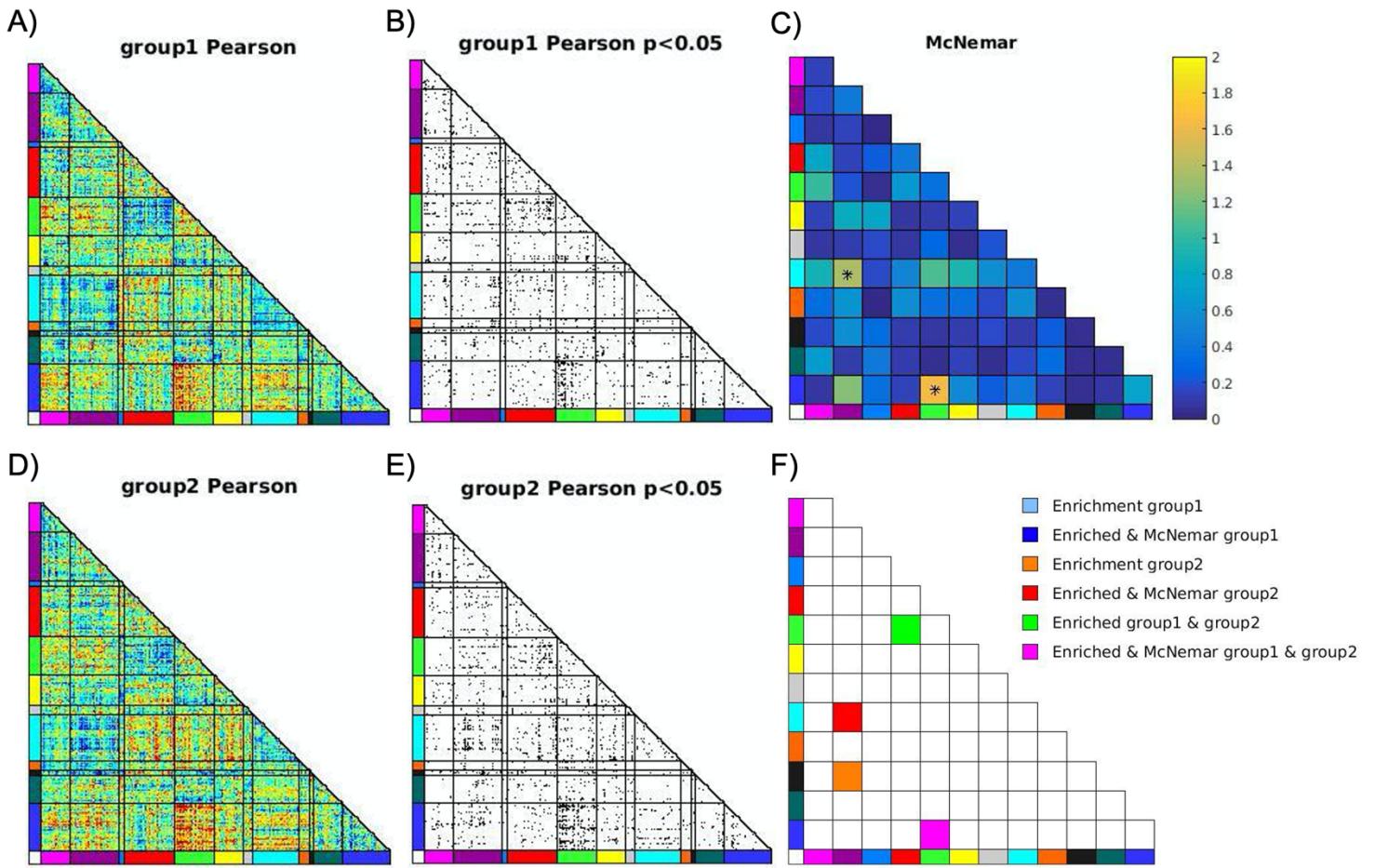
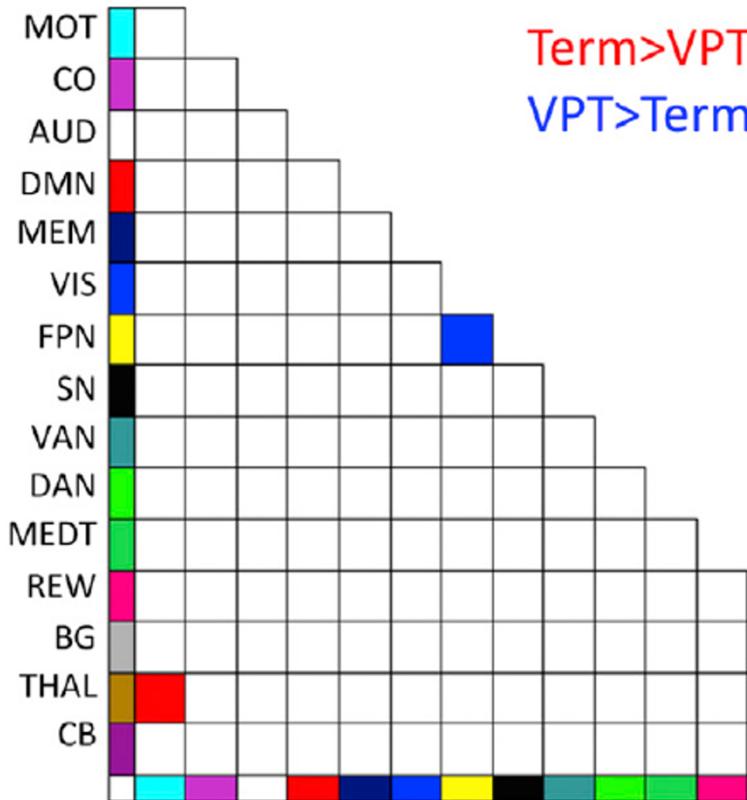


Figure 12. Plots illustrating the differences between groups. A) Connectivity matrix and B) thresholded connectivity matrix for group 1. C) McNemar χ^2 test. D) Connectivity matrix and E) thresholded connectivity matrix for group 2. F) McNemar group differences. Green squares indicate network pairs that are not significantly different between groups but are enriched in both groups. Pink squares indicate network pairs that have significant ROI in both groups, but the ROI differ between groups. Red squares indicate network pairs that are significantly more enriched in group 2 than group 1. Orange squares indicate network pairs that are enriched for group 2. Not shown in this figure are light and dark blue squares, which indicate network pairs that are enriched for group 1, and network pairs that are significantly more enriched in group 1 than group 2, respectively. Created using `fxBx_Chi_McNemar_TwoGroups` function.



Term>VPT

VPT>Term

Figure 13. Plot illustrating a typical output demonstrating McNemar group differences between groups.¹⁵ Blue squares indicate networks exhibiting strong enrichment of brain-behavior correlations in group 1 (very preterm) than group 2 (term). Red squares indicate network pairs that were significantly more enriched in group 2 (term) than group 1 (very preterm).

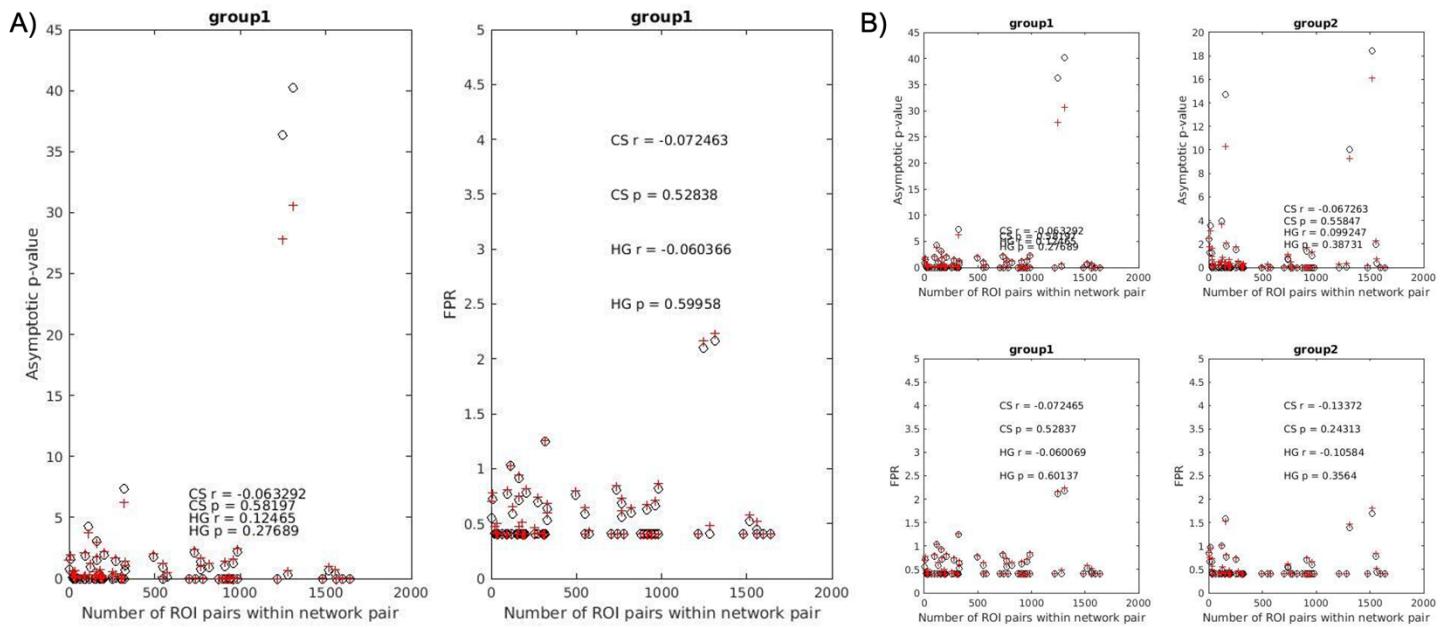


Figure 14. Diagnostic plots used to assess the within-network-pair-only FPR to ensure no asymptotic *p*-values showed up in the randomizations more than 5% of the time, as well as to determine whether

there is a relationship between the permutation-based FPR and the number of ROI within a network pair. Red '+' denote χ^2 ; Black circles denote hypergeometric. CS stands for chi-square; HG stands for hypergeometric. r -values are correlation coefficients; p -values are significance levels. Non-significant p -values and weak correlation coefficients indicate that there is no relationship between the number of ROI pairs within a network pair and the permutation-based FPR or the asymptotic p-values. A) The plot created in the one timepoint code using the fxBx_Enrich_1tp function and B) the plot created in the two timepoint code using the fxBx_Chi_McNemar_TwoGroups function.

5.4 Re-visualizing Data

Results can be replotted using new network cutoff values.

- `Perm_1tp_Figs2(dataOut,IM,params.B,params);`

6. Generating Results Tables

6.1 Extracting Chi-square and McNemar test p-values

The NLA toolbox includes code to easily make results tables exploring significant and enriched network pairs. First, network pairs of interest should be indicated, which can be found using the McNemar matrix exemplified in Figure 12F. To find network pair numbers, count down rows and then across columns. For example, in Figure 12F, the network pair numbers would be green (5,4), orange (10,2), red (8,2), and pink (12,5).

After specifying network pairs and running the full section of code, open NetStats to find enrichment and McNemar statistics, including the χ^2 , ϕ , and p -values for group 1, group 2, and the McNemar test. Below is an example of an NLA statistics table that can be created using these values. Colors correspond to those noted in Figure 12F.

		Enrichment group 1			Enrichment group 2			McNemar			
Network-Pair		N	χ^2	ϕ	p value	χ^2	ϕ	p value	χ^2	ϕ	p value
DorsalAttn	DMN	1312	179.6416	0.3700	0.0068	41.9227	0.1788	0.0409	9.6267	0.0857	0.2231
Salience	CinguloOperc	160	11.0161	0.2624	0.1210	63.1941	0.6285	0.0263	7.5294	0.2169	0.2698
SMhand	CinguloOperc	1520	1.3797	0.0301	0.3021	79.9688	0.2294	0.0199	41.0889	0.1644	0.0427
Visual	DorsalAttn	1248	161.7572	0.3600	0.0079	645.5162	0.7192	1E-06	59.4106	0.2182	0.0231

References

1. Rivals, I., Personnaz, L., Taing, L. & Potier, M.-C. Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics* **23**, 401–407 (2007).
2. Khatri, P., Sirota, M. & Butte, A. J. Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges. *PLoS Comput. Biol.* **8**, e1002375 (2012).
3. Backes, C. *et al.* Systematic permutation testing in GWAS pathway analyses: identification of genetic networks in dilated cardiomyopathy and ulcerative colitis. *BMC Genomics* **15**, 622 (2014).
4. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
5. Sporns, O., Tononi, G. & Kötter, R. The Human Connectome: A Structural Description of the Human Brain. *PLoS Comput. Biol.* **1**, e42 (2005).
6. Gordon, E. M. *et al.* Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. *Cereb. Cortex* **26**, 288–303 (2016).
7. Power, J. D. *et al.* Functional Network Organization of the Human Brain. *Neuron* **72**, 665–678 (2011).
8. Gratton, C. *et al.* Functional Brain Networks Are Dominated by Stable Group and Individual Factors, Not Cognitive or Daily Variation. *Neuron* **98**, 439-452.e5 (2018).
9. Forman, S. D. *et al.* Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold. *Magn. Reson. Med.* **33**, 636–647 (1995).
10. Friston, K. J., Worsley, K. J., Frackowiak, R. S. J., Mazziotta, J. C. & Evans, A. C. Assessing the significance of focal activations using their spatial extent: Assessing Focal Activations by Spatial Extent. *Hum. Brain Mapp.* **1**, 210–220 (1994).

11. Vieira, S., Pinaya, W. H. L. & Mechelli, A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci. Biobehav. Rev.* **74**, 58–75 (2017).
12. Greene, D. J. et al. Multivariate pattern classification of pediatric Tourette syndrome using functional connectivity MRI. *Dev. Sci.* **19**, 581–598 (2016).
13. Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V. & Greicius, M. D. Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cereb. Cortex* **22**, 158–165 (2012).
14. Eggebrecht, A. T. et al. Joint Attention and Brain Functional Connectivity in Infants and Toddlers. *Cereb. Cortex* **27**, 1709–1720 (2017).
15. Wheelock, M. D. et al. Altered functional network connectivity relates to motor development in children born very preterm. *NeuroImage* **183**, 574–583 (2018).
16. Wheelock, M. D. et al. Sex differences in functional connectivity during fetal brain development. *Dev. Cogn. Neurosci.* **36**, 100632 (2019).
17. Rudolph, M. D. et al. Maternal IL-6 during pregnancy can be estimated from newborn brain connectivity and predicts future working memory in offspring. *Nat. Neurosci.* **21**, 765–772 (2018).
18. Wheelock, M. D. et al. Altered brain-behavior relationships underlie attention impairments in very preterm children. *Cereb. Cortex* **31**, 1383–1394 (2021).
19. Maron-Katz, A., Vaisvaser, S., Lin, T., Hendler, T. & Shamir, R. A large-scale perspective on stress-induced alterations in resting-state networks. *Sci. Rep.* **6**, 21503 (2016).
20. Marrus, N. et al. Walking, Gross Motor Development, and Brain Functional Connectivity in Infants and Toddlers. *Cereb. Cortex* **28**, 750–763 (2018).
21. Feczko, E. et al. Subtyping cognitive profiles in Autism Spectrum Disorder using a Functional Random Forest algorithm. *NeuroImage* **172**, 674–688 (2018).
22. Thomason, M. E. et al. Prenatal lead exposure impacts cross-hemispheric and long-range connectivity in the human fetal brain. *NeuroImage* **191**, 186–192 (2019).

23. McKinnon, C. J. *et al.* Restricted and Repetitive Behavior and Brain Functional Connectivity in Infants at Risk for Developing Autism Spectrum Disorder. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **4**, 50–61 (2019).
24. Marek, S. *et al.* Identifying reproducible individual differences in childhood functional brain networks: An ABCD study. *Dev. Cogn. Neurosci.* **40**, 100706 (2019).
25. Shehzad, Z. *et al.* A multivariate distance-based analytic framework for connectome-wide association studies. *NeuroImage* **93**, 74–94 (2014).
26. Sharma, A. *et al.* Common Dimensional Reward Deficits Across Mood and Psychotic Disorders: A Connectome-Wide Association Study. *Am. J. Psychiatry* **174**, 657–666 (2017).
27. Raichle, M. E. The Brain’s Default Mode Network. *Annu. Rev. Neurosci.* **38**, 433–447 (2015).
28. Zalesky, A., Fornito, A. & Bullmore, E. T. Network-based statistic: Identifying differences in brain networks. *NeuroImage* **53**, 1197–1207 (2010).
29. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52**, 1059–1069 (2010).
30. Bahrami, M., Laurienti, P. J. & Simpson, S. L. A MATLAB toolbox for multivariate analysis of brain networks. *Hum. Brain Mapp.* **40**, 175–186 (2019).
31. Ginestet, C. E., Fournel, A. P. & Simmons, A. Statistical network analysis for functional MRI: summary networks and group comparisons. *Front. Comput. Neurosci.* **8**, (2014).
32. Simpson, S. L., Lyday, R. G., Hayasaka, S., Marsh, A. P. & Laurienti, P. J. A permutation testing framework to compare groups of brain networks. *Front. Comput. Neurosci.* **7**, (2013).
33. Nichols, T. & Hayasaka, S. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res.* **12**, 419–446 (2003).
34. Bellec, P. *et al.* Impact of the resolution of brain parcels on connectome-wide association studies in fMRI. *NeuroImage* **123**, 212–228 (2015).
35. Rosvall, M. & Bergstrom, C. T. Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci.* **105**, 1118–1123 (2008).

36. Thomas Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
37. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
38. Shen, X., Tokoglu, F., Papademetris, X. & Constable, R. T. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage* **82**, 403–415 (2013).
39. Craddock, R. C., James, G. A., Holtzheimer, P. E., Hu, X. P. & Mayberg, H. S. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* **33**, 1914–1928 (2012).
40. Ackermann, M. & Strimmer, K. A general modular framework for gene set enrichment analysis. *BMC Bioinformatics* **10**, 47 (2009).
41. van den Heuvel, M. P. *et al.* Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations. *NeuroImage* **152**, 437–449 (2017).
42. Xia, C. H. *et al.* Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat. Commun.* **9**, 3003 (2018).
43. Mootha, V. K. *et al.* PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
44. Zahn, J. M. *et al.* Transcriptional Profiling of Aging in Human Muscle Reveals a Common Aging Signature. *PLoS Genet.* **2**, e115 (2006).
45. Chen, J. J., Lee, T., Delongchamp, R. R., Chen, T. & Tsai, C.-A. Significance analysis of groups of genes in expression profiling studies. *Bioinformatics* **23**, 2104–2112 (2007).
46. Newton, M. A., Quintana, F. A., den Boon, J. A., Sengupta, S. & Ahlquist, P. Random-set methods identify distinct aspects of the enrichment signal in gene-set analysis. *Ann. Appl. Stat.* **1**, (2007).

47. Yaari, G., Bolen, C. R., Thakar, J. & Kleinstein, S. H. Quantitative set analysis for gene expression: a method to quantify gene set differential expression including gene-gene correlations. *Nucleic Acids Res.* **41**, e170–e170 (2013).
48. Efron, B. & Tibshirani, R. On testing the significance of sets of genes. *Ann. Appl. Stat.* **1**, (2007).
49. Seitzman, B. A. *et al.* A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum. *NeuroImage* **206**, 116290 (2020).