From c-Fos density to functional networks: introducing ProgBCT for the rodent brain connectivity analysis

Théo Brunel, Mathias Cavelius, Marc Thévenet, Anne Didier.

Abstract

Understanding the complex connectivity between neuronal populations remains central to elucidate brain function. In rodents, brain connectivity has been studied using c-Fos cellular mapping, c-Fos density correlation matrices and tools of the graph theory. Here, we introduce ProgBCT, a computational pipeline to easily extract connectivity patterns from c-Fos-positive cell counts in the brain. ProgBCT substantially enhances existing tools by integrating detailed preprocessing, advanced statistical validation, comprehensive graph-theoretical analyses, and intuitive visualizations. We provide a detailed description of the pipeline using c-Fos data obtained in the mouse and outline extensive applicability in neurobiological research.

1. How to start?

This manual provides every information and details on how to use the « ProgBCT » program to analyze functional connectivity between brain structures. "BCT » comes from « Brain Connectivity Toolbox ». The program is scripted using MATLAB.

The first step is to verify that data are in the right format and that default parameters are well defined.

1.1. Format of original data file

Original data file is an Excel spreadsheet typically containing values of cell density in every structure. The first row corresponds to the names of columns. First columns are the experiment characteristics (Group, Side of the brain, Animal, etc), the others correspond to brain structures (ROI1 to ROI10) whereas every raw is an animal. Excel table can contain cell density values from only one hemisphere (left in this example, or mean densities of the two) (Mono mode), or cell density values of the two hemispheres (Bi mode). Specific formats relative to Mono or Bi mode are presented below:

1. Mono mode: density of one hemisphere only (6 animals, 10 structures).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Groupe	Side	Animal	ROI1	ROI2	ROI3	ROI4	ROI5	ROI6	ROI7	ROI8	ROI9	ROI10
2	G1	Left	1	0,00018953	0,00011558	0,00164851	0,0005088	0,00026053	4,8953E-05	6,8776E-05	0,00059659	0,00031211	0,00013084
3	G1	Left	2	0,00011803	0,00015883	0,00183167	0,00045792	0,00016002	5,3809E-05	3,8975E-05	0,00057202	0,00042856	0,00025593
4	G1	Left	3	0,00014194	0,00018779	0,00201484	0,00041629	0,00028456	0,00011996	0,00010355	0,00114854	0,00020309	0,0002354
5	G1	Left	4	0,00017884	0,00050175	0,00256434	0,00032708	0,00027043	0,00023362	0,00015773	0,00090434	0,00125268	0,00021347
6	G1	Left	5	0,00014739	0,00015992	0,00238118	0,00035225	0,00017894	6,0946E-05	4,1632E-05	0,00032409	0,00070161	0,00021884
7	G1	Left	6	0,00024476	0,0006083	0,00265593	0,00031581	0,00040551	0,00033701	0,00019284	0,00065039	0,00094319	0,00049069

Raws 2 to 7 contain cell density values for 6 animals and 10 brain regions. First 3 columns mention respectively the experimental group, the brain's side and the mice label. D to H columns contain cell density for different structures (from ROI 1 to 10). Parameters to enter in ProgBCT interface are presented below.

2. Bi mode: Density of the two hemispheres (6 animals and 10 structures).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Groupe	Side	Animal	ROI1	ROI2	ROI3	ROI4	ROI5	ROI6	ROI7	ROI8	ROI9	ROI10
2	G1	Left	1	0,00018953	0,00011558	0,00164851	0,0005088	0,00026053	4,8953E-05	6,8776E-05	0,00059659	0,00031211	0,00013084
3	G1	Left	2	0,00011803	0,00015883	0,00183167	0,00045792	0,00016002	5,3809E-05	3,8975E-05	0,00057202	0,00042856	0,00025593
4	G1	Left	3	0,00014194	0,00018779	0,00201484	0,00041629	0,00028456	0,00011996	0,00010355	0,00114854	0,00020309	0,0002354
5	G1	Left	4	0,00017884	0,00050175	0,00256434	0,00032708	0,00027043	0,00023362	0,00015773	0,00090434	0,00125268	0,00021347
6	G1	Left	5	0,00014739	0,00015992	0,00238118	0,00035225	0,00017894	6,0946E-05	4,1632E-05	0,00032409	0,00070161	0,00021884
7	G1	Left	6	0,00024476	0,0006083	0,00265593	0,00031581	0,00040551	0,00033701	0,00019284	0,00065039	0,00094319	0,00049069
8	G1	Right	1	0,00190332	0,00111283	0,00160673	0,00069974	0,00104606	0,00096527	0,00105279	0,00111915	0,00074142	0,00018544
9	G1	Right	2	0,00229928	0,00190281	0,00201596	0,00069898	0,00151513	0,00128073	0,00076665	0,00097169	0,00012196	0,0007751
10	G1	Right	3	0,01405825	0,01749195	0,01569179	0,00411487	0,01626732	0,00852895	0,00903935	0,01222469	0,01263257	0,00141245
11	G1	Right	4	0,00058518	0,00202887	0,00401336	0,00107906	0,00295256	0,00343077	0,00365594	0,0015038	0,00105325	0,00106063
12	G1	Right	5	0,0039047	0,00590055	0,00398105	0,00096562	0,00385096	0,00276459	0,0031599	0,00253547	0,00303299	0,00127186
13	G1	Right	6	0,00583641	0,01327302	0,03120307	0,01351192	0,01077083	0,01924398	0,02300815	0,01505037	0,01645221	0,00228756

Columns D to H still contain information relative to 10 structures (from ROI1 to 10) and column B mentions the laterality (left or right hemisphere). Data of the left hemisphere must be presented first (as represented in raw 2 to 7) and followed by data of the right hemisphere for the same animals (raw 8 to 13). This is mandatory for Bi mode. However, in Excel spreadsheet, columns relative to the experiment information (A, B or C in the examples above) are not mandatory and unlimited as they are not retrieved in the program.

1.2. Initial program launch

To launch the ProgBCT program, navigate to right folder and double-click the MATLAB file **ProgBCT.m**. This opens the MATLAB interface and displays the source code in the editor. Alternatively, type **ProgBCT** in the **Command Window** below the editor window. The original data file is displayed in red at the top corner of the interface. Below is the description of the first items:

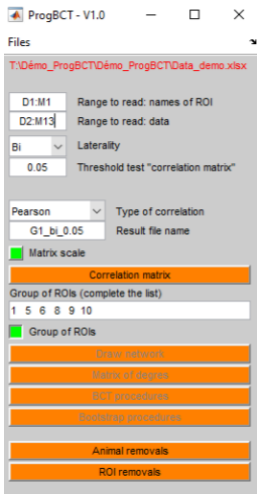
Bar menu / menu **[Files]** enables to select original data file (original Excel spreadsheet), leave program (**[Leave]**), compare two or N connectivity networks or matrices from their **Data_Excel_corr** file (but networks must be computed first).
[Cells to read: name of the structures]: select columns in the Excel spreadsheet. For the first example, put "D1:M1" to analyze all ROI.
[Cells to read: data]: select data in the Excel spreadsheet. All data can be selected ("D2:M7" in the first example or "D2: M13" in the second one).
[Laterality]: select structures of one hemisphere (Mono) or the two (Bi). **Warning:** data file must be organized accordingly (see § 1.1). **Warning:** Left hemisphere always before the right one.
[Statistic threshold]: choose the threshold value of the statistical test (0.05 for instance).
[Correlation type]: choose Pearson or Spearman test according to data.
[Result file nomenclature]: enter the name used to generate result file. It will be generated in the same folder as the original data file.

Once first parameters are set-up, first computations can be done. All these parameters also have a default mode.

1.3. Default parameters file

Parameters can be predefined in the configuration file **PBCT_ParametresParDefaut.m**. To change the default parameters, edit the file and modify the corresponding raw.

```
% Parameters linked to the data
%==============================
% Original data file name:
% Folder and Excel file name
DataRep = 'C:\Data\Data_Density\';
DataFile = 'data_G1.xlsx';
% Cells to read: Names of the structures followed by the data
GroupeStr = 'D1:M1';
GroupeData = 'D2:M13';
% Laterality mode: 1=Mono / 2=Bi (2 hemispheres)
ModeMonoBi = 1;
% Correlation type : 1=Pearson  / 2=Spearman
TypeCorr = 1;
% Statistical treshold
PG_SeuilTest = 0.05;
% Generic name for result files
title = 'Example_ProgBCT_0.05';
```



3. ProgBCT interface window.

Warning: Strict adherence to syntax is essential because this is a MATLAB file that will be executed when the program starts. Any syntax error will result in a runtime error.

2. Computation of the correlation matrix

One the parameters are entered, calculations can be initiated, starting with the computation of the correlation matric via **[Compute correlation matrix].** At this stage, other buttons remain disabled as their use depends on the outcome of the matrix calculation. The result is a correlation matrix of activity (c-Fos positive cell density) between all possible pairs of regions of interest, without directions or direct anatomical connections.

Two outputs are generated: an Excel file containing the numerical values and a correlation matrix represented as a heatmap.

2.1. Result File

Result file is an Excel spreadsheet with a name generated automatically (see le §2.2). If the chosen name is **'Example_ProgBCT_0.05'**, then file will be named **'Example_ProgBCT_0.05_Corr.xlsx'**. It contains 4 sheets:

Sheet 1: every correlation value with pvalue lower than the previously defined threshold (see §2.2). Every raw show in column A (ROI1) and B (ROI2) the two correlated structures.
Sheet 2: computations linked to the first sheet such as the percentage of correlated regions and the mean and standard deviation of the correlation coefficients (with correlations exhibiting pvalue lower than the defined threshold).

Sheet 3: all possible correlations regardless of pvlaue threshold= non-thresholded correlations.
Sheet 4: sum of positive and negative correlation coefficients for each ROI.

2.1.1. Laterality mode: Mono

In the case of only one hemisphere (Mono), in sheet 1 and 3, every raw of the C and D columns has the value of the correlation and the pvalue regarding the 2 corresponding structures.

4. Excel spreadsheet, Sheet 1, Mono mode thresholded correlations.

	A	B	C	D
1	ROI1	ROI2	corr	pvalue
2	ROI1	ROI5	0,88115359	0,02034738
3	ROI2	ROI3	0,83705676	0,03766264
4	ROI2	ROI6	0,98173285	0,00049749
5	ROI2	ROI7	0,93758532	0,00572182
6	ROI2	ROI9	0,81968019	0,04584128
7	ROI3	ROI4	-0,9929611	7,4145E-05
8	ROI3	ROI9	0,85831182	0,02869108
9	ROI4	ROI9	-0,8192596	0,0460485
10	ROI5	ROI6	0,85107929	0,03161473
11	ROI5	ROI7	0,9004659	0,01436751
12	ROI6	ROI7	0,96996392	0,0013397

The number of raw of the first sheet therefore depends on the statistical threshold. If N is the total number of structures, then the maximum number of raw (without head row) is N x (N+1) /2. This total number of raw is retrieved in Sheet 3.

2.1.2. Laterality mode: Bi

Structures are renamed for each hemisphere with an extension « L » for left and « R » for right and analysis is processed for inter and intra-hemisphere connectivity.

In Sheet 1, A and B columns correspond to the extended names of the pairs of structures. Column C (hemisphere) indicate whether the correlation is computed between structures of the same (intra) or different (inter) hemispheres while column D (side), specifies the corresponding side for intra correlation (or null if the correlation is inter). Finally, columns E and F respectively display the correlation value and the pvalue of the correlation; only correlations with a pvalue below the selected statistical threshold are listed.

5. Excel spreadsheet, Sheet 1, Bi mode.

	A	B	C	D	E	F
	ROI1	ROI2	hemisphere	side	corr	pvalue
1	ROI1_L	ROI5_L	intra	left	0,88115359	0,02034738
2	ROI2_L	ROI3_L	intra	left	0,83705676	0,03766264
3	ROI2_L	ROI6_L	intra	left	0,98173285	0,00049749
4	ROI2_L	ROI7_L	intra	left	0,93758532	0,00572182
5	ROI2_L	ROI9_L	intra	left	0,81968019	0,04584128
...	...	...	...	...	...	...
7	ROI8_R	ROI9_R	intra	right	0,99800398	5,9722E-06
8	ROI8_R	ROI10_R	intra	right	0,83495114	0,03861363
9	ROI9_R	ROI10_R	intra	right	0,85162227	0,03139059

In the second sheet, the different computations are retrieved (percentage of correlations, mean correlation coefficient and standard deviation) for each type (intra left, intra right, inter and total).

6. Excel spreadsheet, Sheet 2, Bi mode (calculation's results). Sheet 3 contains every correlation of the matrix, regardless of the statistical threshold.

	A	B	C	D
1		N%	R_mean	R_std
2	intraL	20	0,56589172	0,73069358
3	intraR	50,9090909	0,93112336	0,0481309
4	inter	11	0,8819229	0,03304487
5	Total	23,8095238	0,8399483	0,36396112

7. Sheet 4 contains the sums of positive and negative correlation coefficients for each ROI.

	A	B	C
1	NomStruct	ForcePos	ForceNeg
2	ROI1_L	0,88115359	0
3	ROI2_L	3,57605512	0
4	ROI3_L	1,69536858	0,99296111
5	ROI4_L	0	1,81222072
6	ROI5_L	6,12027213	0
7	ROI6_L	4,46497669	0
8	ROI7_L	2,80801513	0
9	ROI8_L	0	0
10	ROI9_L	1,67799201	0,81925962
11	ROI10_L	4,5513779	0
12	ROI1_R	1,86096113	0
13	ROI2_R	3,76560808	0
14	ROI3_R	7,53238562	0
15	ROI4_R	7,39239532	0
16	ROI5_R	3,72114751	0
17	ROI6_R	8,36911756	0
18	ROI7_R	8,34724321	0
19	ROI8_R	7,34950314	0
20	ROI9_R	7,38293021	0
21	ROI10_R	6,12276832	0

Warning: correlation computations can be done with different statistical thresholds. In case of a new threshold is set, the previously generated result file will be erased, or the name file must be changed.

2.2. Correlation matrix

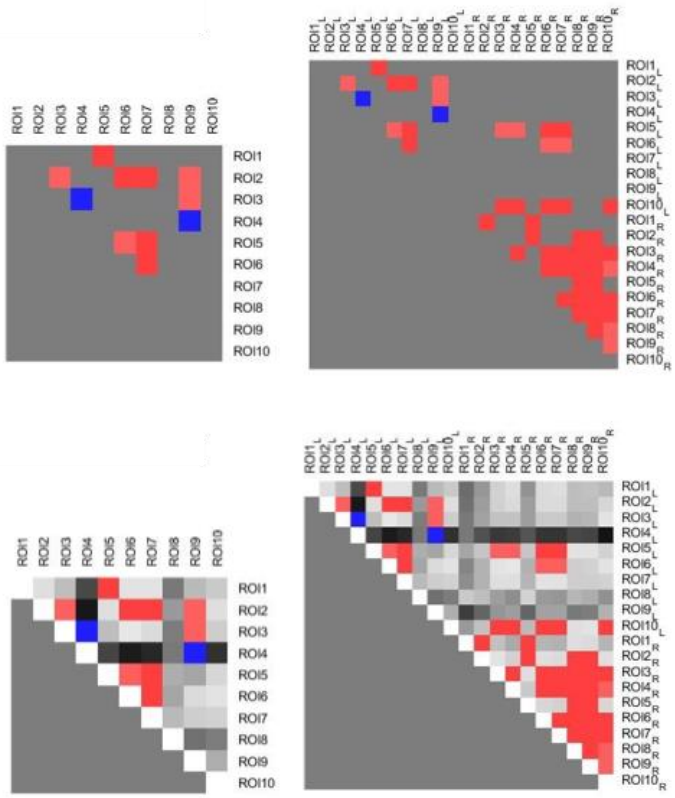
Again, it is necessary to distinguish between the unilateral (Mono) and the bilateral (Bi) modes. In the unilateral mode, only intra-hemispheric correlations are present. The correlation map then displays an upper diagonal matrix with the following color coding. Background color: either a grayscale gradient reflecting the correlation values, ranging from -1.0 (black) to 1.0 (white), with 0.0 represented as medium gray; or a fixed color. This choice is made through default parameters. Overlay color: if the pvalue of a cell is below the set threshold, the cell is colored using a gradient depending on the correlation value. A gradient from dark blue to light

blue is used for negative values (from -1 to 0), and a gradient from light red to dark red is used for positive values (from 0 to 1)

In the bilateral mode (Bi), the matrix is more complex as structures are doubled (left and right).

The upper-left quarter of the matrix corresponds to intra-hemispheric correlations of left hemisphere structures (diagonal matrix); the upper-right quarter corresponds to inter-hemispheric correlations (square matrix); and the lower-right quarter corresponds to intra-hemispheric correlations of right hemisphere structures (diagonal matrix).

8. Correlation matrices (maps) on the left for Mono mode and on the right for Bi mode. Blue indicates negative correlations, red indicates positive correlations and grey indicates non-significant correlations. In the matrices below, the intensity of grey reflects the strength of non-significant correlations.



[Display matrix scales] option allows visualization in a separate window of both the background scale (grayscale) and the overlay scale (color), each ranging from -1 to 1, corresponding to the extreme correlation values.

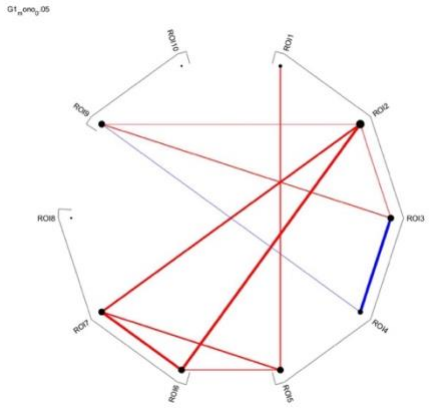
3. – Representing functional connectivity networks

Once the correlation matrices are complete, the next step is to plot the corresponding network in the form of a circle, on which the different structures will be positioned. To initiate the calculations, press the **[Plot network]** button.

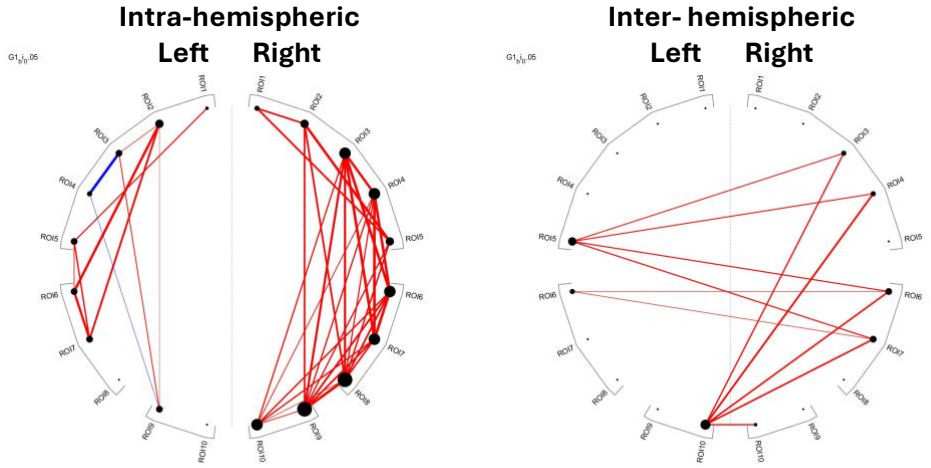
3.1. Data plotting

Around the periphery of the circle are positioned the various studied structures. The sizes of the markers representing the structures (black circles) are directly proportional to the number of connections this structure is showing (degree). Connections are depicted by lines whose thickness corresponds to the correlation coefficient value (the thicker the line, the stronger the correlation). A red line indicates a positive correlation, while a blue line indicates a negative correlation.

9. In Mono mode, only on circle will be draw, with structures evenly distributed along its perimeter, as illustrated in the example.



10. In Bi mode, two circles are plotted: one for intra data and the other for inter ones, as in the following example.



In this example, the central line separates the two hemispheres. The left figure (intra-hemispheric correlations) depicts connections each hemisphere and the right figure (inter-hemispheric correlations), depicts connections linking one hemisphere to the other.

3.2. Plotting options

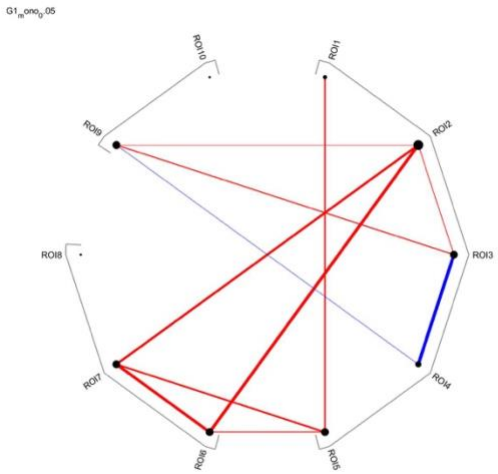
Various visualization options can be modified through the configuration file **PBCT_ParametresParDefaut.m**. To adapt the default settings, edit this file, adjust the desired lines, and restart the program. Below is an overview of the adjustable parameters in the MATLAB file.

```
% General parameters for the different visualizations
%=========================================================
PG_couleurGriseDeFond = 32;      % Gray level number for the lower diagonal
PG_couleurGriseImage = -1;               % Gray level number for the map
background or the value -1 if the background is a gradient based on the correlation
PG_flagVisuNetwork_cercle = false;       % Circle visualization flag
PG_flagVisuNetwork_groupesROI = true;    % ROI grouping visualization flag
PG_rayon = 2;                            % Circle radius
PG_coefRayonTxt = 1.1;                   % Text offset (structure names)
PG_coefTaillePoint = 100;          % Size coefficients for plotting points
PG_coefEpaisseurLigne = 3;               % Line thickness coefficient
PG_tailleTxtStr = 15;                    % Text size for structure names
PG_tailleTxtTitre = 15;             % Text size for title text(group name)
```

It is also possible to display contours grouping set of regions together, for example those belonging to a specific functional system (memory, olfactory or else). To do this, enter in the [**ROI Grouping**] field the list of pairs defining the boundaries of the groupings to be used. If there are 10 structures and you want to group the first 5, the next 3, the last 2, and the remainder separately, enter the following values: 1 5, 6 8, 9 10 (implying the pairs [1,5], [6,8], and [9,10]). To enable this display, check the **[Display ROI grouping]** button. The ROI grouping can be preset in the default configuration file. Likewise, some visualization parameters can be modified if necessary.

```
% Regions grouping
%======================
groupageROI = [];
PG_coefRayonReg = 1.08;          % Spacing of the ROI groups' line
PG_coefRepereReg = 0.08;      % Coefficient defining the size of the marker
PG_colorReg = [0.5 0.5 0.5];     % Color of the ROI groups
PG_angleOffsetReg = 3;           % Angular offsets for start/end markers
```

11. Representation of the correlation network computed for 10 structures. Structures 1 to 5, 6 to 8 and 9 to 10 are visually grouped together by an arc of a circle as belonging to a functional system.

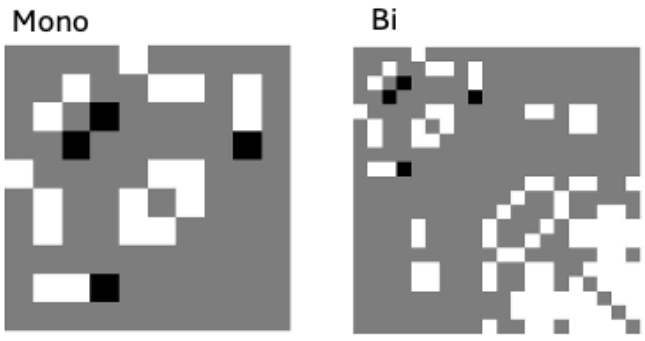


3.3. Computation of the degree matrix

To run this calculation, click the **[Compute degree Matrix]** button. The calculation uses the previously generated Excel correlation file like **"Exemple_ProgBCT_0.05_Corr.xlsx"** described in section 2.1. The resulting degree matrix is thus an N×N square matrix, where N is the number of structures studied. Only three values are possible (regarding the chosen statistical threshold): 0 if there is no significant correlation between two structures, 1 if there is a significant positive correlation, and -1 if there is a significant negative correlation. This degree matrix differs from the original correlation matrices by encoding only the sign and statistical significance of the correlations (−1, 0, or 1), whereas the original matrices contain continuous correlation values ranging from −1 to 1.

The matrix degree can be visualized by modifying the default parameter indicated below (also located in **PBCT_ParametresParDefaut.m**).

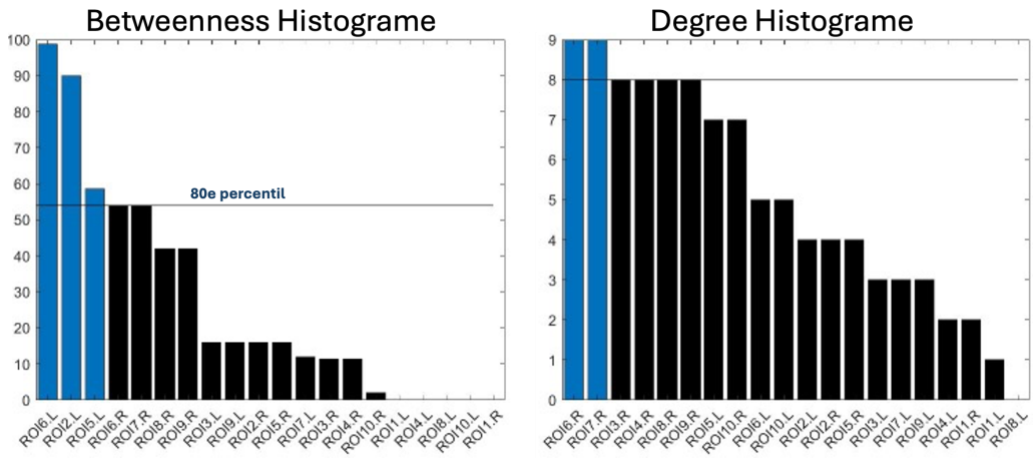
12. Representation of the matrix degree (Mono and Bi-mode).



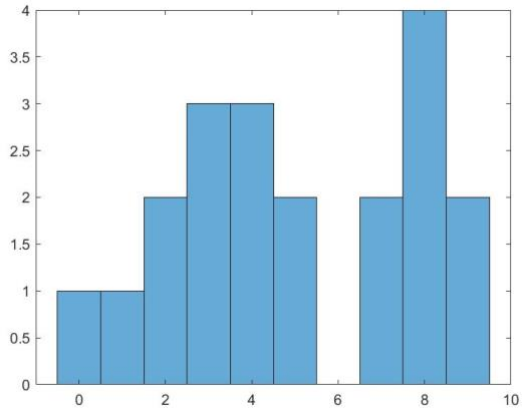
4. BCT procedures

To launch all procedures, click on **[Launch BCT procedures]** button and two graphs are generated.

13. BCT graphs: histogram of the degree distribution (left) and histogram of the betweenness distribution (right). In these graphs, blue bars highlight the regions that fall within the top 80th percentile (note: this does not mean 80% of the data, but rather the threshold below which 80% of the data fall—meaning it includes only the highest 20% of values). These ROIs, which rank in the 80th percentile for degree and betweenness, are considered hubs (Bullmore & Sporns, 2009; Rubinov & Sporns, 2010).



In the next step, an analysis of the degree distribution frequency within the network is performed, offering more details about the network properties.



14. Histogram of the degree distribution frequency

4.1. Some definitions

Graph theory provides tools to identify and characterize the network properties. In neuroscience, nodes which are brain structures, may have connections -functional correlations- that determine their organization into subnetworks or functional modules. The goal is to analyze the topological properties of these networks to apprehend their overall functioning and enable comparisons across different states or experimental conditions. Measured parameters are defined in BCT and slightly expanded definitions are proposed below.

Degree: This measures the sum of direct connections (edges) a node has with others in the network. A high degree indicates that the structure occupies a central position in the information flow.

Centrality: This evaluates the importance of a structure within a network, based not only on direct connections but also on its strategic position. A node is central in terms of betweenness if two other structures have to go through it to pass information. Such a node acts as a control point or bridge in the network. The betweenness centrality of a node k is the sum of shortest paths between any two nodes that pass-through k. This role as an intermediary or relay can make it particularly important for network organization.

Hubs: These are key structures in a network that are both highly connected and strategically positioned. A region is considered a hub if it ranks in the top 80th percentile for both degree and centrality. Hubs often drive network dynamics by facilitating interactions among other structures of the system.

Clustering coefficient: This quantifies, for a given node, the likelihood that its immediate neighbors are also connected to each other. In other words, it measures the local density of connections around a node, reflecting redundancy or cohesion within its direct neighborhood. It emerges from a ratio of the number of actual connections among a node's neighbors to the total number of possible connections between them. The average clustering coefficient of a network is the mean of the individual coefficients of all active nodes, providing insight into the network's overall tendency to form tightly interconnected local clusters.

Assortativity: It measures the tendency of nodes to connect to others with similar properties such as degree (number of connections). A positive assortativity coefficient indicates that highly connected nodes tend to link directly to each other, whereas a negative coefficient suggests preferential connections between nodes of dissimilar degree. This parameter is important for understanding the hierarchical or stratified organization of a network.

Transitivity: It is a global measure of the network's tendency to form triangles, i.e., to coherently connect neighbors of a node. It represents the probability that two nodes sharing a common neighbor are themselves connected, independently of individual node degrees. Unlike the local clustering coefficient, transitivity weighs the entire graph and is less affected by sparsely connected nodes. It serves as a robust indicator of the functional cohesion of the network, reflecting link density within potential modules.

Efficiency: It measures a network's capacity to transmit information rapidly and economically. It is based on the concept of shortest paths between nodes: a network is efficient if, on average, information can flow between any two nodes in a minimal number of steps. Local efficiency evaluates this property within a node's immediate neighborhood, while global efficiency

measures the performance of the entire network. Unlike degree, which only considers direct connections, efficiency accounts for all indirect paths, providing a more systemic perspective on network function.

4.2. Network characteristics computations

ProgBCT procedure first computes the degree, clustering coefficient, global efficiency and transitivity of the experimental network (experimental data) then it will compute several times the same values for a random network. This function generates random networks while preserving the number of nodes and the degree distribution (but not the distribution strength) of the experimental network. The function also ensures that the randomized network maintains the ability for every node to reach every other node in the network (Maslov and Sheppen, 2002). The number of computations of this random network is fixed to one thousand by default but can be modified in the file. Values from this random computation will be then averaged and the 95% confident interval is calculated.

For the experimental network, a confidence interval is calculated based on a bootstrap procedure. At each iteration, a new sample is constructed by randomly selecting individuals therefore creating a new correlation matrix and its corresponding network. These bootstrap-derived networks provide an empirical distribution of topological metrics, capturing the variability expected under the assumption of random sampling. Unlike classical randomization methods altering the structure of the network itself, this approach preserves the statistical properties of the original data while assessing the robustness of the extracted network measures.

Results of computation are reported in an Excel spreadsheet with a name corresponding to the generic one (see §2.1), for instance in the file: **Exemple_ProgBCT_0.05_GT.xlsx**.

File is composed of different Sheets:

Sheet 1: column 2 contains betweenness for each structure. These values are sorted in ascending order, with the structure names listed in column 1. Column 3 mentions whether the betweenness value exceeds the 80th percentile (1) or not (0).
Sheet 2: column 2 contains degree for each structure. These values are sorted as those of the betweenness and column 3 also mentions if the value exceeds the 80th percentile (1) or not (0).
Sheet 3: contains 4 calculated values, for the initial network (raw Network) and for the means of randomized networks (raw Random). These values correspond to the following parameters:

Mean_C: mean of « clustering coefficient »
r: measures « assortativity »
E: measures « global efficiency »
Trans: measures « transitivity »

Sheet 4: contains 4 columns corresponding to the 4 previous parameters; raw stand for the calculation of each parameter which have been bootstrapped, and values are sorted in an ascending order. These values are used for computations that complete Sheet 3.

Once these initial calculations are completed, certain parameters will be extracted from Sheet 4 and added to Sheet 3.

In Sheet 3, for each of the 4 parameters previously discussed (columns B to E):

-2.5% max (x16): values found at the 97.5% percentile in Sheet 4 (i.e., the 975th row if there are 1000 randomizations, for example).

-2.5% min (row 17): values found at the 2.5% percentile in Sheet 4 (i.e., the 25th value if there are 1000 randomizations).

-CI max (row 19): difference between the values in row 16 and those in row 3.

-CI min (row 20): difference between the values in row 3 and those in row 17.

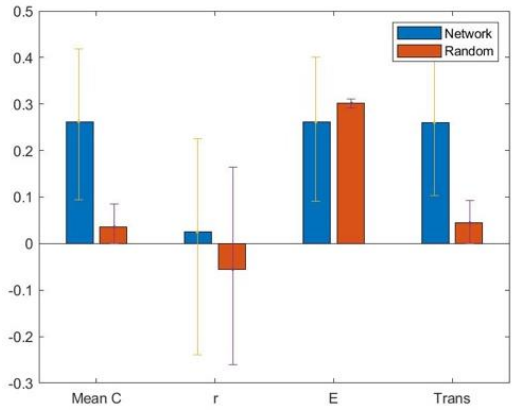
Sheet 5: contains the same 4 columns than Sheet 4; row correspond to the calculation of every randomization and values are sorted by ascending order.

4.3. Bootstrap procedures

This procedure will once again compute the 4 parameters of interest that we previously described (Mean_C, r, E and Trans) and complete the result file. This time, randomization is performed on the raw of the experimental matrix.

To assess the robustness and variability of graph-theoretical metrics describing network topology, we applied a non-parametric bootstrap procedure. This method involves generating a large number of simulated datasets (n = 10000) by random resampling from the original dataset. For each bootstrap iteration, a correlation matrix (Pearson or Spearman, depending on user selection) is computed and converted into a weighted connectivity matrix, from which a network is constructed. Several key topological metrics are then extracted from each network, including the average clustering coefficient, assortativity, global efficiency, and transitivity. The empirical distributions of these metrics across bootstrap iterations are used to derive non-parametric confidence interval, providing a robust estimate of the variability. These results are compared to the metrics computed from the original dataset and visualized as bar plots with confidence intervals, allowing evaluation of both the statistical significance and the stability of the observed topological features.

15. The graph compares the features of the experimental network with those of a randomized network generated to match the underlying attributes of the original dataset.



16. Sheet 3 of the result file of BCT procedures.

	A	B	C	D	E
		Mean_C	r	E	Trans
1	Network	0,63452381	0,49901381	0,48654135	0,68727273
2	Random	0,34471429	-0,1572465	0,55357211	0,35912727
3	Bootstrap	0,74533298	0,54543159	0,5720956	0,8320347
4					
5					
6	Network/Boot				
7	2.5% max	0,96474931	0,98266864	0,91929825	0,98120301
8	2.5% min	0,53333333	0,05400914	0,25938596	0,67213115
9					
10	CI max	0,21941633	0,43723704	0,34720265	0,14916831
11	CI min	0,21199965	0,49142245	0,31270963	0,15990355
12					
13	Random				
14	2.5% max	0,44821429	0,06508876	0,56184211	0,42545455
15	2.5% min	0,25420635	-0,3648915	0,54254386	0,29454545
16					
17	CI max	0,1035	0,22233531	0,00827	0,06632727
18	CI min	0,09050794	0,20764497	0,01102825	0,06458182

5. Modularity and Louvain communities

This procedure will generate a new Excels file named **Exemple_ProgBCT_0.05_COM**, which will serve to identify clusters and their composition thanks to the modularity and Louvain communities. Modularity is a quantitative measure that assesses the quality of a network partition by comparing the density of connections within communities to that expected in a random network. In contrast, the Louvain method is an optimization algorithm designed to detect communities by maximizing modularity, although it does not guarantee a unique solution.

The Louvain algorithm is a popular heuristic method for community detection in graphs (groups of structures more densely connected internally than with the rest of the network). It is based on optimizing a metric called modularity, which assesses how well a network can be partitioned into communities or modules. This metric quantifies the degree of functional compartmentalization within the network. High modularity indicates that the network is organized into coherent subunits where interactions occur more frequently within modules than between them. Unlike the clustering coefficient, which focuses on local connections between a node and its immediate neighbors, modularity reflects a macroscopic organization of the network. The communities identified by the Louvain algorithm reflect the modular organization of the brain, where groups of brain structures preferentially collaborate, suggesting local functional specialization within a globally integrated system. This method is valued for its efficiency and its ability to uncover communities in large networks.

The Louvain algorithm operates in two iterative phases: initially, each node is assigned to its own community, then it is moved to the neighboring community that maximizes the modularity gain. A new graph is then created by aggregating the detected communities into single nodes, and the process repeats until no further increase in modularity is detected.

This approach serves two main purposes. First, it distinguishes whether the overall network consists of a single hyperconnected unit or is composed of multiple smaller, more or less connected subnetworks. Second, within each network or subnetwork, it identifies meaningful functional clusters.

These clusters represent groups of structures with stronger connectivity, offering insight into how brain regions may be participating to the whole network and/or functionally compartmentalized.

The Excel file contains two sheets:

Sheet 1: Each column displays one module or subnetwork and raw 3 contains the number of brain regions composing the subnetwork. Subnetwork consisting of only one ROI are the one that are not connected to any other structures. If multiple such single-ROI subnetworks exist, it means that within the main functional connectivity network, several regions are disconnected from each other. If there is only one subunit, the functional connectivity network is therefore defined solely by this unified network.

Sheet 2: the organization is similar to the first sheet, but subteworks are measured using the Louvain community algorithm. Each column contains ROIs grouped by clusters. The first row mentioned the maximum modularity score for Louvain community. The modularity score ranges from 0 to 1: values

	A	B
1	Components	
2	SSReseau1	SSReseau2
3	19	1
4	ROI1_L	ROI8_L
5	ROI2_L	
6	ROI3_L	
7	ROI4_L	
8	ROI5_L	
9	ROI6_L	
10	ROI7_L	
11	ROI9_L	
12	ROI10_L	
13	ROI1_R	
14	ROI2_R	
15	ROI3_R	
16	ROI4_R	
17	ROI5_R	
18	ROI6_R	
19	ROI7_R	
20	ROI8_R	
21	ROI9_R	
22	ROI10_R	

	A	B	C	D
1	CommunityLouvain		Qcommunity	0,3272
2	SSReseau1	SSReseau2	SSReseau3	SSReseau4
3	8	8	3	1
4	ROI1_L	ROI10_L	ROI1_R	ROI8_L
5	ROI2_L	ROI3_R	ROI2_R	
6	ROI3_L	ROI4_R	ROI5_R	
7	ROI4_L	ROI6_R		
8	ROI5_L	ROI7_R		
9	ROI6_L	ROI8_R		
10	ROI7_L	ROI9_R		
11	ROI9_L	ROI10_R		

below 0.3 indicate weak community structure, scores between 0.3 and 0.6 suggest a moderate to meaningful organization, scores from 0.6 to 0.8 reflect strong modularity, and values above 0.8 may indicate very pronounced community structure.

17. The table on the left shows the composition of the networks and their subnetworks. The table on the right presents the clustering according to Louvain communities.

Finally, this analyze enable the identification of brain regions community which are strongly interconnected. They can correspond to the whole network or functional subunits inside the network, maintaining connections between communities to ensure the integration of specialized modules.

6. Suppression effect

On the interface's bottom, « ROI suppression » and « animal suppression » buttons are displayed. Each will lead to the generation of an Excel spreadsheet. The purpose of these two approaches is to assess the robustness of the network. They allow determining whether the network is consistently represented by all individuals and/or ROIS.

6.1. Animal suppression effect

This analysis aims to evaluate the impact of interindividual variability on the configuration of the functional connectivity network.

To do so, a random leave-one-out procedure is performed, whereby one subject is removed from the sample at a time, and the impact of this exclusion on network parameters is observed. The goal is to determine the extent to which the overall network represents the entire population. If removals cause minimal variation in topological measures, this suggests a robust and homogeneous network structure, not driven by a single animal. Conversely, if certain removals lead to significant fluctuations, this indicates high sensitivity to sample composition and thus reduced network stability.

It is important to note that Pearson correlations, often used to construct the network, are particularly sensitive to extreme values. As a result, some individuals may disproportionately influence the overall functional structure.

18. This table present in order: the number of suppressed animals, the maximum degree of a region, the total number of correlations; the global efficiency of the network; the mean clustering coefficient; assortativity; transitivity; the sum of connections.

	A	B	C	D	E	F	G	H	I
1	Nb	NumSuppr	MaxDeg	SumDeg	E	Mean_C	r	Trans	SumCorr
2	6	0	9	100	0,48654135	0,63452381	0,49901381	0,68727273	50
3	5	1	8	88	0,38815789	0,61083333	0,51390476	0,70588235	44
4	5	2	8	70	0,2627193	0,59642857	0,44091487	0,68181818	35
5	5	3	10	106	0,31666667	0,74583333	0,83828489	0,94475138	53
6	5	4	11	116	0,50359649	0,64949134	0,33681202	0,71498771	58
7	5	5	8	92	0,34578947	0,67488095	0,4471738	0,75675676	46
8	5	6	8	80	0,21754386	0,45	0,99346832	0,99212598	40

Network robustness is estimated based on the variability of these measures (centrality, efficiency, modularity etc) following each suppression. The variability is thus particularly critical when the sample size is small, as the influence of a single individual is amplified, potentially biasing global estimates. This approach directly challenges the interpretation and generalization of the results to the population by testing the network's dependency to one or several specific profile and/or individual.

6.2. Effet des suppressions successives des ROI's

An aother parameters used to evaluate the network's robustness is ROI suppressions. A procedure of successive ROI removals is used. Two strategies are compared:

Targeted suppression, where ROIs are eliminated in a descending order based on their values of degree and centrality (Sheet 1: Network).

Random suppression of ROIs, repeated with a bootstrap procedure to obtain a stable estimate of network decay (Sheet 2, 3 and 4: respectively Rand1, Rand2 and Rand3). A fifth sheet named RandMean compiles the means values of the columns Nb and SizeComp from the three randomized sheets (Rand1, Rand2, and Rand3), providing a summary measure of the network across bootstrap iterations.

The purpose is to measure, at every step of the procedure, the progressive decay of connectivity network characteristics. The underlying hypothesis is that a network relying on hubs (nodes with high centrality) will decay faster when these key regions are targeted first compared to a random suppression. If the network size decreases faster in the targeted condition, this suggests a hierarchical organization of the network centered around its hubs.

In each sheet, parameters are presented at every step of the suppression procedure:

Sheet 1: Targeted Suppression

	A	B	C	D	E	F	G	H	I	J
1	Nb	NomStruc	MaxDeg	SumDeg	E	Mean_C	r	Trans	SumCorr	SizeComp
2	20	Toutes	9	100	0,48654135	0,63452381	0,49901381	0,68727273	50	100
3	19	ROI6_R	8	82	0,4660401	0,61077694	0,44132175	0,63934426	41	82
4	18	ROI7_R	6	66	0,43015095	0,5962963	0,37157455	0,6	33	66
5	17	ROI3_R	5	54	0,4113708	0,5745098	0,19767442	0,56	27	54
6	16	ROI4_R	4	44	0,27388889	0,52083333	0,16455696	0,6	22	44
7	15	ROI8_R	4	36	0,26539683	0,46666667	0,17647059	0,54545455	18	36
8	14	ROI9_R	4	30	0,23754579	0,52380952	0,51612903	0,6	15	30
9	13	ROI5_L	4	24	0,19444444	0,58974359	0,33333333	0,66666667	12	24
10	12	ROI10_R	4	22	0,21464646	0,63888889	0,08333333	0,66666667	11	22
11	11	ROI6_L	3	18	0,1969697	0,51515152	0,06896552	0,69230769	9	18
12	10	ROI10_L	3	18	0,24074074	0,56666667	0,06896552	0,69230769	9	18
13	9	ROI2_L	2	12	0,16666667	0,66666667		1	6	12
14	8	ROI2_R	2	8	0,14285714	0,375	1	1	4	8
15	7	ROI5_R	2	6	0,14285714	0,42857143		1	3	6
16	6	ROI3_L	1	2	0,06666667	0			1	2
17	5	ROI7_L	1	2	0,1	0			1	2
18	4	ROI9_L	0	0	0	0			0	0
19	3	ROI4_L	0	0	0	0			0	0
20	2	ROI1_R	0	0	0	0			0	0

Sheet 2: Random Suppression

Nb	NomStruc	MaxDeg	SumDeg	MeanBC	E	Mean_C	r	Trans	MaxMcom	SumCorr
54	Toutes	7	152	190,555556	0,26075427	0,26190476	0,02456912	0,26020408	9	76
53	Audi_L	7	152	194,150943	0,27078328	0,26684636	0,02456912	0,26020408	8	76
52	Pir_L	7	150	188	0,27344741	0,28479853	0,00131283	0,2628866	7	75
51	S2_L	7	144	196,27451	0,26732341	0,27404295	0,00451272	0,25945946	8	72
50	GL_L	7	140	168,48	0,25609552	0,27952381	-9,853E-05	0,26519337	8	70
49	ACo_L	7	138	171,918367	0,26591583	0,28522838	-0,049182	0,26519337	9	69
48	PLCo_R	7	128	152,166667	0,26022338	0,22242063	0,06542056	0,23780488	10	64
47	LS_R	6	120	150,595745	0,25574996	0,19574468	0,03530895	0,2260274	9	60
46	AOB_L	6	114	102,956522	0,23212752	0,1673913	0,07166124	0,2189781	10	57
45	Pir_R	6	114	105,244444	0,24267877	0,17111111	0,07166124	0,2189781	9	57
44	Tub_R	6	110	101,045455	0,24290287	0,15530303	0,09489852	0,20930233	9	55
43	Ecx_L	5	104	109,813953	0,23067205	0,16046512	0,10582011	0,22881356	9	52
42	Ecx_R	5	100	106,714286	0,22959787	0,11587302	0,0888667	0,21428571	8	50
41	VTA_R	5	100	109,317073	0,24107776	0,11869919	0,0888667	0,21428571	7	50
40	LS_L	5	94	134,5	0,22294947	0,12916667	0,05781522	0,21	7	47
39	BLA_R	5	92	129,641026	0,22321729	0,13675214	0,04976526	0,21649485	7	46
38	Mot_R	5	86	133,631579	0,21371332	0,12719298	0,12468193	0,21176471	7	43
37	AOB_R	5	80	74,4864865	0,18071145	0,13063063	0,17155139	0,2278481	9	40
36	Gr_R	5	78	76,3888889	0,18812776	0,13425926	0,15873016	0,23076923	9	39

19. This table represent in order: the number of remaining regions; the name of suppressed structures; the maximum degree of a region; the total number of correlations; the global efficiency of the network; mean clustering coefficient; assortativity; transitivity; the sum correlations and size component.

The primary parameter examined is the size of the residual network at each step, calculated as follows: (Number of remaining correlations ÷ Number of initial correlations) × 100.

This tracking quantifies the structural resilience or vulnerability of the network by comparing the impact of random removals to that of targeting the most central nodes. If the area under the curve is greater for the targeted group than for the random group, this indicates a non-random hierarchical organization of the network.

7. Networks comparison

Another advantage of ProgBCT is the possibility to compare functional networks across different, groups or conditions. This enables to identify differences in topologies and/or co-activation patterns. By clicking on « File », several options are available such as « Data file selection » and most importantly « Compare 2 results files (_Corr.xlxs) : similarity and difference » ; « Compare N results files (_Corr.xlxs) » : similarity only» ; « Compare 2 matrix of correlations (_Corr.mat) : difference and Mantel test ».

These comparative analyses are crucial is studies aiming to characterize the effect of one experimental condition (such as stress, pharmacological treatment, enrichment or genetic mutation) on the functional brain organization.

In the examples, we illustrate matrix comparisons for Mono networks. For Bi networks, the matrices are twice as large, and in the case of circular network comparisons, both intra- and inter-group networks are computed separately for each group as well as jointly, allowing exploration of shared and group-specific patterns of connectivity.

7.1. Correlation matrices comparison

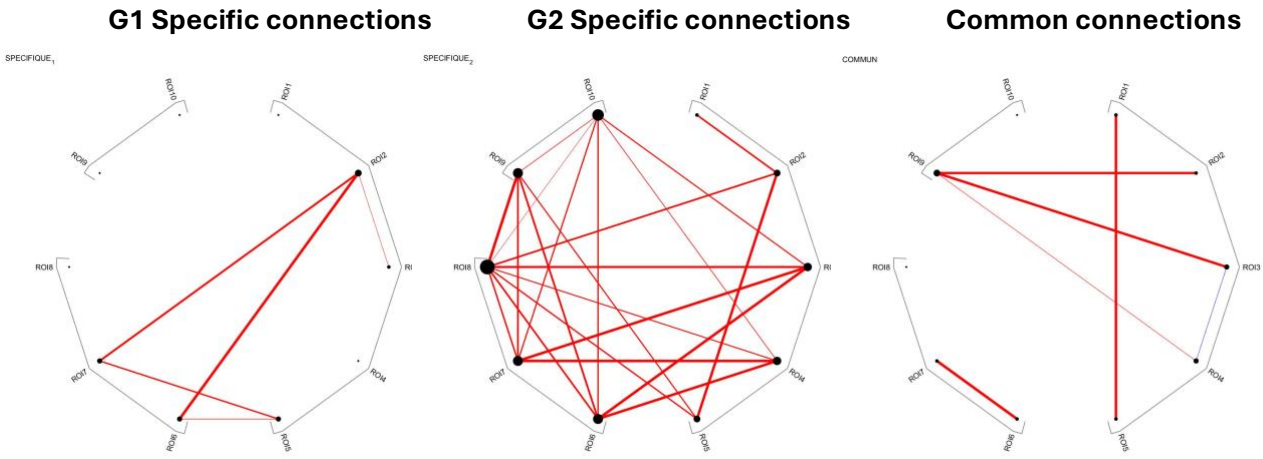
Starting from the correlation matrices generated for each group, it is possible to compare correlation coefficients region by region. The script performs a partial Mantel test with restricted permutations. Permutations are carried out while preserving the sampling group structure defined by the Euclidean distances, in order to maintain spatial dependencies. The correlation between effective distances is thus assessed while controlling for the effects of Euclidean distances, with statistical significance estimated via 1,000 Monte Carlo permutations within groups. The distance matrix is constructed by computing, for each pair of observations, a dissimilarity measure derived from the values extracted from the two input matrices. This approach allows the identification of common connections (present in both groups) as well as specific connections to a given condition.

20. On the left, it's a comparison matrix for data from a single hemisphere (Mono), and on the rightc for data from both hemispheres (Bi).

7.2. Network comparison

ProgBCT enables visualization, for each group, of two circular graphs (inter- and intra-hemispheric) where nodes are distributed by hemisphere (left/right) and edges represent significant connections as previously described. When comparing groups, multiple circular networks are generated: two networks showing common inter- and intra-hemispheric links; two networks with connections specific to group 1; and two others specific to group 2. This approach is particularly useful for studying hemispheric reorganization, functional asymmetries, or the effects of treatments on brain integration.

18. Comparison network mode mono. The two networks on the left correspond respectively to the



connections specific to Group 1 and 2 (found exclusively in on or the other group). The network on the right represents the connections shared between the networks of Group 1 and 2.

8. Conclusion

To conclude, these analyses provide a way to extract a comprehensive set of metrics are useful to extract a series of indicators describing the topology of a functional connectivity networks, such as centrality, efficiency, clustering, assortativity or modularity. These parameters enable characterization of the brain's functional organization within the context of a specific tasks, pathologies or any experimental conditions. This framework offers an interpretative foundation to relate complex brain network organization to their underlying neurobiological mechanisms trough integrative network-level analysis.

ProgBCT is an innovative, flexible and efficient methodological tool for analyzing brain functional connectivity from immunochemistry data. It delivers a rigorous solution to explore brain networks based on cellular activation, extending functional connectomics to ex vivo experiments.

Thanks to its customizable interface and compatibility with diverse experimental paradigms, ProgBCT enables the quantification of functional brain graphs, extraction of relevant topological properties, and comparison across different groups or behavioral manipulations.

This tool aligns with current efforts to extend connectomic approaches beyond temporal recordings. ProgBCT will thus help systematic analysis of functional connectivity in protocols previously less accessible to connectomics and could become a valuable resource in preclinical studies on brain plasticity, aging, pathology, or treatment effects.

9. Discussion, limits and perspectives

Nature of the data and interpretative framework. ProgBCT is based on the analysis of inter-individual correlations of c-Fos+ density between brain regions, to infer functional co-activations from post-mortem data (Wheeler et al 2013; Vetere et al., 2017). Importantly, this method does not permit causal inference or directional interpretation, as it relies solely on statistical dependencies between regions (Friston, 2011; Stephan et al., 2009). Thus, a strong correlation between the hippocampus and prefrontal cortex may be interpreted as joint engagement in a task, without implying direct influence or information flow between them. Such co-activations may also result from shared influence by a third region or from a global systemic state (Honey et al., 2009). Therefore, the generated graphs should be understood as correlational functional networks rather than representations of effective connectivity.

Extension to other markers of neuronal activation. If ProgBCT has been initially thought and conceived to manipulate data from c-Fos+ cells quantification, its modular structure makes it highly flexible and adaptable to the use of other immediate early genes (IEGs) such as Zif268/Egr1 (Bozon et al., 2003), Arc, the synaptic protein linked activity-dependent synaptic (Guzowski et al., 2001) or pCREB, the transcription factor activated by AMPc under an extracellular signal such as a neurotransmitter (Silva et al., 1998). Integrating these markers would allow exploration of neuronal functional networks across complementary temporal scales and molecular mechanisms, thereby enhancing the scope of network analyses in complex paradigms such as learning, stress, addiction, and pathology.

Data reliability and preparation. Results generated with ProgBCT strongly depend on original data quality, starting with the precision of c-Fos+ cells counting. This relies on immunochemistry quality, detection threshold and regional segmentation (Fürth et al., 2018 ; Kim et al., 2015, Midroit et al 2018). A strict standardization of these steps is crucial to ensure inter-individual comparisons and approach reproducibility. Results can also be affected by the atlas used to reconstruct the whole brain from tissue slices. Anatomical precision of region's limits directly influences the validity of density comparisons between individuals, therefore reliability of the result. Mathematically, size of experimental groups also directly influences statistical power of correlation analyses. Sample size inferior to n=8 induces more variability and reduces the capacity to detect stable and reliable co-activations (Ceyhan Ceran Serdar et al., 2021; Costantini et al. 2015; Vetere et al., 2017). It is therefore recommended to use at least 8 to 10 subjects by group to improve metrics robustness and limit bias.

Network generation and functional interpretation. When the network is being generated, it is strongly recommended to include regions specifically involved to the behavioral task of interest as non-specific regions or non-task-related. This approach grounds the functional interpretation of the

graphs in a solid neurobiological basis, while providing internal points of comparison (Bijsterbosch et al., 2019). Such strategy also enables the verification of co-activations specificity and avoids interpretation biased by an overestimation of general connectivity (Hayasaka, 2013). The absence of control regions may bias the analysis by producing an overly connected representation of the brain, regardless of the effect under study.

Graph comparisons. Graph comparisons between groups are valid only if the networks have the same number of nodes. Many metrics -such as global efficiency, clustering or centrality- depend directly on the graph's size and density (van Wijk, Stam & Daffertshofer, 2010 ; Ginestet et al., 2011). Comparing networks of different sizes introduces structural bias that can affect the results independently of the biological effect of interest. Van Wijk et al. (2010) demonstrated that efficiency and clustering tend to decrease artificially as the number of nodes increases. Therefore, it is crucial to maintain a similar number of regions across groups and to apply consistent selection criteria and data preprocessing.

Threshold sensitivity. The choice of statistical significance threshold applied to correlation matrices (e.g., $p < 0.05$, 0.025, or 0.001) directly influences the resulting graph structure by determining the network's density and the relative connectivity between regions (van Wijk et al., 2010; Zalesky, Fornito & Bullmore, 2012). A permissive threshold might introduce noise while a stringent one may smooth biologically relevant co-activations.

A recommended practice is to compare unthresholded (weighted) and thresholded (binarized) matrices by analyzing average correlation coefficients per region or system. An appropriate threshold likely preserves the network's average structure while eliminating weak or non-significant links. It is also advisable to assess complementary indicators: the effective correlation proportion, defined as the number of significant connections for a given region relative to the total number of connections in the network; and the theoretical connectivity proportion, as the number of significant connections for a region divided by the maximum possible connections (N–1). These two measures help detect under- or over-connectivity effects related to thresholding and provide a robust comparative basis for group analyses (Rubinov & Sporns, 2010). Such indicators offer a way to verify that observed connectivity differences are not merely artifacts of threshold choice but reflect genuine network-level effects. They help ensure the interpretability and reproducibility of graph-based analyses.

Detection and processing of influent individuals. ProgBCT allows to evaluate the impact of each animal on the structure of the functional network by reanalyzing the metrics after random or targeted exclusion of individuals. This approach of sequential exclusion reinforces the robustness of the method by identifying highly influent individuals with disproportionate contributions, especially in the presence of outliers (Wilcox, 2012).