Supplementary Information: Methods

Network Model

A network *G(N,K)* is the combination of two sets: a set of nodes (N) and a set of links (K), in which each link has two endpoints that represent a connection between two nodes. The most common representation of a network is a drawing of dots joined by lines; the way used to represent a network is trivial as long as the same relations between nodes are kept.

Model Descriptors

The number of nodes and links are the raw descriptors of network models: nodes represent the elements of the system, bones; while links represent their relations, suture contacts. Thus, a link connecting two nodes indicates the presence of a mutual relation.

The Adjacency Matrix

The adjacency matrix $A_{i,j}$ is the mathematical structure that codifies the information on nodes' connection. For undirected, unweighted networks this is a symmetric binary matrix of size $N\ x\ N$, where 1 indicates presence and 0 indicates absence of connection. The adjacency matrix defines the connectivity pattern of the network: the number and the particular distribution of connections between nodes. It also defines the neighborhood, connectivity context, of each node as all the nodes to which it connects.

$$A_{i,j} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,j} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,j} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,j} \\ a_{i,1} & a_{i,2} & a_{i,1} & a_{i,j} \end{bmatrix}$$

Element Descriptors

The element descriptors are the connectivity parameters measured for individual nodes, that is, the connectivity context that defines the role of each node in the network.

Node Connectivity

The node connectivity (Eq. 1) is the sum of connections a specific node has to other nodes in the network:

$$k_i = \sum_{i=1}^{j=n} A_{i,j} \tag{1}$$

Clustering Coefficient

The node clustering coefficient (Eq. 2) measures the presence of connections between the neighbors of a node: the amount of neighbors that are also connected between them. Formally, the clustering coefficient is the ratio between the total number of links connecting its nearest neighbors and the total number of all possible edges between all these nearest neighbors, which is $k_i(k_i-1)/2$:

$$C_i = \frac{\sum(\tau_i)}{\sum k_i(k_i-1)} \tag{2}$$

where τ_i is the number of connections between the neighbors of node i.

Shortest Path Length

The shortest path length between two nodes (Eq. 3) is a pair-wise measure of their shortest distance:

$$\ell_{i,j} = d(n_i, n_j) \tag{3}$$

where $d(n_i, n_j)$ is the minimum distance in number of links to connect nodes i and j. By default, links have length one; thus, the shortest path between two connected nodes is 1. Non-connected nodes have higher shortest path length depending on how well connected they are to others.

System Descriptors

System descriptors are parameters that capture properties of the entire network and are derived from the combination of node parameters. These descriptors are useful to compare organization between different systems: (1) biological and non-biological systems, (2) biological systems at different scales, and (3) different systems of the same kind.

Density of Connections

The density (Eq. 4) is the number of existing connections with respect to the maximum possible, which is $N(N-1)/2$:

$$Density = \frac{2K}{N(N-1)} \tag{4}$$

Density measures completeness: how many connections are realized of the total possible; a complete network has a density equal to 1.

Connectivity Distribution

The connectivity distribution (Eq. 5) is the frequency of occurrence of nodes with k connections:

$$P(k) = \frac{N_k}{N} \tag{5}$$

$P(k)$ is the probability to find a node with a given number of connections in the network. In the same manner, the cumulative connectivity distribution (Eq. 6) captures the frequency of nodes with connectivity equal or greater than k; it is a common methodological choice to avoid the statistical fluctuations of the $P(k)$ in networks with a small number of nodes (Dorogovtsev and Mendes, 2003, Appendix A):

$$P_{cum}(k) = \frac{N_{k' \geq k}}{N} \tag{6}$$

Mean Clustering Coefficient and Clustering Distribution

The mean clustering coefficient ($\overline{C}$) is the arithmetic mean of the clustering coefficient of all nodes in the network (Eq. 2). The $C(k)$ is the distribution of the clustering coefficient mean of all nodes with k connections; similarly to the $P(k)$, this distribution describes the structure of the network. For instance, a right-skewed distribution indicates the presence of highly clustered groups of nodes or *blocks* (Mason and Verwoerd, 2007).

Mean Shortest Path Length

The mean shortest path length ($\overline{L}$) is the arithmetic mean of the shortest path length between all pairs of nodes in the network (Eq. 3).

Organization Descriptors

Organizational descriptors capture the connectivity patterns that promote hierarchical and modular structures.

The Small-World Effect

The small-world effect is an organization feature of networks characterized by highly clustered nodes, like in a regular network, and a small mean shortest path length, like in a random network (Watts and Strogatz, 1998). This is a pervasive feature of many real networks; to identify the presence of the small-world effect in the network, the values of $\overline{C}$ and $\overline{L}$ are compared with those of a random network with the same number of nodes and connections (Watts and Strogatz, 1998). Small-world networks have higher values of $\overline{C}$ than regular and random networks, and similar or lower values of $\overline{L}$. However, a common problem in the analysis of the small-world effect in networks with few nodes (i.e., $n < 100$) is that the value of $\overline{C}$ cannot be sufficiently higher than for random networks; hence, the detection of the small-world is unreliable. To circumvent this problem, the analysis can be corrected so that for any number of nodes, a network is small-world if $[(\overline{C}/\overline{C}_{rand})/(\overline{L}/\overline{L}_{rand})] \geq 0.012n^{1.11}$ (Humphries and Gurney, 2008). The heterogeneity of the connectivity pattern in small-world networks is related with the emergence of modular organizations.

Community Structure

Networks have community structure if nodes are grouped into densely connected sets. In this context, the definition of module is simple: a group of nodes with more connections among them than to other nodes outside the module; however, detecting modules is very tricky. Indeed, this is an NP-complete problem (non-deterministic polynomial time problem), which means that evaluationing all possible partitions to find the best solution takes too much computational time. This is because the number of partitions on which we look for an optimal one is extremely huge. The number of possible partitions of a set of n elements is called the nth Bell number; for instance, the Bell number for a small network of 21 nodes, like the human skull, is aprox. $4.75 \cdot 10^{14}$. The use of heuristic tools overcome this problem, for example, traditional clustering techniques based on similarity/dissimilarity matrices, oriented algorithms based on prior information, and spin-glass models (see Porter et al., 2009, for an extensive review). The method used here to detect modules in skull networks is explained below.

Analysis of Modularity in Networks

A heuristic approach was used to identify connectivity modules in skull networks: an agglomerative hierarchical clustering analysis on a similarity matrix, which was derived from the connectivity pattern of all nodes in the network. The outcome is a dendrogram of nested groups of nodes. The best of all partitions was then identified using an index of quality. The following sections describe the steps taken in this thesis to carry out a modularity analysis: (1) quantification of nodes similarities, (2) agglomerative grouping of nodes, and (3) identification of the best partition.

Topological Overlap

The topological overlap (Eq. 7) is a normalized measure of similarity that estimates the extent to which two nodes connect to the same other nodes:

$$TO(n_i, n_j) = TO(n_j, n_i) = \frac{J(n_i, n_j)}{min(k_i, k_j)} \tag{7}$$

where $J(n_i, n_j)$ is the total amount of neighbors in common between two nodes and $min(k_i, k_j)$ the lowest connectivity of both nodes. Two nodes that share all their connections (with the same other nodes) have a TO of 1, whereas two nodes without any neighbor in common have a TO of 0.

The TO was the similarity matrix used in the agglomerative cluster analysis; This has been extensively used to analyze modularity in different types of networks (Ravasz et al., 2002; Solé et al., 2007), because it is expected that nodes in the same module connect to the same nodes (nodes with the same neighbors belong to the same module). This measure of similarity has been extensively used in social networks, where it is called *structural equivalence* (Breiger et al., 1975).

Hierarchical Clustering Analysis

The hierarchical clustering analysis brings together nodes with a higher TO in single branches until all nodes form one single group. After each match, the TO matrix is recalculated; grouped nodes act as a new element in the grouping process. The final outcome is a dendrogram that shows hierarchically nested partitions. We can use the inconsistency coefficient to know which partitions are consistent enough (Jain and Dubes, 1988). However, inconsistency cannot identify the best cut-off point of the dendrogram; to do this, we used a quality index based on connectivity, Newman-Girvan Q value, which is similar to an optimization factor.

Newman-Girvan Q Value

The modularity Q-value (Eq. 8) compares the within-module connectivity with a random distribution of connections, as an index of quality:

$$Q = \sum_{m=1}^{M} \left[\frac{k_s(m)}{K} - \left(\frac{k_m}{2K} \right)^2 \right] \tag{8}$$

where M is the number of modules, $k_s(m)$ is the sum of connections from nodes to the same module m, and k_m is the sum of all connections of nodes in module m (Newman and Girvan, 2004). The Q-value is calculated for each bifurcation in the dendrogram; the highest value indicates the best partition of the network (Fig. 1).

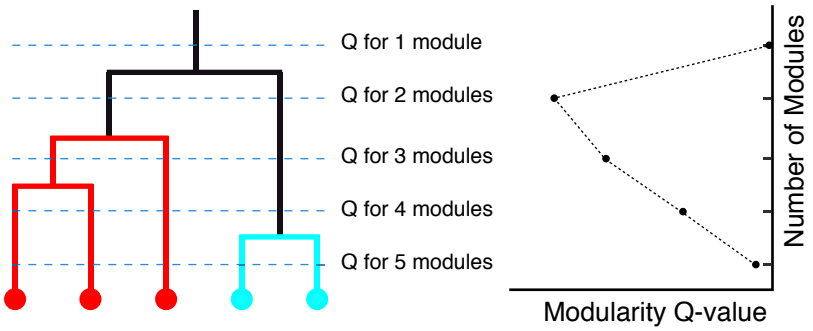

Figure 1: Schema of the identification process of the best partition. The five possible partitions are indicated with dashed lines; the modularity Q-value is measured in each bifurcation. The highest value indicates the best partition according to connections within and between modules.

Supplementary References

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97.

Breiger, R., Boorman, S., and Arabie, P. (1975). An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. *J. Math. Psychol.*, 12:328–383.

Dorogovtsev, S. N. and Mendes, J. F. F. (2003). *Evolution of Networks: From Biological Networks to the Internet and WWW*. Oxford University Press, Oxford, UK.

Esteve-Altava, B., Marugán-Lobón, J., Botella, H., and Rasskin-Gutman, D. (2011). Network models in anatomical systems. *J. Anthropol. Sci.*, 89:175–184.

Humphries, M. D. and Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PLoS One*, 3:e0002051.

Jain, A. and Dubes, R. (1988). *Algorithms for Clustering Data*. Prentice-Hall, New Jersey.

Mason, O. and Verwoerd, M. (2007). Graph theory and networks in biology. *IET Syst. Biol.*, 1:89–119.

Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E.*, 69:026113.

Porter, M. A., Onella, J.-P., and Mucha, P. J. (2009). Communities in networks. *Not. Am. Math. Soc.*, 56:1082–1097.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. (2002). Hierarchical organization of modularity in metabolic networks. *Science*, 297:1551–1555.

Solé, R. V., Valverde, S., and Rodriguez-Caso, C. (2007). Modularity in biological networks. In Képès, F., editor, *Biological Networks. Complex Systems and Interdisciplinary Science Vol. 3*, pages 21–40. World Scientific.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, 393:440–442.