Network Models in Anatomical Systems

Borja Esteve-Altava¹, Jesús Marugán-Lobón², Héctor Botella³ & Diego Rasskin-Gutman¹

1) Theoretical Biology Research Group, Institute Cavanilles for Biodiversity and Evolutionary Biology, University of Valencia, 46071 Valencia, Spain
   e-mail: diego.rasskin@uv.es

2) Unit of Palaeontology, Department of Biology, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain

3) Area of Palaeontology, Department of Geology, University of Valencia, 46100 Burjassot, Valencia, Spain

Summary - Network theory has been extensively used to model the underlying structure of biological processes. From genetics to ecology, network thinking is changing our understanding of complex systems, specifically how their internal structure determines their overall behavior. Concepts such as hubs, scale-free or small-world networks, common in the complexity literature, are now used more and more in sociology, neurosciences, as well as other anthropological fields. Even though the use of network models is nowadays so widely applied, few attempts have been carried out to enrich our understanding in the classical morphological sciences such as in comparative anatomy or physical anthropology. The purpose of this article is to introduce the usage of network tools in morphology, specifically by building anatomical networks, dealing with the most common analyses and problems, and interpreting their outcome.

Keywords - Network Theory, Modeling, Integration, Modularity.

Introduction

In recent years, network theory has been widely used as an operational framework to analyze complex systems and their relational properties (Watts & Strogatz, 1998; Barabási & Albert, 1999; for a review see Newman, 2003). Indeed, network theory has been applied to a wide range of complex biological phenomena, from the self-organization of genetic regulatory pathways to patterns of community assembly in ecosystems (Jeong et al., 2001; Sálaraz-Ciudad & Jernvall, 2002; Képés, 2007; Manson & Verwoerd, 2007; Dunne et al., 2008; Tyler et al., 2009). Furthermore, networks have been used to understand many aspects of human biology, including genetics (Franke et al., 2006), neurosciences (Sporns et al., 2004; Hagmann et al., 2008), and social relations (Milgram, 1962; Wasserman & Faust, 1994). In fact, it was in the context of studies on human social relationships where networks came into use to address biological problems, e.g. the spread of diseases (Klovdahl, 1985).

Ever since the fundamental principles of comparative anatomy were laid down in the 19th century by classic anatomists like George Cuvier, Geoffroy St. Hilaire, or Richard Owen, connections among anatomical elements have been essential for the recognition of biological homologies. Few studies have addressed the possibility of implementing an adequate methodological tool to use connectivity patterns to study morphological organization, although ideas emerged in such direction. For example, Woodger (1945) proposed a systematic use of topological information by creating a theoretical framework for the identification of anatomical parts as being distal, proximal, or articulated to others in their neighborhood, i.e. recalling
the original Geoffroy’s proposals. Mathematically, Woodger’s framework was based on the use of group theory to analyze phenotypic transformations during development, as well as for the identification of homologies. Rashevsky (1954) was a pioneer in representing the complexity of biological organization by means of graphs, a step beyond Woodger’s use of group theory. Although he briefly discussed their possible application in anatomical systems, all his work revolves around functional and physiological problems, making it difficult to apply them to purely anatomical systems. Riedl (1978) introduced graph diagrams only for the use of representation for the anatomy of a mammal skull using positional relationships (i.e. connectivity) to identify homologies. The first attempt at using graphs and network theory in comparative anatomy is Rasskin-Gutman & Buscalioni (2001) and Rasskin-Gutman (2003). The former is a descriptive exploration of the theoretical morphospace of archosaurs pelvic girdles, whereas the latter uses cellular automata to build graphs according to a series of rules that are based on how skull bones are connected on archosaurs skulls. In these works, several suggestions about complexity measures were made, along with quantitative ways to study the structural relationship among skeletal parts. In particular, Rasskin-Gutman (2003) showed a consistent approach at using graphs with vertebrate skulls.

An efficient characterization of the many levels of morphological information that describe any anatomical system ought to unveil key properties involved in their evolvability (Bastir, 2008). Here, we show a new topological methodology based on network theory that offers a suitable way to characterize complex anatomical structures, providing an independent assessment of integration and modularity issues based on connectivity rather than size and shape, and hence complementing existing morphometrics tools.

Methodology and Theory

Network Models

The simplest type of network is a set of vertices connected by edges, in which vertices correspond to the elements of a system while edges indicate the presence of an interaction that puts them in relation. Other information may be introduced in a network by using more than one type of vertices and edges, or properties associated to them. Thus, the edges can carry weights that measure the strength of the interaction, i.e. ‘weighed’ networks, or be directed to represent non-reciprocal interactions, i.e. ‘directed’ networks. Building a network model of an anatomical system of the type we propose here can provide useful information about the organization of such system. For example, the network model of a skeletal part might show a high degree of clustering indicating a specific pattern of integration and modularity that would otherwise go unnoticed (i.e., by using conventional morphometrics tools in which only size and shape information is captured, whereas connectivity information is only present collaterally when landmarks are located within sutures). Digging deeper into the properties of such a network could show the presence of a small-world structure (i.e. a particular arrangement of connections that brings closer the parts of the system), which would further indicate that some specific parts, and no others, are essential for the integrity of the system, providing important clues about its evolvability. For example, it is reasonable to expect that a highly connected bone would be more likely to be conserved than a poorly connected one, which could be lost or fused to another bone in the course of evolution. The power of network models lies in their apparent simplicity, able to capture the essential relations among their parts. In order to model skeletal systems we use unweighted, undirected networks, which will be the focus of the next sections, and from which more complicated forms can be derived.

An Anatomical Network Model

To start building the network model of an anatomical system the first step is to identify the elements of the system and the interaction we want to model. Elements and interactions must have unique definitions that allow them to be traceable in all regions of our system and, in case
we want to use networks in a comparative framework, in other related systems. However, global organizational patterns could also be subject to comparisons, thanks to the level of abstraction of network theory, e.g. to compare the internet with the neural systems of animals. In Figure 1 we sketch the kind of abstraction needed to build anatomical networks of bone connectivity. Here, the entire bones are abstracted as the nodes or vertices of the network and the sutures or physical junctions are the edges or links that connect them. Information about elements and interactions is codified in matrix form. An adjacency matrix is a square $N \times N$ matrix, where $N$ is the total number of vertices in the network, and each interaction value between two connected (i.e. adjacent) elements is noted down in the matrix cells. The standard notation is a binary code of 1 for presence of connection and 0 for absence; in networks without loops (i.e. connection from a vertex to itself) the diagonal cells are marked with 0. Some software for network analysis may demand other kinds of file formats, but all of them can be obtained from matrix files (some of these programs are listed later on in the Info on the web section). The adjacency matrix of an anatomical system is the raw material for further analyses. The degree of a vertex ($k$) is the number of edges connected to that vertex (e.g. a vertex connected to other four vertices has $k = 4$). Density is the total amount of existing connections in a network in relation to the total maximum possible, given its size. This has been taken as a measure of network complexity in metabolic and ecological networks, because a higher connectivity is related to the achievement of many functional responses in such networks. However, the way these connections are disposed provides also important information about the organizational principles that generate them.

![Fig. 1 - Abstraction process in the construction of anatomical network models. (A) We depart from an anatomical system, e.g. the human skull or a bird hindlimb, identifying our anatomical elements that will be represented by vertices. In these cases, the relation we model is the presence of a bone junction, e.g. fibrous joints in the skull and synovial joints in the hindlimb. (B) Labeled vertices and edges may be drawn as a graph. (C) Other relationships between parts, such as composition, position, orientation, shape, and size, are meaningless in the network model, and hence, vertices can be represented in any way as long as their connections are maintained properly. (D) The network connectivity pattern is codified in a binary adjacency matrix of presence/absence ready to be analyzed. 1s for presence of connection and 0s for absence. Notice that in unweighted networks adjacency matrices are symmetric.](image-url)
The Degree Distribution

The degree distribution, $P_k$, is the frequency of vertices in the network that have degree $k$. We can plot our network $P_k$ as a histogram of the degree of vertices. Although it is the simplest statistical information of network organization, it usually is enough to determine its basic properties. The functional form of this distribution is informative about the structural organization of connections, e.g. random or biased, and also about processes involved in how the network grows, e.g. preferential attachment rules. In random networks, in which edges are set with equal probability between any two vertices, the $P_k$ fits Poisson functions. In contrast, almost all biological network distributions are right-skewed, having an exponential or potential decay of frequency as degree increases. Moreover, the $P_k$ may show a different distribution if we focus in either the network as a whole or subparts of it.

Almost all automatic routines in open license software (e.g. Pajek and Network Workbench Toolkit; see Info on the web section) output directly the frequency of each degree value in the network, so the analysis of this data requires statistical packages. However, anatomical networks face statistical problems due to their small size, i.e. they have a total number of vertices of around two orders of magnitude. Because of this, goodness-fit tests to identify the functional form of a distribution may be affected. A good solution to this problem is to work with cumulative degree distributions (see Appendix II in Dorogovtsev & Mendes, 2003), which describe the number of vertices with degree greater than or equal to $k$. This procedure allows fluctuations to be less pronounced, and improves the identification of the functional forms by reducing noise in the tail of the distribution. In Figure 2 we show the whole network of a human skull of $n=21$ bones,
in which the adult mandible an isolated, not sutured bone, is not part of it. Both degree distribution profiles are shown as a way of comparison.

The Small–World Effect: Paths and Clusters

In network theory edges are also used as length units. If we consider all the edges in a network to have length 1, then the distance between two vertices is equal to the number of edges that separate them. The distance that separates any two vertices is called their ‘path length’ (e.g. vertices connected by one edge have a path length equal to 1). When there is more than one possible path between two vertices, the shortest is considered the path length. The characteristic path length of a network, $L$, is the arithmetic mean of all paths among its vertices. Surprisingly, most networks seem to be connected by a shorter $L$ than random networks with the same number of vertices and edges. This drop in $L$ occurs because of the linking up of two distant vertices; such short-cuts connect vertices that would otherwise be much farther apart than in random networks. Moreover, these new edges also increase the number of triangulations between vertices, also called vertex clustering. The clustering coefficient characterizes the density of connections in the environment close to a vertex. It is the ratio between the total number of edges connecting its nearest neighbors and the total number of all possible edges between all these nearest neighbors. The characteristic clustering coefficient, $C$, is the mean of the clustering coefficient of all vertices. Networks with lower $L$ and higher $C$ than random equivalent networks are known as small-world networks; so called by analogy with the small-world phenomenon used by Stanley Milgram (1967) to describe the structure of real social networks. In a fully connected network the characteristic clustering coefficient is equal to 1, whereas in a random equivalent network (REN), it is proportional to its size. In contrast to both extremes, small-world networks are highly clustered-like regular networks, but with a small characteristic path length like random ones (Fig. 3).

The easiest way to assess the presence of the small-world effect in an anatomical network is by comparing its $C$ and $L$ values with those of random equivalent networks; a ratio higher than 1 indicates that the network presents the small-world effect. RENs have the same number of vertices, edges, and average degree of links per node than our network. Because a random network is only one member of a statistical ensemble of all possible realizations, we have to use a sample of such RENs for our comparisons. However, as we pointed out for the degree distribution, the small size of most anatomical networks makes small-world recognition sometimes difficult. Some authors have solved this problem for real networks in an heuristic way (Dunne et al., 2002); by plotting the $C$ ratios as a function of size, they showed that small-world networks follow a unique linear ($C=edges/vertices$) relationship that diverges from the one expected for random networks ($C=edge/vertices^2$). Thus, networks that follow this linear relation would be expected to be small-world, even though, as noted before, the $C$ and $L$ ratios are not greater than 1. An anatomical network showing a small world structure will likely have few hub bones that are more connected and many poorly connected ones; if this is so, this kind of arrangement would provide robustness to random loss of bones during evolution.
Network Resilience

Many real networks show a kind of integrity robustness against vertices removal, also called network resilience. The effect of vertices deletion in network cohesiveness may be traced on different network parameters; the most common is the characteristic path length, but others may also be good measures, in particular cases such as clustering coefficient, diameter, or density. In general, any vertex deletion may lead to an increase in path lengths, and after successive deletions most network vertices become disconnected and network integrity lost. The behavior of network resilience seems to vary depending on the network degree distribution, as well as the way in which vertices are removed. For example, in networks where edges are not homogeneously distributed, and few vertices attach most of the connections (i.e. exponential and scale-free distributions), the removal of such vertices have higher impact in the characteristic path length increase than the removal of poorly connected ones. Thus, most real networks have reported to be robust against random vertex deletion, but fragile to highly-connected vertex deletion (Albert et al., 2000). Again, network size is important for statistical treatment of vertices removal because test controls require random removals, and the probability to randomly delete a highly connected node increases as network size decreases. The analysis of node removal in a network model of a skeletal system indicates which bones are more likely to be fused or lost along the course of evolution.

Detection of Modules and Motifs

A widespread feature of almost all biological networks is a structural organization in parts, or modules; groups of vertices more connected within the group than to other vertices outside the group. Although the definition of module may be very intuitive, procedures to detect them are numerous and complex, e.g. traditional hierarchical clustering, optimization algorithms of modularity values, minimizing energy functions by spin-glass methods, or topological overlap of vertices (for extensive reviews about modularity detection methods in networks see Danon et al., 2005; Fortunato, 2009; Porter et al., 2009). Differences in module detection methods come down to the precise definition of what it means to be “more connected” used to build module detection strategies and algorithms.

At a scale between vertices and modules, networks can be organized in small recurrent connectivity patterns called motifs; minimal community structures composed of a few vertices that form the building blocks of networks. In undirected networks the simplest motifs are triangles and squares, whereas in directed networks most common motifs are one to three vertices loops. Some motifs are characteristic of a particular kind of networks, such as three-vertices-chain in food webs, diamonds in neural networks, and feed-forward loops in genetic regulatory networks (Milo et al., 2002). In anatomical systems, such as the skull, triangular motifs have been found more frequently, whereas serial, squared, and pentagonal motifs are also present but less frequent (Rasskin-Gutman, 2003). It has been suggested that an iterative assembly of such motifs may be responsible of the formation of network modules at a higher scale with characteristic biological meanings, hence, many efforts are devoted to identify such motifs and develop network growing algorithms to simulate their formation (Zhang et al., 2005).

Network Models to tackle Morphological Integration and Modularity

Integration and modularity are two concepts strongly linked in biological systems as two faces of the same coin; modules appear as a consequence of heterogeneous integration in a system (Bastir, 2008). The uniting principle of modularity refers to the pattern of connectedness in which elements are grouped into highly connected groups that are more loosely connected to other such groups (Rasskin-Gutman, 2003; Wagner et al., 2007; for general reviews of the modularity concept, see Callebaut & Rasskin-Gutman, 2003; Klingenberg, 2004; 2008; Schlosser & Wagner, 2004). Similarly, the network definition
of module can match the one used in morphology, in which modules are defined as parts that have more or stronger connections among themselves than with other parts of the same system, with which they hold fewer or weaker connections (see chapters 8-11 of Callebaut and Rasskin-Gutman, 2006; and, more recently, Klingenberg, 2010). It will be very interesting to check if detected network modules and motifs fit in with morphological modules, or anatomical regions, assessed by other criteria, such as developmental origin, genetic determination, functional activity, or evolutionary behavior. Indeed, our preliminary results using such approach indicate that network modules in the human skull network strongly resemble the facial and cranial vault modules assessed by functional modules.

Whereas integration can be measured directly by network parameters values, the organization of connectivity patterns are characterized qualitatively by comparing degree distribution functions with those of theoretical models; such as Poisson, linear, exponential, or potential. For example, a distribution function that follows a power-law is characteristic of networks with a scale-free structure, whereas random networks have a Poisson distribution. The fitting of anatomical networks to any theoretical model reveal several testable properties, such as the tolerance to structural changes, i.e. network resilience, which is achieved by anatomical systems during development and evolution. These degree distribution functions are dependent on the growth rules of the networks, such as preferential attachment, i.e. the more connections a vertex already has, the more connections will attach during network growth. Thus, different network parameters allow characterizing different aspects of morphological integration during evolution and development at the level of connectivity between anatomical elements. Indeed, such organizational traits have varied during the evolution of anatomical structures. In the tetrapod skull they seem to follow a trend of increasing integration pointed out by the density of connections in the skull networks of different species (unpublished results). Moreover, we have found that tetrapod skulls fit right-skewed distributions, i.e., they are characterized by presenting few highly connected bones (hubs) and many poorly connected ones, by which the skulls achieved robustness to random failures.

Although general network organization is conservative within taxa, intraspecific variation among individuals is also present. See for example in Figure 4 the variation present in the pterion region among hominid skulls or the presence of supernumerary bones (Berry & Berry, 1967). These features characterize patterns of connectivity at an individual level; but also at a population level since some of these patterns are inheritable (Wang et al., 2006). Thus, integration and modularity in anatomical networks can be traced in population distributions, as well as in taxa evolution, within an unique comparative framework.

**Concluding Remarks**

The use of network analysis has grown in disparate fields such as genetics, systems biology,
and ecology, and this has been facilitated by the intense theoretical research in graph theory in the last decades. This multidisciplinary synergy occurs because network theory enables the study of complex systems by analyzing relational properties as connections in a simpler manner. A quantitative characterization of anatomical systems by means of network theory brings out important features of its structural organization, such as morphological integration and modularity. Since similarities in the organization of systems reflect similarities in the generative process (Wutchy et al., 2006), the use of network theory in anatomy can enhance our understanding of the generation of form during evolution and development, in many complex structures. In a broader conceptual context, networks offer a unique methodology applicable to systems at all scales, whether they are biological or non-biological. This multi-scale framework can address causal and compositional relationships between levels of complexity, helping to highlight useful information about the rules that organize them (Rasskin-Gutman & Esteve-Altava, 2009).

Acknowledgements

This research project was supported by grant (BFU2008-00643) from the Spanish Ministerio de Ciencia e Innovación.

Info on the web

There are many programs designed for analysis and visualization of networks. The following are just a small sample.

http://pajek.imfm.si/doku.php

PAJEK is a classic program for analysis and visualization of large networks. This software is mainly devoted to social networks, including several common routines such as block modeling, hierarchical clustering, or coloring. It is of easy use, with an extensive documentation.

http://nwb.slis.indiana.edu

Network Workbench Toolkit is a very friendly network analysis, modeling, and visualization tool. It performs basic network analyses (e.g. density, cluster, or paths) but more complex routines related to modularity are not implemented. It is a standalone desktop application requiring Java 1.4+ JRE, running on Windows, Mac, and Linux platforms.

http://www.mathworks.com/products/neuralnet

Neural Network Toolbox™ provides tools for designing, implementing, visualizing, and simulating neural networks in MATLAB. In order to use this platform, some previous knowledge is required. Neural Network Toolbox supports feed-forward networks, radial basis networks, dynamic networks, self-organizing maps, and many others.

http://igraph.sourceforge.net

Igraph is a free software package for creating and manipulating undirected and directed graphs. It includes implementations for classic graph theory problems like minimum spanning trees and network flow, and also implements algorithms for some recent network analysis methods, like community structure search. It can be installed as a C library, R package, Python extension module, and Ruby extension. Previous training in these platforms is needed.
References

Network models in anatomical systems


Associate Editor, Markus Bastir