Check for updates

METHOD ARTICLE

Estimation of the fractal network properties of multicellular life by cellular box-counting [version 1; peer review: 2 approved with reservations]

József Prechl

Diagnosticum Zrt, Budapest, Hungary

V1 First published: 09 Nov 2017, 6:1980
<https://doi.org/10.12688/f1000research.13070.1>

Latest published: 09 Nov 2017, 6:1980
<https://doi.org/10.12688/f1000research.13070.1>

Abstract

Multicellular life is based on the ability of cells to divide, differentiate, cooperate and die in a controlled and organised manner, generating and maintaining an organism. The temporal distribution of division, differentiation and death determines the cellular composition of the organism at any particular point in time. Like these ontogenetic events, phylogenetic development takes place with the changes in total cell numbers, the allocation of these cells to different tissues and the disappearance of certain tissues.

Fractal properties of complex networks, a result of growth, can be estimated by box counting, whereby the topological properties of the network are mapped by changing the resolution of examination, that is changing the size of the boxes used to identify and group network components. Here we develop the concept of cellular box-counting, referring to the fact that cells can be grouped on various levels of hierarchy and these various levels can be interpreted as boxes of different linear sizes. We apply the method to data representing distinct stages and groups of evolution of life and interpret the network properties of brown algae, green plants and animals. The results are in agreement with previously established values of degree exponent of biological networks and provide clues to the differences in the network organization of multicellular life.

Keywords

network, fractal, complexity, multicellular, life

Open Peer Review

Approval Status

	1	2
version 1
09 Nov 2017	view	view

1. **George W. Bassel** , University of Birmingham, Birmingham, UK

2. **Bela Suki**, Boston University, Boston, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: József Prechl (jprechl@gmail.com)

Author roles: Prechl J: Conceptualization, Data Curation, Formal Analysis, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2017 Prechl J. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Prechl J. **Estimation of the fractal network properties of multicellular life by cellular box-counting [version 1; peer review: 2 approved with reservations]** F1000Research 2017, 6:1980 <https://doi.org/10.12688/f1000research.13070.1>

First published: 09 Nov 2017, 6:1980 <https://doi.org/10.12688/f1000research.13070.1>

Introduction

Molecular pathways responsible for cellular complexity in a given multicellular organism are those that arose and have been selected during evolution leading to that organism. By going backwards in time, or alternatively by taking contemporary representative organisms of those backward steps, ancestor cells and tissues, of every organ and organ system of the examined organism can be traced. These tissues are the result of different expression patterns, different branches of molecular pathways. Going back far enough, a eukaryotic unicellular ancestor stage can be reached. By recording the relationships (lineage) and cellularity of these stages a biological network (Albert, 2005) representing the evolutionary development of multicellular organisms could be drawn. In this fractal network a node represents a cell belonging to a given tissue in the examined organism, the tissue being the virtual descendent of a single cell in an earlier organism, and the cell being the precursor of a tissue developing in an organism to appear later (Figure 1A). Expansion of a given cell type generates new nodes connected to the same hub, which represents a progenitor cell. Differentiation generates new hubs.

The structure of such a fractal network can be described by rate of change of connectivity (node degree, k_B) and the rate of change of node number (N_B) while moving along the temporal dimension, represented by l_B , the linear box size (Song *et al.*, 2007; Song *et al.*, 2006). Zooming in means decreasing l_B , looking at pathways responsible for more and more specific cell types. Zooming out, called renormalization, means drawing simpler, more general, shared molecular networks, until reaching homeostatic networks shared by all cells in the multicellular organism that are present in the common ancestor. The number of hubs at various levels of development in an organism corresponds to cellular complexity, that is levels of organization as cells, tissues, organs, and organ systems.

Methods

The properties of such a network can be estimated by using cellular complexity expressed as the number of different cells in an organism and total cell number (N) in an organism. The dataset compiled by Bell and Mooers (Bell, 1997) was supplemented with vertebrate data compiled by Schad *et al.* (Schad *et al.*, 2011) to increase the representation of complex animals, estimating the cellularity of chordates based on average weight (see Supplementary file for dataset). The network scaling relations are interpreted after Song *et al.* as follows (Song *et al.*, 2006; Song *et al.*, 2005).

Boxes provide a fractal dimension d_B that describes how relative box numbers are changing with scaling:

$$N_B(l_B)/N \sim l_B^{-d_B} \quad 1)$$

where $N_B(l_B)$ is the number of boxes identified at linear box size l_B in a network with N nodes. When l_B is equal to N , the box contains all the nodes and N_B equals 1. Consequently, plotting $\log(N_B(l_B)/N)$ against $\log(l_B)$ for a number of different networks a straight line (Figure 1B) corresponding to $d_B=1$ is obtained.

Boxes have a degree exponent d_k that describes the relationship of degrees outside and inside of the boxes (backwards and forwards in time) as scaling changes:

$$k_B(l_B)/k_{hub} \sim l_B^{-d_k} \quad 2)$$

where $k_B(l_B)$ is the degree of the most connected box and k_{hub} the degree of the most connected node inside the boxes. When l_B is equal to N then the box has no links and $k_B(l_B)$ is 1, while k_{hub} will be what is observed as complexity. Consequently,

$$1/k_{hub} \sim N^{-d_k} \quad 3)$$

Assuming that a single cell has a complexity of one and all the examined networks are descendants of this ancestral network the rule governing the generation of related (ideally linear descendant) networks can be identified by finding an average degree exponent d_k for these networks. Here linear regression weighted with $1/N$ to compensate for the underestimation of complexity in organisms with very high cell numbers (Figure 1B) was used.

Thus, the average value of d_k relative to d_B can be obtained from the equation of the linear regression. This in turn provides the degree distribution exponent gamma:

$$\gamma = 1 + d_B/d_k \quad 4)$$

using

$$\gamma = 1 + d_B/(d_B - ds) \quad 5)$$

where ds stands for the difference of slopes obtained from the regression (Figure 1C).

Results

The different relationships between fractal dimension and degree distribution exponent in three independently evolved multicellular groups should correspond to structural differences of cellular and molecular networks in these phylogenetic groups. At any value of gamma the corresponding fractal dimension is lower, in the order of brown algae, green algae and plants, and animals. Thus, the development of complexity corresponds to a decrease of difference between d_B and d_k , a trend towards $\gamma = 2$, which would represent maximal diversification with a hub linked to non-identical nodes. At any particular value of γ , plants show higher values of d_B in agreement with the observed fractal anatomy of plants. Indeed, increasing complexity corresponds to decreasing self similarity. Curves representing steps of growing complexity gradually deviate from that representing fractal dimension:

$$d_B = (\gamma - 1)/(\gamma - 2) \quad 6)$$

(Kim *et al.*, 2007), corresponding to:

$$\gamma = 1 + d_B/(d_B - 1) \quad 7)$$

Figure 1. Cellular box-counting of multicellular life forms. **A:** Evolutionary fractal network. A molecular network represented by a single cell seeds the evolution of a novel tissue, which becomes an organ system with time. By using adjusted box sizes ($I_B=N$) we normalize fractal box dimension d_B according to organism size. Complexity of the organism is then defined by k_{hub} , degree of the most connected node at a subsequent time point in evolution. Ellipses represent networks: organisms organized at various levels, the basic unit of organization being a cell. With the development of novel molecular pathways new types of cells appear; further steps, brought about by duplication or alternative splicing, refine these pathways leading to subtypes of cells in ever more complex organisms. **B:** Difference of slopes of fractal dimension and degree exponent. Complete renormalization generates a single box, which represents the organism at $I_B=N$. Here $N_B(I_B)/N$ is $1/N$ and $d_B=1$. Dots represent individual species (see [Supplementary file](#)), color lines are weighted regression lines representing groups as indicated. Double headed arrows indicate the difference in slopes (ds). Inset shows corresponding original approach described by *Song et al.* to define network fractality. **C:** Relationship between degree distribution and fractal dimension in different multicellular organisms. 95% confidence intervals of regression slope values are shown for the examined three different groups of multicellular life. Curves were generated online using [fooplot.com](#).

(Figure 1C). The observed differences are not in agreement with observations on the metabolic network of bacteria and eukaryotes, where key parameters of network topology were found to be identical (Jeong *et al.*, 2000). However, here we are not looking at conserved protein networks of metabolic activity, rather protein networks responsible for the multicellular organization of eukaryotes. Considering the dramatic differences in anatomy, physiology and metabolism in the examined groups the results are expectable.

Conclusions

Relatively simple anatomical and histological data of phylogenetically related organisms can be used to get insight into the fundamental network organisation of cells and molecules responsible for multicellularity. This method is a simple top-down approach for the investigation of cellular and molecular networks, which complements bottom-up approaches used by proteomics, metabolomics and genomics.

Further elaboration of the methodology based on network science (Jin *et al.*, 2013) and systems biology along with further refinement of phylogenetic groups, cell and cell type numbers can

provide more accurate estimations for selected organisms. Incorporation of data on the cell numbers in various tissues will allow the estimation of d_b , thereby the full description of fractal network properties. Finally, application of the method for ontologic data, examining the cellularity and complexity during the development of an organism can help draw the cell and molecular networks for any particular multicellular form of life.

Competing interests

No competing interests were disclosed.

Grant information

The author(s) declared that no grants were involved in supporting this work.

Acknowledgments

I wish to thank Krisztián Papp (MTA-ELTE Immunology Research Group, Budapest) for his help with data analysis. The book “A csodák logikája (The logic behind miracles)” by László Méréő gave me inspiration to search for the fractal properties of life.

Supplementary material

Supplementary File 1: Cellularity and complexity data. Data on the cellularity and complexity of a number of different organisms were compiled from the following two publications:

Non-vertebrate data: Bell, G. 1997. Size and complexity among multicellular organisms. *Biological Journal of the Linnean Society* 60(3), pp. 345–363. <https://doi.org/10.1006/bijl.1996.0108>

Vertebrate data: Schad, E., Tompa, P. and Hegyi, H. 2011. The relationship between proteome size, structural disorder and organism complexity. *Genome Biology* 12(12), p. R120. <https://doi.org/10.1186/gb-2011-12-12-r120>.

Taxonomical grouping follows what was used by Bell *et al.* Data on vertebrates in the first publication were replaced by more recent estimates in the second publication. Base 10 logarithm values were replaced by base 2 logarithm. Inverse values ($1/\text{complexity}$) of numbers describing cellular complexity were calculated in accordance with the proposed interpretation of k_{hub} . Cellularity of vertebrate organisms was calculated by using the average mass values shown and using conversion that assumes the mass of an average cell to be 7 nanograms (7×10^{-9} g). References, with first author and date, are shown for Bell *et al.*'s dataset.

[Click here to access the data.](#)

References

- Albert R: **Scale-free networks in cell biology.** *J Cell Sci.* 2005; 118(Pt 21): 4947–4957.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Bell G: **Size and complexity among multicellular organisms.** *Biol J Linn Soc.* 1997; 60(3): 345–363.
[Publisher Full Text](#)
- Jeong H, Tombor B, Albert R, *et al.*: **The large-scale organization of metabolic networks.** *Nature.* 2000; 407(6804): 651–654.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Jin Y, Turaev D, Weinmaier T, *et al.*: **The evolutionary dynamics of protein-protein interaction networks inferred from the reconstruction of ancient networks.** *PLoS One.* 2013; 8(3): e58134.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kim JS, Goh KI, Kahng B, *et al.*: **Fractality and self-similarity in scale-free networks.** *New J Phys.* 2007; 9(6): 177.
[Publisher Full Text](#)
- Schad E, Tompa P, Hegyi H: **The relationship between proteome size, structural disorder and organism complexity.** *Genome Biol.* 2011; 12(12): R120.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Song C, Gallos LK, Havlin S, *et al.*: **How to calculate the fractal dimension of a complex network: the box covering algorithm.** *Journal of Statistical Mechanics: Theory and Experiment.* 2007; 2007(03): P03006.
[Publisher Full Text](#)
- Song C, Havlin S, Makse HA: **Origins of fractality in the growth of complex networks.** *Nat phys.* 2006; 2(4): 275–281.
[Publisher Full Text](#)
- Song C, Havlin S, Makse HA: **Self-similarity of complex networks.** *Nature.* 2005; 433(7024): 392–395.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 17 January 2018

<https://doi.org/10.5256/f1000research.14172.r29484>

© 2018 Suki B. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bela Suki

Department of Biomedical Engineering, Boston University, Boston, MA, USA

This paper describes an analysis method and some results related to multicellular organization and complexity. While the study could be interesting, it is not well written. It is difficult to understand what is measured and what those measurements mean as described in the specific comments.

Specific comments:

1. In the introduction there are a few general statements that should be supported by general references.
2. The “eukaryotic unicellular ancestor” in the introduction represents a single cell, the ancestor of all eukaryotic cells. Is this the same as the Last Common Universal Ancestor, or LUCA? Maybe not. LUCA has been studied a lot and it appears as the ancestor of all cells including bacteria and archaea. But is there a eukaryotic ancestor as well? Is it known that there was only one ancestor and only one transition from bacteria/archaea to the eukaryotic cell? If so, is there any study that can back up such a statement?
3. Is d_B on the right hand side of equation 1 supposed to be d_B (i.e., B in subscript)? Similarly, d_k is likely supposed to be d subscript k in equation 2.
4. Please explain where equation 4 comes from.
5. Please explain better what the data represent. There is a statement that N is the total number of cells in an organism. What is then N_B ? Is it the number of different types of cell? And what is the network itself? There is a statement that “boxes have a degree exponent d_k ” (page 2, right column, first line). But what is a box? Is it a cell itself? And what is the network inside the box? Or is it that the smallest box is a single cell and inside a box can only be looked at when the box represents an organ? Yet towards the end of the Result, it is stated that “rather protein networks responsible for the multicellular organization of

eukaryotes". But I do not see any analysis on protein networks. All these questions arise because the methods is unclear and incomplete.

6. If complexity here means intercellular network complexity characterized by the degree distribution, then it is obviously different than the complexity of protein networks. It is not obvious why one should expect protein network to be self-similar with cell networks.

Is the rationale for developing the new method (or application) clearly explained?

Partly

Is the description of the method technically sound?

Partly

Are sufficient details provided to allow replication of the method development and its use by others?

No

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Biomechanics, mechanobiology, complexity

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 27 November 2017

<https://doi.org/10.5256/f1000research.14172.r27994>

© 2017 Bassel G. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

George W. Bassel

School of Biosciences, University of Birmingham, Birmingham, UK

The authors present a method to measure cellular complexity in different species by examining fractal network properties. This is potentially a very interesting method, yet there are several

things which should be addressed:

1. As this is a methods paper, it would be nice to see how this particular method of measuring cellular complexity relates to others. There is no benchmark presented for the comparison of results.
2. It may not immediately be obvious to readers how fractals relate to complexity of an organism. The method illustrated in Fig. 1A could be more clear for readers who are not familiar with fractals, showing an example of how box-counting is actually carried out, and how this relates to complexity.
3. The results presented in Fig. 1 could be discussed in greater detail. It is not clear why there is a fairly noisy relationship between fractal dimension and degree exponent between species within different evolutionary backgrounds. Could this variance be due to the data that were used?
4. Though it is possible to measure cellular complexity with this approach, it is not clear why this method is of particular value, and what purposes it would serve aside from the example presented. The addition of more relevant context in the introduction and conclusions could help with this.

Is the rationale for developing the new method (or application) clearly explained?

Partly

Is the description of the method technically sound?

Partly

Are sufficient details provided to allow replication of the method development and its use by others?

Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research