

UNIVERSAL SCALING IN PHYLOGENETIC BRANCHING

E. Alejandro Herrada^{1}, Claudio J. Tessone^{2}, Víctor M. Eguíluz^{1}, Emilio Hernández-García^{1}, and Carlos M. Duarte^{3}

^{1} IFISC, Instituto de Física Interdisciplinar y Sistemas Complejos (CSIC-UIB), Palma de Mallorca, Spain
^{2} ETH Zurich, Systemgestaltung, Zurich, Switzerland
^{3} IMEDEA, Instituto Mediterráneo de Estudios Avanzados (CSIC-UIB), Esporles, Mallorca, Spain



ABSTRACT

The evolutionary patterns of diversification of life in the planet are summarized in the concept of the Tree of Life, a synoptic depiction of the pathways of biological differentiation reconstructed by phylogenetic methods. Its main branches separate the major kingdoms and the terminal apices include the lower taxonomic levels down to species, and thanks to the modern development of high-resolution markers, even down to populations and other intraspecific entities.

Here we examine the topological properties of different portions of the Tree of Life. We quantify them in terms of quantities developed for the study of complex trees and networks, and show that the branching patterns conform to allometric rules, relating phylogeny size and form, greatly conserved among kingdoms and also from the main to the terminal branches. Our findings suggest the action of similar evolutionary forces driving phylogenetic diversification across a broad range of scales, from macro-evolution to speciation and population differentiation, which turn out to give tree patterns quite distinct from the ones produced by the simplest random branching models.

Phylogenies

- 134 phylogenies: 67 intra- and 67 inter-specific phylogenies:
 - Organisms (Virus, Monera, Protocista, Fungi, Plantae and Animalia).
 - Environments (terrestrial, marine and fresh water).
 - Climatic regions (from polar to desertic).
 - Phylogenetic tree reconstruction methods (neighbour-joining, maximum parsimony and maximum likelihood).
- TreeBASE [2]: 4,020 phylogenetic trees.

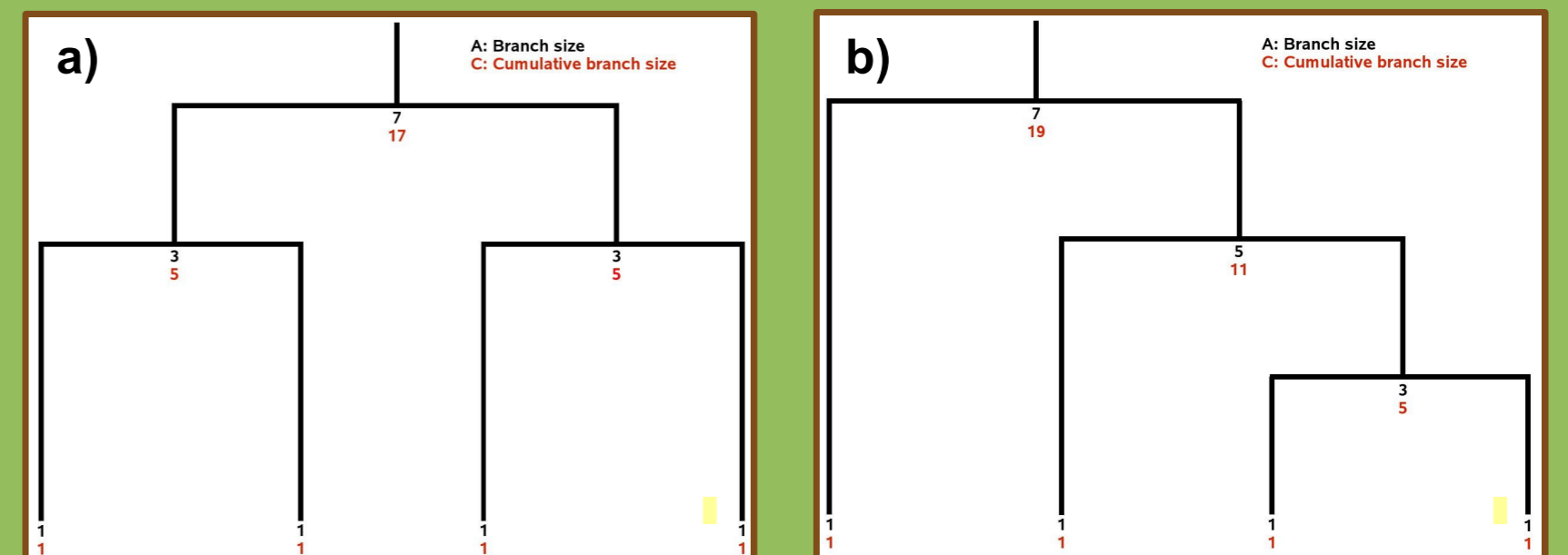
METHODS

Allometric scaling analysis

We apply the definitions used in Banavar's theorem [1] (Fig.1):

- For each node i of a tree, S_i is the subtree rooted at the node i containing all nodes below i .
- The number of nodes in the subtree is called *branch size*, A , which is defined, as $A_i = |S_i|$.
- A parameter to describe the shape of a subtree is C , the *cumulative branch size*.

Fig. 1. A_i and C_i values for a completely imbalanced (a) and a completely balanced (b) tree.



Null model (Equal-rates Markov model, ERM model)

- ERM model:
 - The phylogenies are the product of random branching.
 - The "effective speciation rate" (the difference between extinction and speciation rate) is equal for all species.
 - The process is memoryless.
- We generated a random tree, whose size was ranged between 2 and 10^6 tips, following the same evolutionary principle of the equal-rates Markov (ERM) model.

RESULTS

Branch size (A) distribution ($1-P(A) = A^{-1}$)

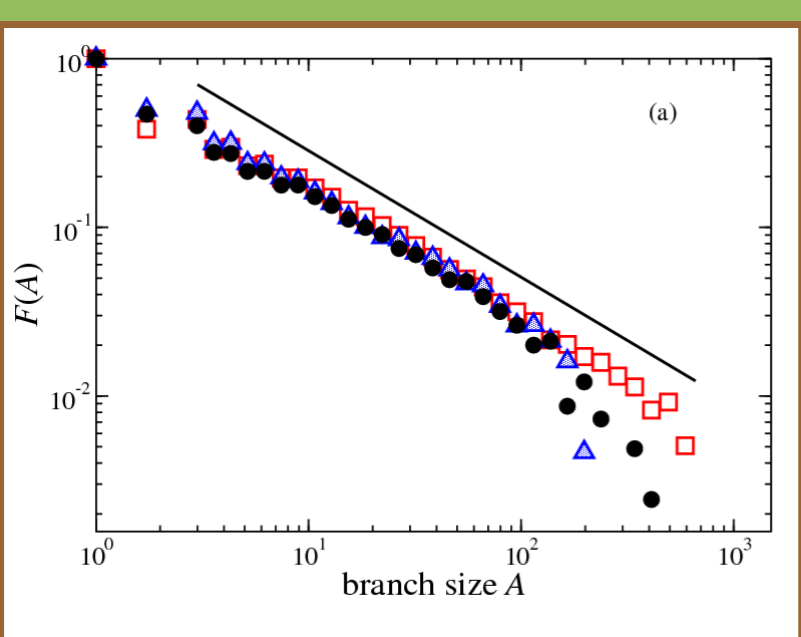


Fig. 2. Scaling of branch size (A) distribution of intra- and inter-specific phylogenies.

Cumulative branch size (C) distribution ($1-P(C) = C^{-1}$)

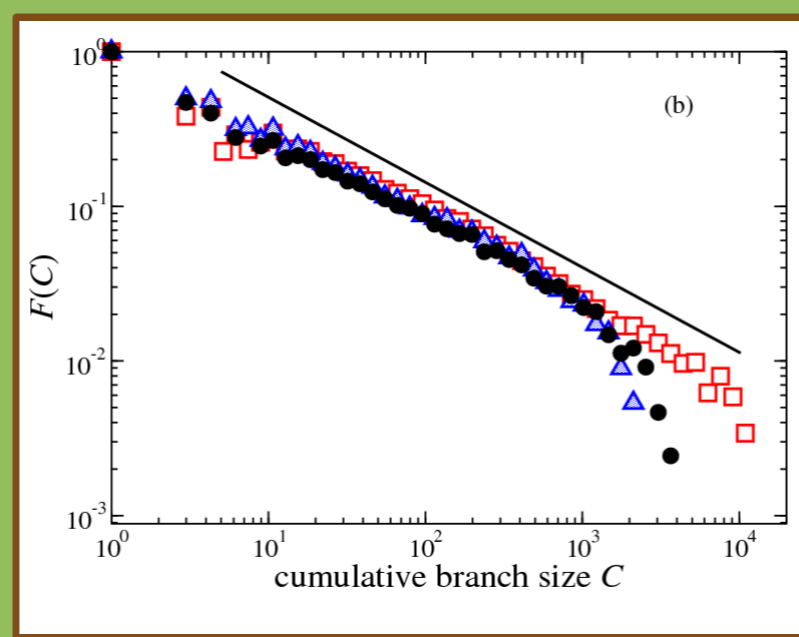


Fig. 3. Scaling of cumulative branch size (C) distribution of intra- and inter-specific phylogenies.

Allometric scaling ($C \approx A^\eta$)

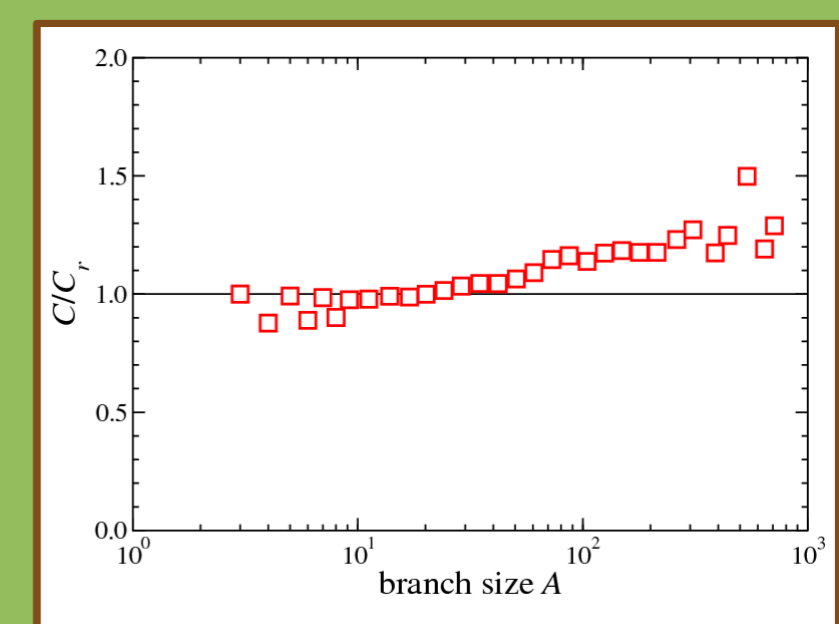
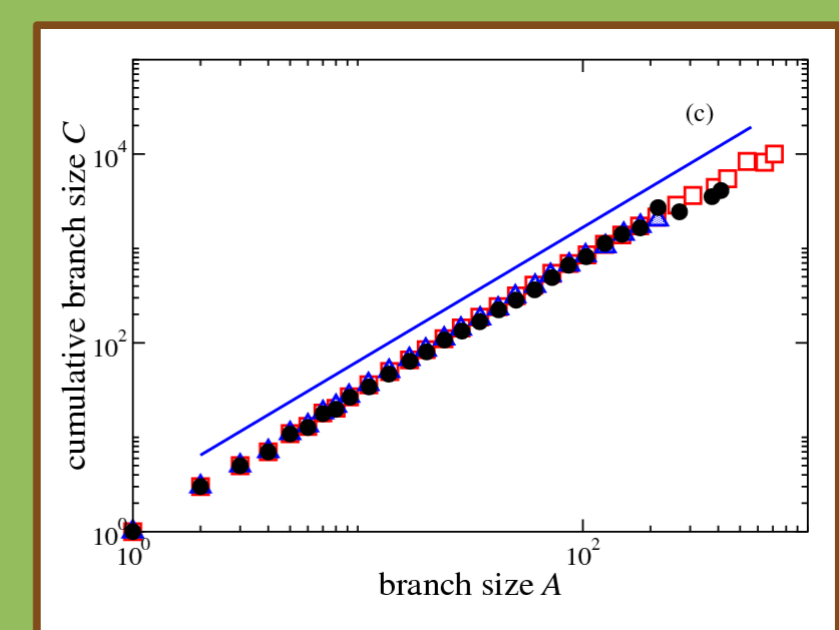


Fig. 4. Allometric scaling of intra- and inter-specific phylogenies.

CONCLUSIONS

- We have found a power-law distribution for the branch size (1.74-1.76), for the cumulative branch size (1.53-1.55), and for the allometric scaling (1.43-1.44).
- We have not found any noticeable difference between the power-law exponents for the intra- and for the inter-specific phylogenies, both in the branch size and the cumulative branch size distribution and in the allometric scaling.
- The ERM model does not explain the allometric scaling. The phylogenies have a more imbalanced topology than that predicted by the ERM model.
- The inclusion of a memory-dependent process in the null model could help us to explain the dynamics of the evolutionary events.

REFERENCES

- [1] Banavar, J. R., Maritan, A., Rinaldo, A. Size and form in efficient transportation networks, *Nature* **399**, 130 (1999).
- [2] TreeBASE. <http://www.treebase.org>.