Defining and testing Complex Multicellularity with developmental and genomic features potentially boosting the number of cell types

Irma Lozada-Chávez, Sonja J. Prohaska and Peter F. Stadler

Interdisciplinary Center for Bioinformatics and Department of Computer Science, University of Leipzig, GERMANY.

ilozada@bioinf.uni-leipzig.de

Printed in Universitätsrechenzentrum Leipzig

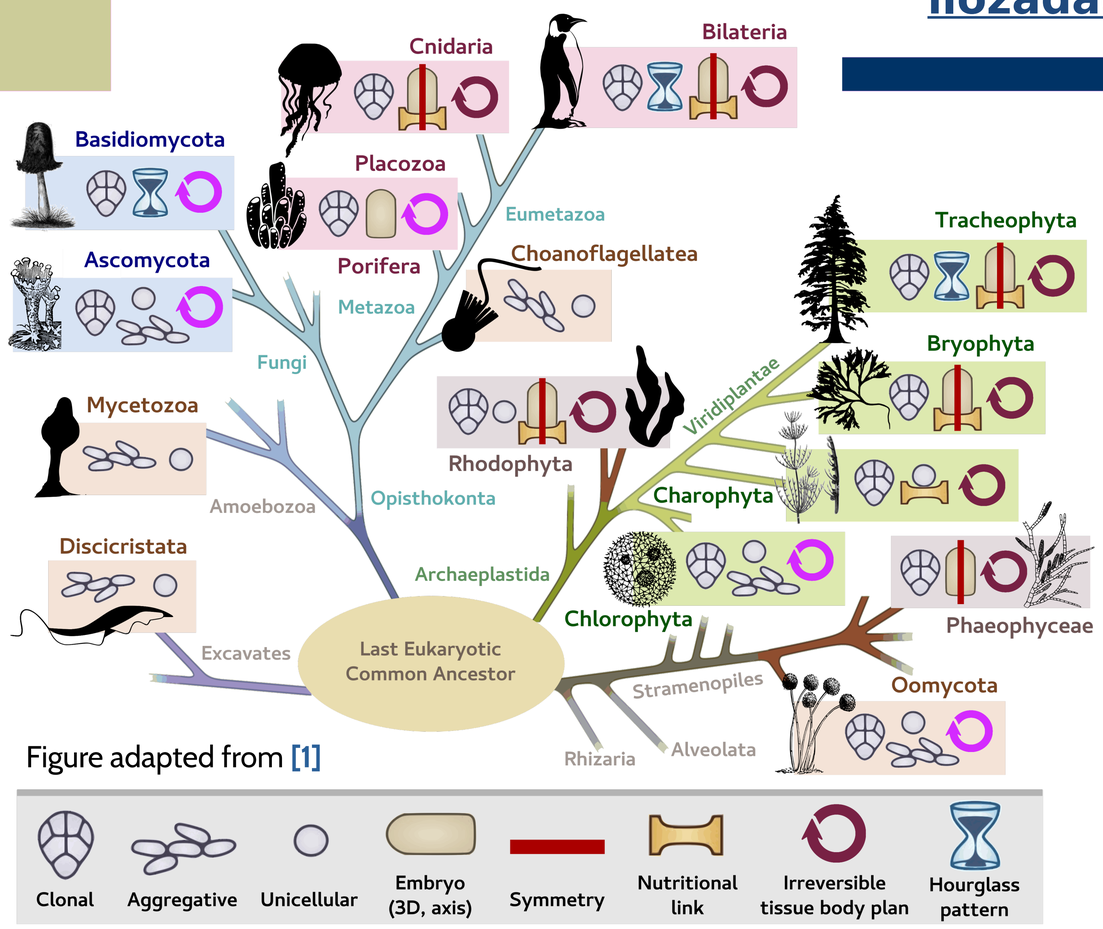
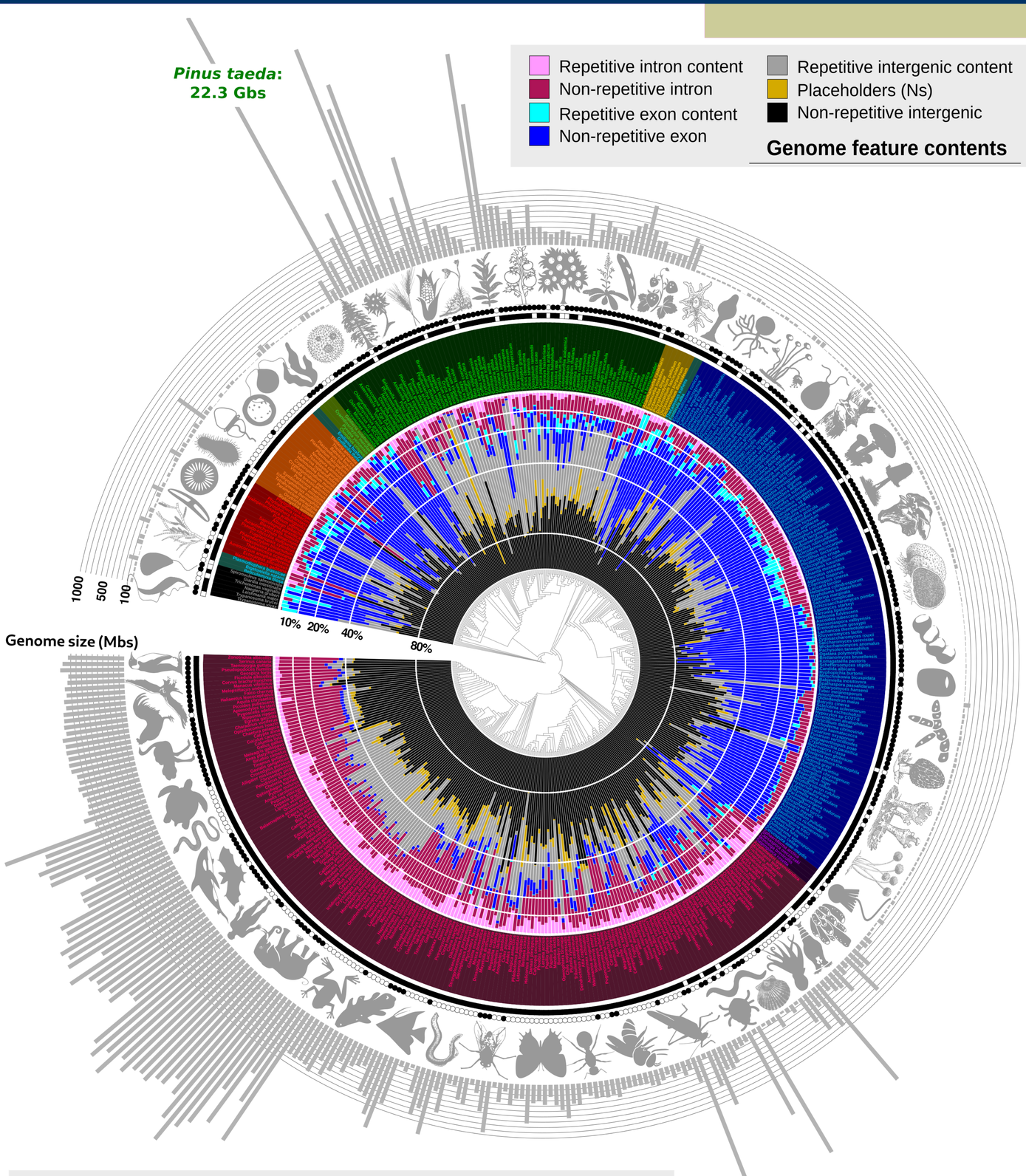
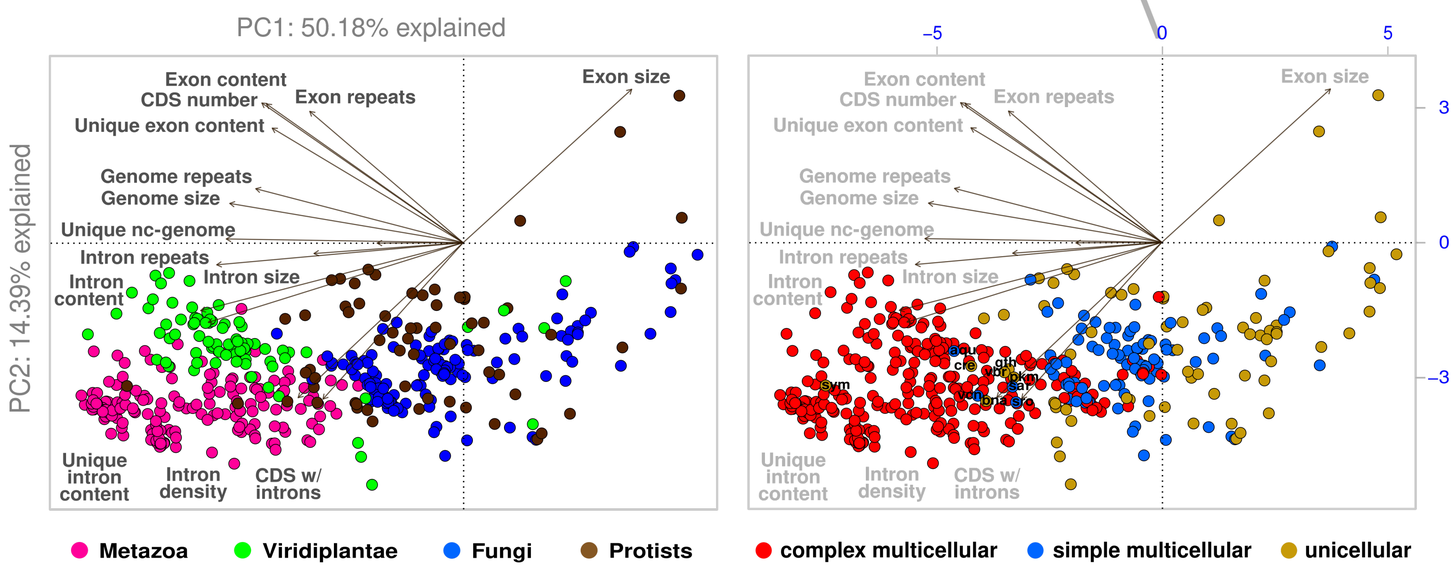


Figure adapted from [1]

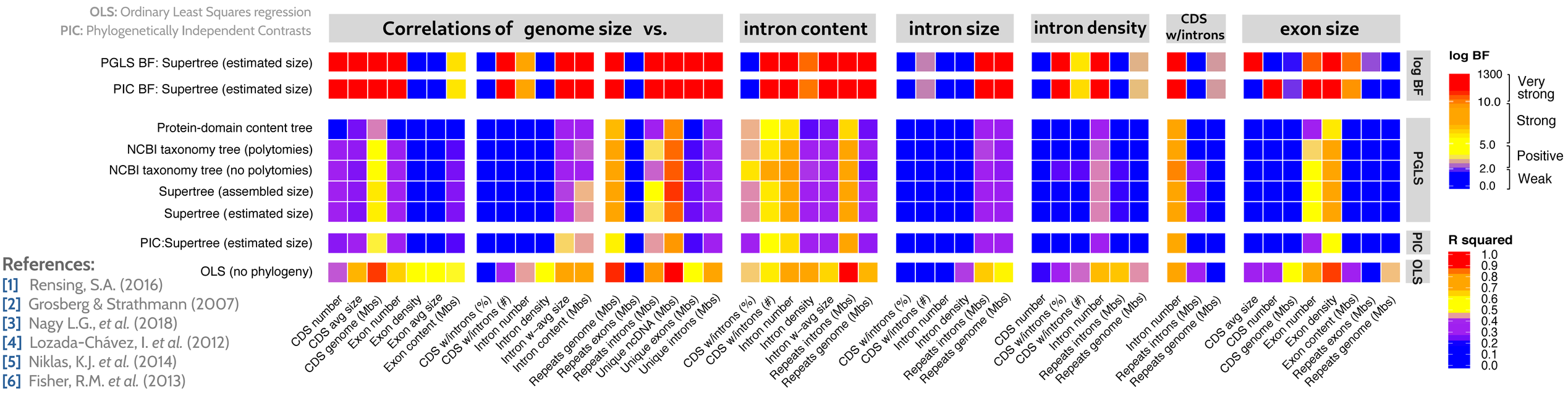


What is Complex Multicellularity (CM)? For some authors, CM is either an arbitrary border in a continuous transitional evolution of biological complexity [2] or a developmental stage coexisting with **simple multicellularity** (SM) in some species [3]. For the authors of this work, **CM is a major evolutionary transition [4], restricted to few independent origins in Eukarya**. **How can we measure and define CM to test these hypotheses?** Typically, the number of differentiated cell types (DCTs) is used as the defining feature of CM. However, accurate estimates of DCTs are only available for a small number of species, and they also fail to capture the underlying principles driving multicellular complexity [5]. We use here phylogenetically based comparative analyses for 500 eukaryotes from all major supergroups to examine whether key features defining body plan development and genome complexity can distinguish between simple and complex multicellular organisms (CMOs).

By using Principal Components Analysis (phytools), Phylogenetic Generalized Least Squares (PGLS), and Bayesian Markov-Chain-Monte-Carlo models (MCMCglmm), we carried out correlations between multicellularity (SM versus CM) and: (1) genome-wide features calculated systematically (*e.g.*, size, density and content of introns, exons, repeats); (2) traits of body plan development compiled from literature (*e.g.*, obligate/facultative multicellularity, clonal/non-clonal origin, axes formation, symmetry breaking, reversible/irreversible tissue-based body plans, number of DCTs by taxa). Our results are robust to different assumptions of causality (threshBayes) and alternative tree topologies (based on supertree construction, NCBI-taxonomy, protein-domain content). We found that CMOs and their closest ancestral relatives are characterized by high intron-richness, regardless their genome size. Indeed, changes in the variation of some intron features (such as size and repeat composition) are only weakly, while other features measuring intron abundance (within and across genes) are not, scaling with changes in genome size at the broadest phylogenetic scale. Also, changes in the length and abundance of introns within a genome are found to be largely evolving independently throughout Eukarya. Collectively, our results suggest that relatedness (in agreement with [6]), intron-richness and contingent irreversibility have influenced the likelihood of some groups to boost the concerted development of multiple cell types. They also endorse a clear distinction between SM and CM in Archaeplastida and Eumetazoa, but not in Fungi (partially due to the lack of developmental data). This calls into question the existence of either CM (as define it here) in fungi or a unifying framework for major transitions in multicellularity.



Lozada-Chávez I, *et al.* 2018. https://www.biorxiv.org/content/10.1101/283549v1



References:

[1] Rensing, S.A. (2016)
[2] Grosberg & Strathmann (2007)
[3] Nagy L.G., *et al.* (2018)
[4] Lozada-Chávez, I. *et al.* (2012)
[5] Niklas, K.J. *et al.* (2014)
[6] Fisher, R.M. *et al.* (2013)