Current Biology, Volume *23*

Supplemental Information

Group Formation, Relatedness, and the Evolution of Multicellularity

Roberta M. Fisher, Charlie K. Cornwallis, and Stuart A. West

Supplemental Experimental Procedures

Data Collection

We searched the literature for information on the evolution and development of multicellularity in as many different taxa as possible. We searched for papers by: (1) searching for combinations of the following key words in the literature with Papers 2 (covering Scopus, Web of Knowledge, JSTOR, PubMed and Google Scholar): multicellular, multicellularity, colony formation, evolution, development, aggregation, subsocial, semisocial, clonal; (2) searching reviews of multicellularity; (3) doing forward and back searches on papers. Bettina Schirrmeister provided information on the life cycles of cyanobacteria.

We found data for 17 out of the 25 independent evolutionary transitions to multicellularity. Several groups (e.g. plants) include both complex and simple multicellular species. When analysing the evolutionary transitions (Fig. 3) we are interested in the highest level of complexity obtained, and so we classified each group according to the more complex species in the group. Whilst a single cell (unitary) stage leads to clonality [1, 2], clonality can also occur in species with multicellular propagules, when the propagules are formed by cells remaining with their parents, as in some Cyanobacteria [3]. This matters because, from an evolutionary theory perspective, the key distinction is whether groups are clonal ($r = 1$) or not ($r < 1$), and not just whether there is a single cell (unitary) stage (although they will be highly correlated [1, 2]).

Statistical methods

We examined whether the evolution of multicellularity was influenced by relatedness by testing if there were differences between clonal and non-clonal taxa for six different social traits: (1) obligate versus facultative multicellularity (binary distribution); (2) total number of cells (Gaussian after log transformation); (4) the number of cell types after controlling for total number of cells (Poisson distribution); (5) presence of sterile cells (binary distribution); (4) % of cells that are sterile in taxa with sterile soma (binomial distribution); (6) complex versus simple multicellularity (binary distribution). Our analyses involved four steps:

1. Differences between clonal and non-clonal taxa in multicellular traits

We analyzed if clonal and non-clonal taxa differed in traits 1 to 6 using Bayesian generalised linear models (BGLM) with Markov chain Monte Carlo (MCMC) estimation in MCMCglmm, R version 2.15.1 [4, 5]. The results are presented in Table S1. In all models relatedness (clonal versus non-clonal) was entered as a fixed effect. For the analysis of cell types, the total number of cells (log and then Z-transformed) was included as a covariate. For some taxa information on the total number of cells was not available. To avoid excluding these data points from the analysis of cell types missing data was imputed as the mean value (=0 after Z transformation)[6].

The parameter estimates we report are the posterior mode and 95% credible intervals (lower CI – upper CI), which were back-transformed to the scale of the response variable and marginalized over the residual variance. In all models the global intercept was removed to gain absolute parameter estimates for each level of the fixed effects. Estimates of the differences between the levels of fixed effects were calculated from a posterior distribution created by subtracting the estimates for each level obtained during each MCMC iteration (labelled difference in Tables S1-S2). Parameter estimates were considered statistically significant when 95% credible intervals did not include 0 and pMCMC values calculated in MCMCglmm (number of simulated cases that are > 0 or < 0 corrected for finite number of MCMC samples) were less than 0.05.

1.2 Model specifications: We ran each analysis for 6000000 iterations with a burn-in of 1000000 and thinning interval of 1000. These iteration settings were chosen as they minimized autocorrelation between posterior samples during test runs where we varied the number of iterations from 1 to 6 million, the burn-in from 0.5 to 5 million and the thinning interval from 100 to 1000. For binary and binomial traits we used models with logit link functions and specified a fixed effect prior of N (0, σ2units, +π2/3), which is relatively uniform on the logit scale. For binary response variables we fixed the residual variance to 1. For Poisson response variables we used log link functions and for Poisson and Gaussian traits we specified a prior of V=1, nu=0.002 for the residual variance.

1.3 Model checking: We checked the convergence of models by visually inspecting trace plots of MCMC chains and using Gelman-Rubin tests in the R package 'coda' [7, 8]. For the Gelman-Rubin test we ran each analysis 3 times and used the Gelman-Rubin statistic (potential scale reduction factor; PSR) to compare within- and between-chain variance. When models have converged the PSR <1.1 and in all our analyses PSR was <1.05. Furthermore, for models using logit link functions we checked that the absolute value of the latent variable did not exceed 20 [5].

2. Phylogenetic analyses of differences between clonal and non-clonal taxa in multicellular traits

We obtained the data on multicellular traits from a very diverse range of taxa. We therefore repeated the analyses outlined in section 1, but included information on the non-independence of data caused by the phylogenetic relationships between taxa [5]. We used Bayesian phylogenetic mixed models (BPMM) with MCMC estimation implemented in MCMCglmm. A phylogenetic tree of the taxa in our dataset (created from published sources – see 2.1) was entered as a random effect, which specifies a covariance matrix describing the relationships between taxa [5]. The output from these models is reported in exactly the same way as section 1 apart from parameter estimates were marginalized over the sum of the random effects (phylogenetic and residual variance) as opposed to only residual variance. The results are presented in Table S2.

2.1 Tree creation: We used phylogenies from the following sources to resolve polytomies in our dataset and create taxonomic structuring for the analyses: (1) between the three domains showing bacteria as the root, [9]; (2) between the major groups of bacteria [10]; (3) between major groups of eukaryotes [11, 12]; (4) relationships within the fungi [13]; (5) relationships within the red algae [14]; (6) relationships within the brown algae [15]; (7) relationships within the Volvocine algae [16]. We used these phylogenies because they were recent, included taxa in our dataset and were highly cited.

We converted the information on taxonomy and topology into a phylogenetic tree using the R package 'ape' [17] In some parts of the tree there were polytomies and we resolved these using the multi2d function in ape. In the areas of the tree with polytomies there was no variation in relatedness (all taxa involved in polytomies were either clonal or non-clonal) and there was no variation in obligate versus facultative reproduction, the presence of sterile soma or complex versus simple classifications of multicellularity. Therefore resolving these polytomies randomly had little influence on our results. We set all branch lengths equal to 1 (see section 4 for tests of robustness of our results to different branch length settings).

2.2 Model specifications: Model settings were the same as in section 1.2 apart from the prior specification for the extra random effect of phylogeny. We ran models using two different priors for random effects. First, we used an inverse gamma prior that is commonly used for random effects (V = 1, nu=0.002). Second, we ran models with parameter expanded priors (half-Cauchy priors following: V=1, nu=1, alpha.mu = 0, alpha.V = 25^2) as sometimes variance components were close to 0 [8] .The inverse gamma prior led to better convergence as measured by autocorrelation between posterior samples and therefore we only present results from these models.

2.3 Model checking: We checked models in exactly the same way as in section 1.3. We also examined the convergence of estimates of random effects by visually inspecting trace plots of the MCMC chains and by examining autocorrelation between posterior samples.

3. Robustness of results to assumptions of causality and allowing relatedness to evolve across the phylogenetic tree

The analyses presented in sections 1 & 2 assume that relatedness does not evolve and that it causally affects each of the social traits (response variables). Theoretically, the causal effect of relatedness on the evolution of social traits is strongly supported, but variation in relatedness amongst cells may change over evolutionary time. We therefore modeled the phylogenetic correlation between relatedness and each of the six social traits allowing relatedness and the social traits to evolve across the tree, whilst also relaxing the assumption of causality. The results are presented in Table S3.

We analyzed phylogenetic correlations using the threshold model described by Felsenstein [18]. The threshold model allows the correlated evolution of discrete characters, such as clonal versus non-clonal and obligate versus facultative reproduction, to be modeled by assuming there is some unobserved quantitative character (liability) that underlies discrete characters: at a certain threshold in liability the phenotype switches from one state to another. This is well suited to modeling evolutionary changes in discrete characters because it allows the underlying probability of displaying a character to evolve rather than assuming changes between states at each node occur with equal probability across tree, as with continuous-time Markov models. For example, the probability of evolving obligate multicellular reproduction is more likely if taxa are within clades dominated obligate reproduction in comparison to the average across the tree [5, 18]. Furthermore, it also allows the phylogenetic correlation between discrete and continuous traits, such as clonal versus non-clonal and total number of cells, to be estimated. We used a Bayesian implementation of the threshold model, threshBayes, in the R package 'phytools' [19] to analyze our data. Prior to analysis the number of cells and the number of cell types were log and then Z-transformed and the proportion of sterile cells was square root arcsine transformed.

3.1 Model specifications: We ran each analysis for 6000000 iterations with a burn-in of 1000000 and thinning interval of 1000. We used the default prior settings that specify exponential priors for the

evolutionary variances associated with each trait, normal priors for the ancestral state of each trait and a uniform prior for the correlation between the two traits.

3.2 Model checking: We checked models with the same diagnostics as section 2.3.

4. Robustness of results to assumptions of tree branch lengths

Phylogenetic analyses using trees with equal branch lengths can over estimate the strength of association between variables [20, 21].We therefore examined the robustness of our results to assumptions of branch lengths in two ways.

4.1 Repeating analyses using different tree transformations: We repeated the analyses outlined in sections 2 & 3 twice, once where the branches of the phylogenetic tree had been transformed to simulate an early burst of diversification (Fig. S2a. Referred to as 'early burst' in tables) and once to simulate a late burst of diversification (Fig. S2c. Referred to as 'late burst' in tables). We created the early and late burst trees using Pagel's δ transformations (Early burst $\delta = 0.2$, Late burst $\delta = 1.5$) implemented through the deltaTree function in the R package 'Geiger' [22-24]. The results of the MCMCglmm analyses are presented in Table S4 and those of threshBayes are given in Table S5.

4.2 Removing branch lengths: ancestral reconstruction using maximum parsimony: We used maximum parsimony, which does not require a tree with branch lengths, to reconstruct ancestral states of relatedness using the MPR function in the R package 'ape'. After reconstructing ancestral states we identified independent evolutionary transitions between clonal and non-clonal multicelluarity. We tested whether clonal and non-clonal taxa differed in their social traits using Wilcoxon paired rank sum tests of the mean phenotype of the clonal taxa and the mean phenotype of the non-clonal taxa involved in each independent comparison. The results of these analyses are presented in Table S6.

Supplemental References

1. Grosberg, R., and Strathmann, R. (1998). One cell, two cell, red cell, blue cell: the persistence of a unicellular stage in multicellular life histories. Trends in Ecology & Evolution *13*, 112–116.

2. Grosberg, R. K., and Strathmann, R. R. (2007). The Evolution of Multicellularity: A Minor Major Transition? Annu Rev Ecol Evol S *38*, 621–654.

3. Schirrmeister, B. E., Antonelli, A., and Bagheri, H. C. (2011). The origin of multicellularity in cyanobacteria. BMC Evol Biol *11*, 45.

4. Hadfield, J. D., and Nakagawa, S. (2010). General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. Journal of Evolutionary Biology *23*, 494–508.

5. Hadfield, J. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Software *33*, 1–22.

6. Nakagawa, S., and Freckleton, R. P. (2010). Model averaging, missing data and multiple imputation: a case study for behavioural ecology. Behav Ecol Sociobiol *65*, 103–116.

7. Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. R news *6*, 1–25.

8. Gelman, A., and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. Statistical science *7*, 457–472.

9. Williams, T. A., Foster, P. G., Nye, T. M. W., Cox, C. J., and Embley, T. M. (2012). A congruent phylogenomic signal places eukaryotes within the Archaea. Proceedings of the Royal Society B: Biological Sciences *279*, 4870–4879.

10. Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., Kunin, V., Goodwin, L., Wu, M., Tindall, B. J., et al. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature *462*, 1056–1060.

11. Burki, F., and Pawlowski, J. (2006). Monophyly of Rhizaria and Multigene Phylogeny of Unicellular Bikonts. Mol Biol Evol *23*, 1922–1930.

12. Paps, J., Medina-Chacón, L. A., Marshall, W., Suga, H., and Ruiz-Trillo, I. (2012). Molecular Phylogeny of Unikonts: New Insights into the Position of Apusomonads and Ancyromonads and the Internal Relationships of Opisthokonts. Protist *164*, 2–12.

13. James, T. Y., Kauff, F., Schoch, C. L., Matheny, P. B., Hofstetter, V., Cox, C. J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., et al. (2006). Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature *443*, 818–822.

14. Freshwater, D. W., and Rueness, J. (1994). Phylogenetic relationships of some European Gelidium (Gelidiales, Rhodophyta) species, based on rbcLnucleotide sequence analysis. Phycologia *33*, 187–194.

15. Silberfeld, T., Leigh, J. W., Verbruggen, H., Cruaud, C., de Reviers, B., and Rousseau, F. (2010). A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the "brown algal crown radiation". Mol. Phylogenet. Evol.

56, 659–674.

16. Hallmann, A. (2011). Evolution of reproductive development in the volvocine algae. Sex. Plant Reprod. *24*, 97–112.

17. Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics *20*, 289–290.

18. Felsenstein, J. (2012). A comparative method for both discrete and continuous characters using the threshold model. The American Naturalist *179*, 145–156.

19. Revell, L. J. (2011). phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution *3*, 217–223.

20. Kaiser, D. (2001). Building a multicellular organism. Annual Review of Genetics *35*, 103–123.

21. Purvis, A., Gittleman, J. L., and Luh, H.-K. (1994). Truth or Consequences: Effects of Phylogenetic Accuracy on Two Comparative Methods. J. Theor. Biol. *167*, 293–300.

22. Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. (2008). GEIGER: investigating evolutionary radiations. Bioinformatics *24*, 129–131.

23. Pagel, M. (1994). Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters. Proceedings: Biological Sciences *255*, 37–45.

24. Pagel, M. (1999). Inferring the historical patterns of biological evolution. Nature *401*, 877–884.

Figure S1: Ancestral reconstruction of the relationships between clonal (black tip labels and edges) and non-clonal (red tip labels and edges) taxa. This figure relates to Figure 2 in the main text.

Figure S2: Early burst, late burst and equal length trees to test the robustness of our assumptions. This figure is related to Figure 2 in the main text.

Figure S3: The graphs show a comparison between clonal and non-clonal multicellular species for different measures of sociality. Graphs (a)-(d) show the raw data, which assumes species as independent data points. This figure is an alternative to Figure 3 in the main text.

Supplementary Tables

Table S1: Analyses of the effect of relatedness (clonal versus non-clonal) on social traits in multicellular groups not accounting for phylogenetic relationships between species conducted using MCMCglmm.

Table S2: Analyses of the effect of relatedness (clonal versus non-clonal) on social traits in multicellular groups taking into account phylogenetic relationships between species conducted using MCMCglmm.

Table S3: Estimates of the phylogenetic correlation between relatedness (clonal versus non-clonal) and different social traits in multicellular group analysed using threshBayes.

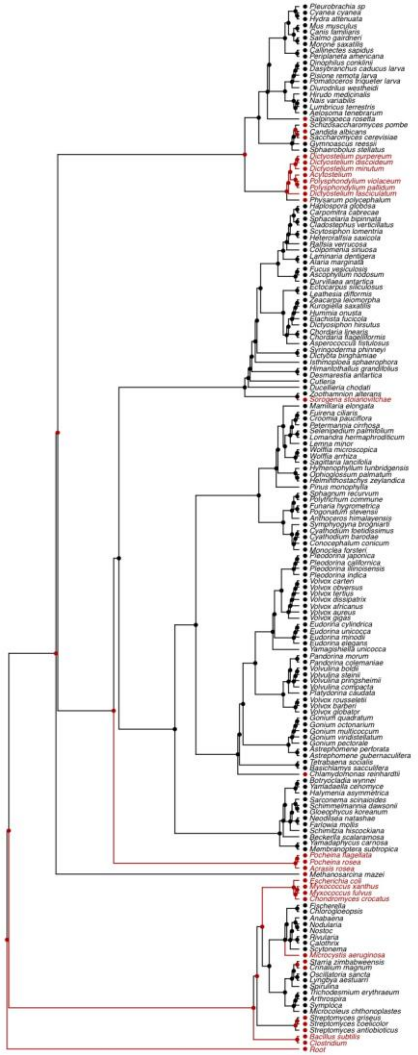
Table S4: The influence of assumptions about branch lengths on the estimated effect of relatedness (clonal versus non-clonal) on social traits in multicellular groups analysed using MCMCglmm.

Table S5: The influence of assumptions about branch lengths on the estimated phylogenetic correlation between relatedness (clonal versus non-clonal) and social traits in multicellular groups analysed using threshBayes.

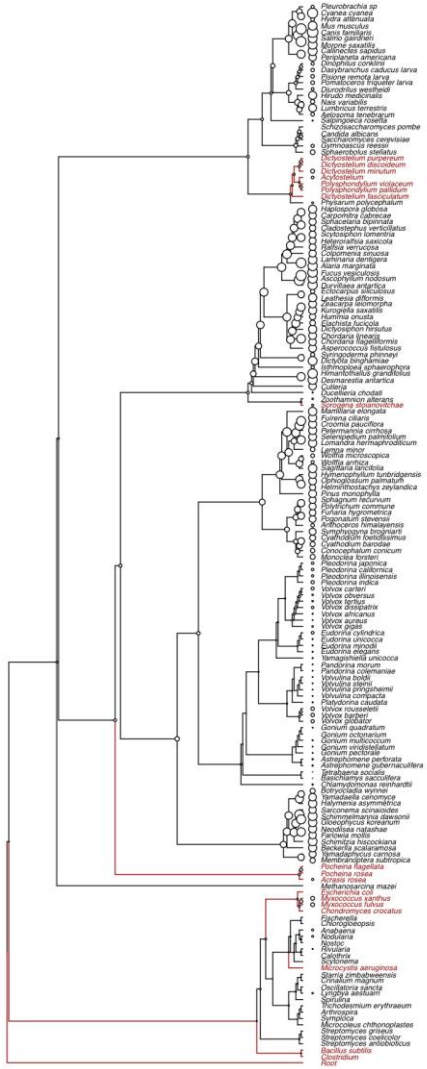
Table S6: Wilcoxon paired rank sum tests of differences in social traits across independent evolutionary transitions between clonal and non-clonal states identified using maximum parsimony ancestral reconstruction.

Table S7: Raw data collected from literature review, which was used for all analyses, included as an Excel file. References for all raw data are included.

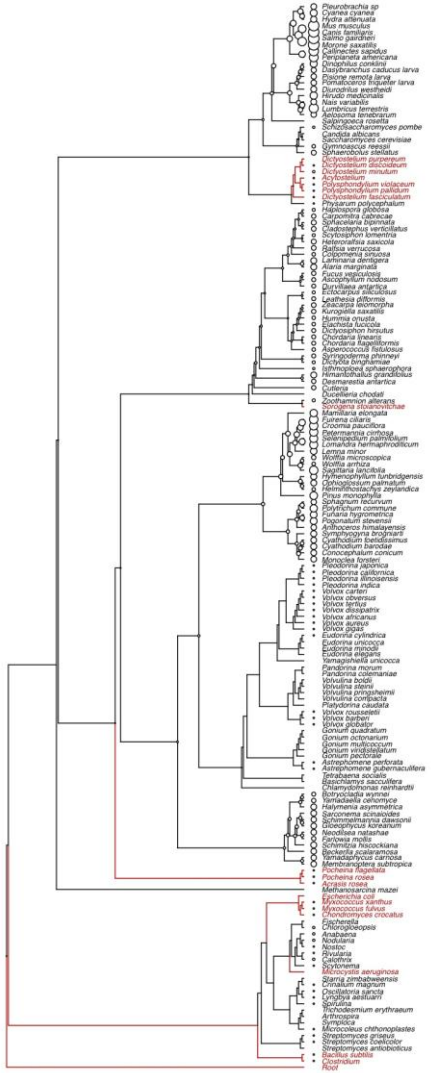
(a) Obligate versus Facultative reproduction
(pMCMC = 0.0002)



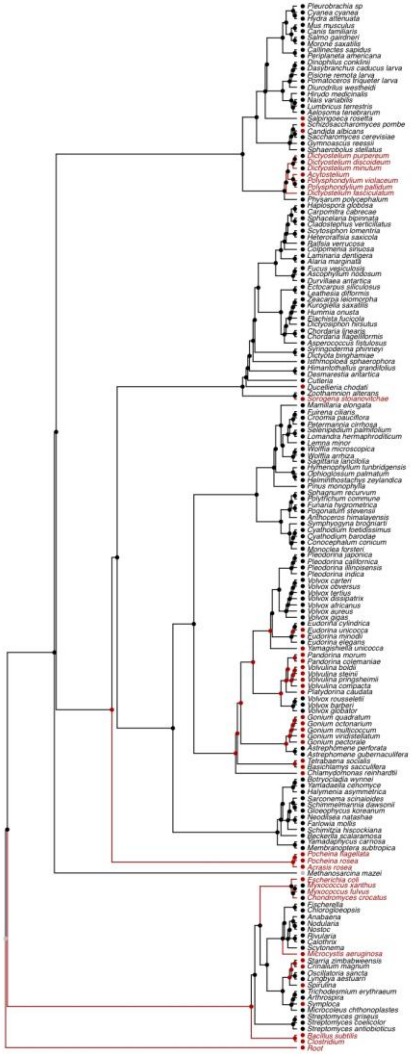
(b) Number of cells
(pMCMC = 0.27)



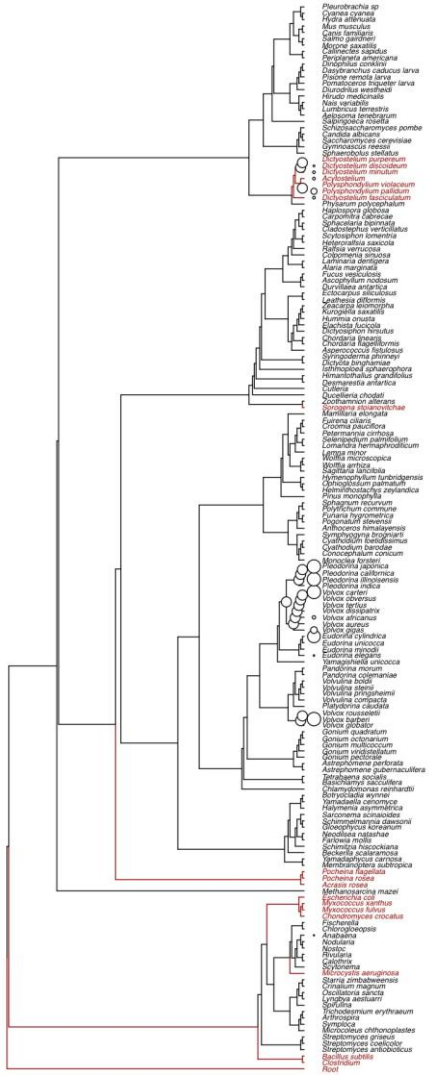
(c) Number of cell types
(pMCMC = 0.0008)



(d) Probability of sterile soma
(pMCMC = 0.02)



(e) % cells that are sterile
(pMCMC = 0.26)



(f) Complex versus simple multicellularity
(pMCMC = 0.21)

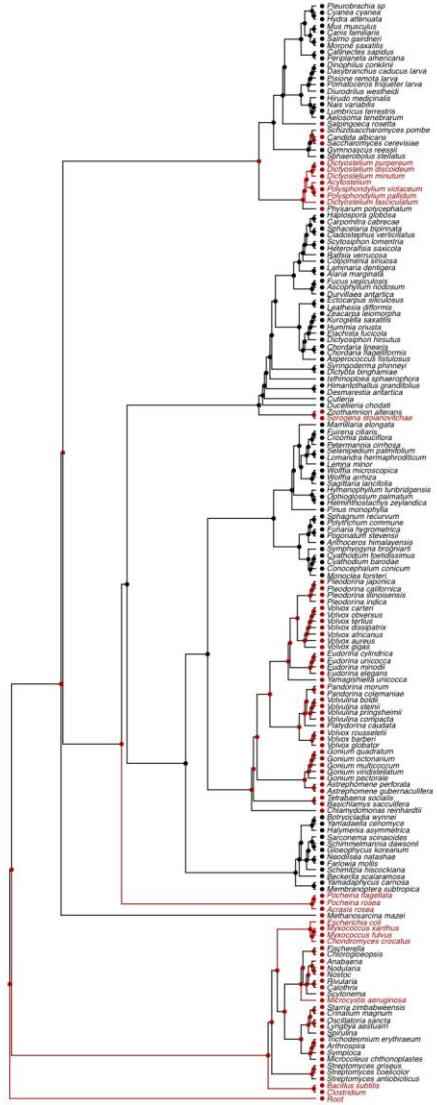


Figure S1. Ancestral Reconstruction of the Relationships between Clonal and Nonclonal Taxa, Related to Figure 2

Clonal, black tip labels and edges; nonclonal, red tip labels and edges. **(a)** obligate versus facultative multicellularity (obligate = black dots at tips and nodes, facultative = red dots at tips and nodes), **(b)** total number of cells (size of dots at tips and nodes = log number of cells), **(c)** number of cell types (size of dots at tips and nodes = log number of cell types), **(d)** probability of sterile soma (sterile = black dots at tips and nodes, non-sterile = red dots at tips and nodes), **(e)** % cells that are sterile (size of dots at tips and nodes = %) and **(f)** complex versus simple multicellularity (complex = black dots at tips and nodes, simple = red dots at tips and nodes). Where data were missing for the tips (e.g for the total number of cells and cell types) no symbols were plotted. Grey dots and lines represent ambiguous reconstructions where nodes could not be assigned to one state or the other (posterior probability between 0.1 and 0.9: state 0 was classified as <0.1 posterior probability and state 1 as >0.9 posterior probability). pMCMC values are taken from Table S2 and represent differences between clonal and nonclonal taxa in each of the social traits.

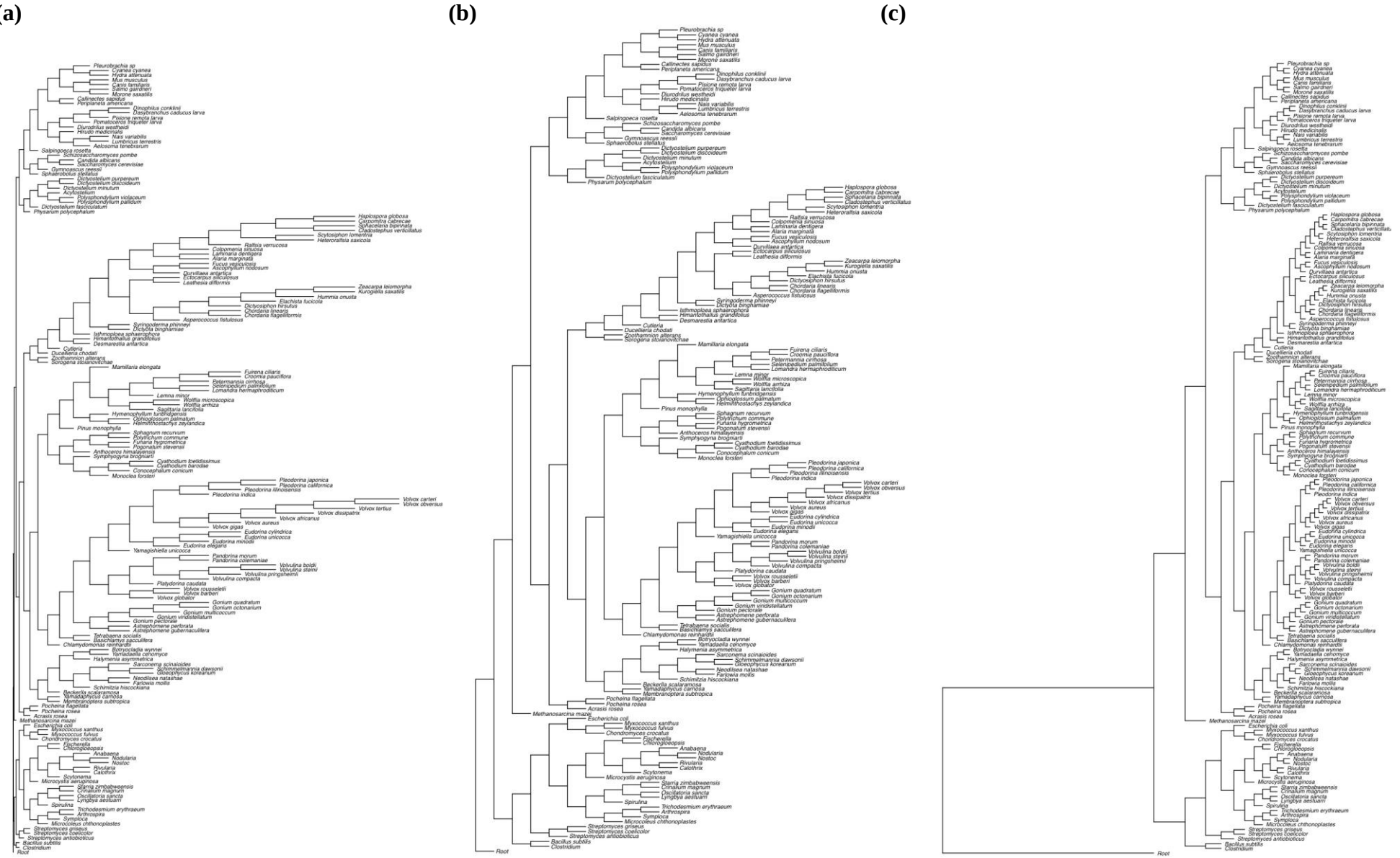
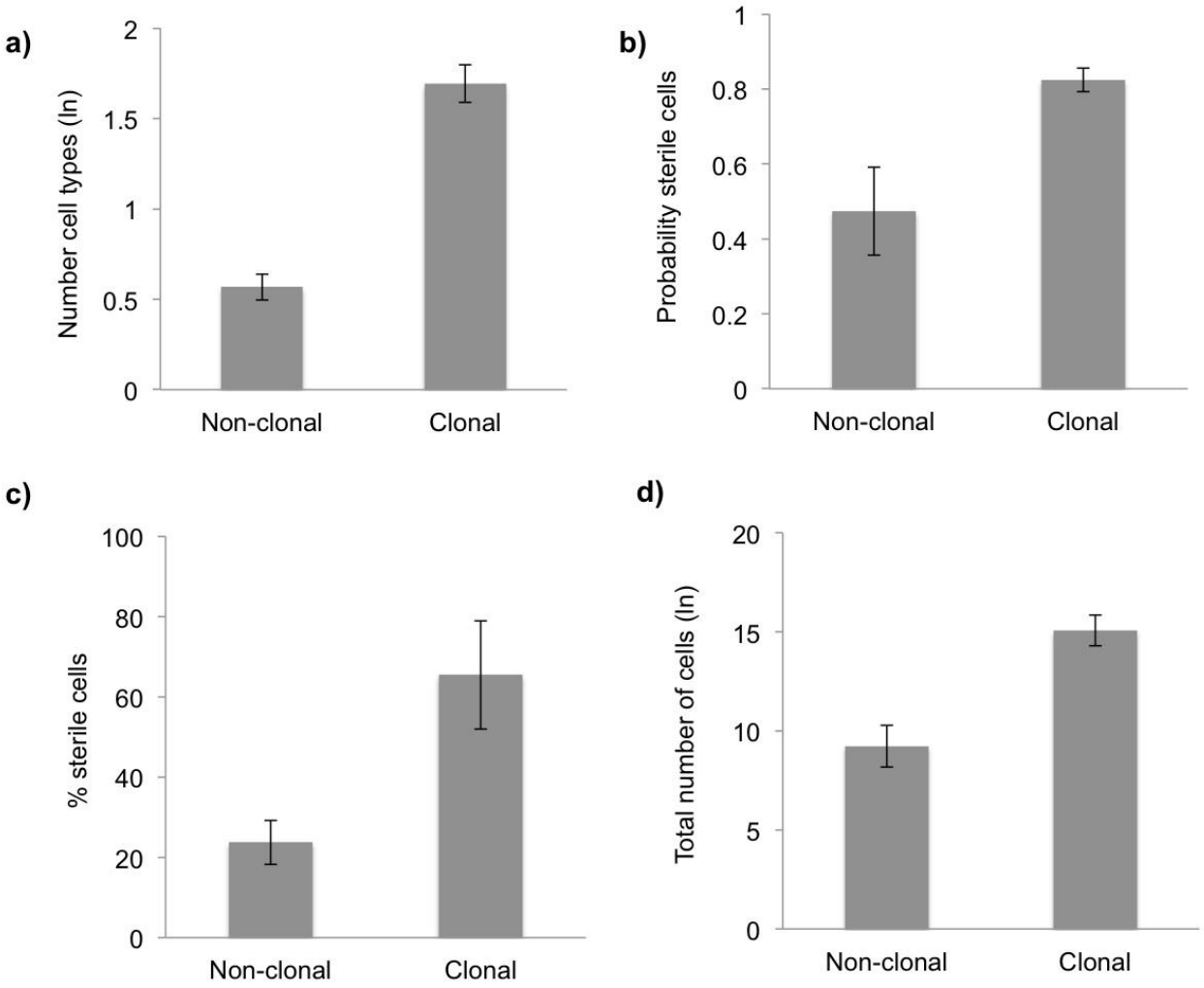


Figure S2. The Early-Burst, Equal-Length, and Late-Burst Trees Used for Analyses Examining the Robustness of Results to Assumptions about Branch Lengths, Related to Figure 2

(a) Early burst (δ=0.2), (b) equal length and (c) late (δ=1.5) burst trees.


a)
Number cell types (ln)
2
1.5
1
0.5
0
Non-clonal
Clonal
b)
Probability sterile cells
1
0.8
0.6
0.4
0.2
0
Non-clonal
Clonal
c)
% sterile cells
100
80
60
40
20
0
Non-clonal
Clonal
d)
Total number of cells (ln)
20
15
10
5
0
Non-clonal
Clonal

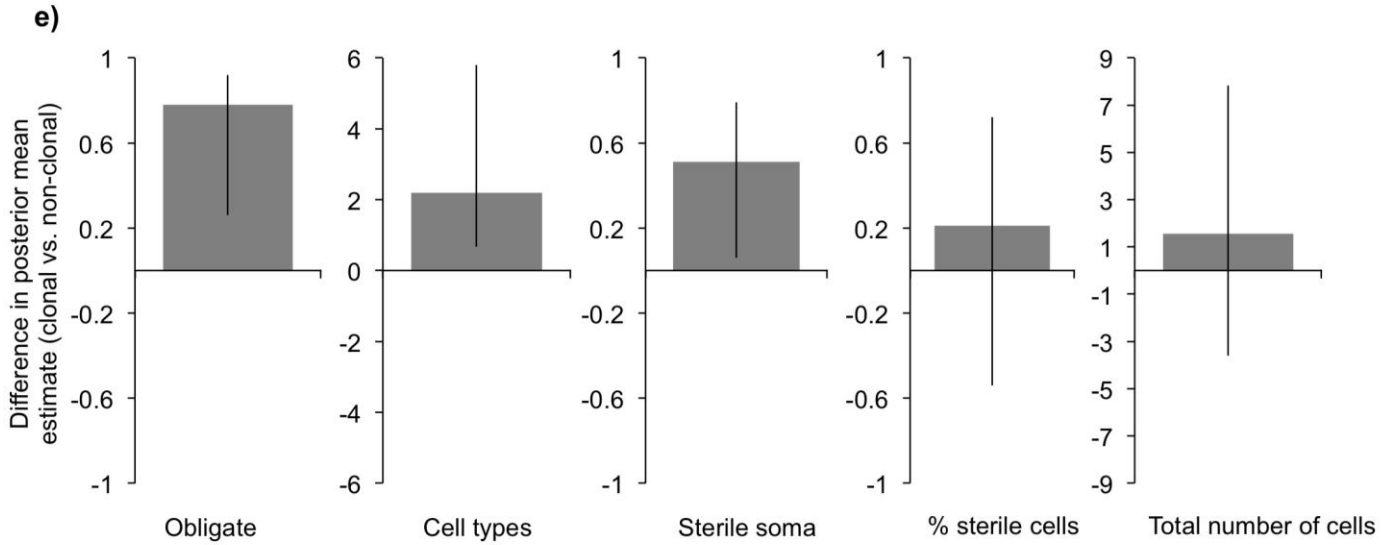


Figure S3. Relatedness and Sociality in Multicellular Groups

The graphs show, for multicellular groups that are either clonal or potentially non-clonal, the: (a) number of cell types; (b) probability of having sterile cells; (c) percentage of sterile cells in species with a sterile cells; (d) total number of cells in the group. The graphs show the raw data, with means and standard errors calculated assuming species as independent data points. (e) shows the difference between clonal and non-clonal (posterior mean estimate from BPMM ± 95% credible interval) species for five measures of sociality: the probability of being obligate, number of cell types, probability of having a sterile soma, % sterile cells in species with a sterile soma and total number of cells (related to Figure 3 in the main text).

Table S1. Analyses of the Effect of Relatedness on Social Traits in Multicellular Groups, Not Accounting for Phylogenetic Relationships between Species, Conducted Using MCMCglmm

Trait	Species	Effect	Posterior mode	Lower CI	Upper CI	P diff
1. Obligate versus facultative	168	Clonal	0.91	0.87	0.95	
		Non-clonal	0.03	0.0006	0.13	
		Difference	0.89	0.77	0.94	**<0.0001**
2. Number of cells	133	Clonal	15.16	13.67	16.73	
		Non-clonal	8.36	1.33	14.39	
		Difference	6.73	0.55	13.99	**0.02**
3. Number of cell types	162	Clonal	7.23	6.36	8.83	
		Non-clonal	2.64	1.42	4.12	
		Difference	4.91	2.95	6.65	**<0.0001**
		Number of cells	3.50	2.88	4.24	**<0.0001**
4. Probability of sterile soma	167	Clonal	0.83	0.76	0.88	
		Non-clonal	0.45	0.26	0.69	
		Difference	0.30	0.14	0.58	**0.001**
5. % cells that are sterile	48	Clonal	0.74	0.49	0.86	
		Non-clonal	0.35	0.17	0.65	
		Difference	0.35	-0.02	0.61	**0.05**
6. Complex versus simple	168	Clonal	0.58	0.50	0.66	
		Non-clonal	0.03	0.00	0.14	
		Difference	0.55	0.41	0.65	**<0.0001**

Table S2. Analyses of the Effect of Relatedness on Social Traits in Multicellular Groups, Taking into Account Phylogenetic Relationships between Species, Conducted Using MCMCglmm

Trait	Effect	Posterior mode	Lower CI	Upper CI	P diff
1. Obligate versus facultative	Clonal	0.80	0.43	0.97	
	Non-clonal	0.02	0.002	0.35	
	Difference	0.78	0.26	0.92	**0.0002**
	Phylogeny[1]	0.36	0.08	0.87	
2. Number of cells	Clonal	4.98	-2.94	12.04	
	Non-clonal	2.55	-3.05	8.96	
	Difference	1.55	-3.60	7.84	0.27
	Phylogeny[2]	0.99	0.34	0.99	
3. Number of cell types	Clonal	4.33	2.12	8.53	
	Non-clonal	1.71	0.80	3.69	
	Difference	2.20	0.67	5.79	**0.0008**
	Number of cells	1.98	1.68	2.23	**<0.0001**
	Phylogeny[2]	0.97	0.65	0.99	
4. Probability of sterile soma	Clonal	0.79	0.42	0.97	
	Non-clonal	0.19	0.03	0.53	
	Difference	0.51	0.06	0.79	**0.02**
	Phylogeny[1]	0.66	0.29	0.95	

5. % cells that are sterile	Clonal	0.61	0.06	0.92	
	Non-clonal	0.34	0.03	0.86	
	Difference	0.21	-0.54	0.72	0.39
	Phylogeny[2]	0.99	0.06	0.99	
6. Complex versus simple[3]	Clonal	0.50	0.20	0.74	
	Non-clonal	0.53	0.38	0.92	
	Difference	-0.10	-0.57	0.17	0.21
	Phylogeny[1]	0.97	0.56	0.99	

[1]Intraclass correlation coefficient of phylogeny.
[2]Proportion of residual variation explained by phylogeny.
[3]Should be interpreted with caution, logit link function overloaded as latent variable exceeded 20.

Table S3. Estimates of the Phylogenetic Correlation between Relatedness and Different Social Traits in a Multicellular Group Analyzed Using threshBayes

Trait	Phylogenetic correlation with relatedness (posterior mode)	Lower CI	Upper CI	P
1. Obligate versus facultative	0.73	0.39	0.97	**<0.0001**
2. Number of cells	0.36	-0.11	0.68	0.08
3. Number of cell types	0.59	0.18	0.74	**0.003**
4. Probability of sterile soma	0.58	0.11	0.73	**0.009**
5. % cells that are sterile	0.31	-0.59	0.81	0.34
6. Complex versus simple	0.26	-0.09	0.60	0.09

P value = proportion of iterations where phylogenetic correlations was greater than 0

Table S4. The Influence of Assumptions about Branch Lengths on the Estimated Effect of Relatedness on Social Traits in Multicellular Groups Analyzed Using MCMCglmm

Trait	Branch lengths	DIC	Posterior mode of difference between clonal and non-clonal	Lower CI	Upper CI	P diff	Phylogenetic posterior mode	Lower CI	Upper CI
1. Obligate versus facultative	Equal length	46.7	0.78	0.26	0.92	**0.0002**	0.36	0.08	0.87
	Early burst	53.8	0.54	0.20	0.81	**0.001**	0.77	0.23	0.96
	Late burst	42.9	0.64	0.22	0.88	**0.001**	0.68	0.32	0.96
2. Number of cells	Equal length	196.8	1.55	-3.60	7.84	0.27	0.99	0.34	0.99
	Early burst	810.6	4.78	-0.31	8.78	0.06	0.47	0.23	0.69
	Late burst	81.5	2.25	-6.82	7.75	0.47	0.99	0.80	0.99
3. Number of cell types	Equal length	730.4	2.20	0.67	5.79	**0.0008**	0.97	0.65	0.99
	Early burst	743.8	2.99	1.48	4.97	**<0.0001**	0.99	0.84	0.99
	Late burst	728.1	1.66	0.07	16.51	**0.003**	0.99	0.89	0.99
4. Probability of sterile soma	Equal length	69.2	0.51	0.06	0.79	**0.02**	0.66	0.29	0.95
	Early burst	61.8	0.27	0.02	0.58	**0.02**	0.98	0.59	0.99
	Late burst	63.3	0.23	-0.04	0.54	0.05	0.98	0.68	0.99
5. % cells that are sterile	Equal length	1002.9	0.21	-0.54	0.72	0.39	0.99	0.06	0.99
	Early burst	1002.9	0.09	-0.35	0.63	0.33	0.99	0.52	0.99
	Late burst	1003.0	0.25	-0.40	0.58	0.32	0.99	0.001	0.99
6. Complex versus simple[1]	Equal length	15.7	-0.10	-0.57	0.17	0.21	0.97	0.56	0.99
	Early burst	23.4	-0.05	-0.35	0.09	0.15	0.99	0.76	0.99

Late burst	11.05	-0.03	-0.31	0.14	0.28	0.99	0.89	0.99

[1]Should be interpreted with caution, logit link function overloaded as latent variable exceeded 20.

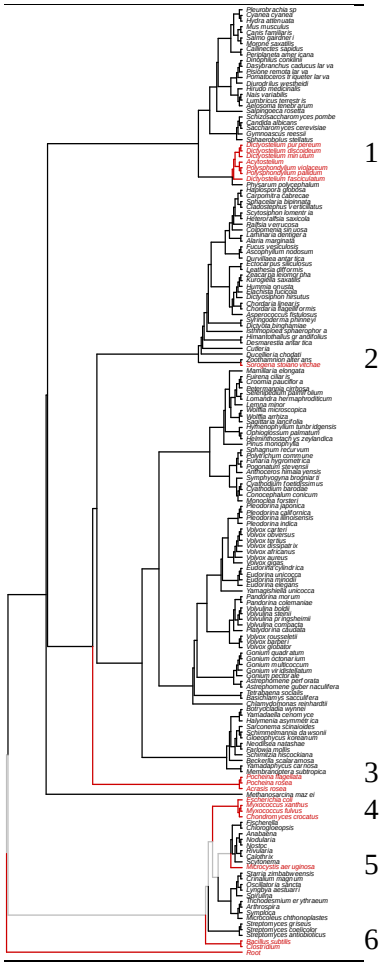
Table S5. The Influence of Assumptions about Branch Lengths on the Estimated Phylogenetic Correlation between Relatedness and Social Traits in Multicellular Groups Analyzed Using threshBayes

Trait	Branch lengths	LogL	Phylogenetic correlation with relatedness (posterior mode)	Lower CI	Upper CI	P diff
1. Obligate versus facultative	Equal length	-606.7	0.73	0.39	0.97	**<0.0001**
	Early burst	-449.9	0.80	0.49	0.94	**<0.0001**
	Late burst	-356.5	0.70	0.32	0.89	**0.002**
2. Number of cells	Equal length	-674.4	0.36	-0.11	0.68	0.08
	Early burst	-645.5	0.51	0.07	0.70	**0.01**
	Late burst	-558.8	0.23	-0.25	0.69	0.18
3. Number of cell types	Equal length	-381.9	0.59	0.18	0.74	**0.003**
	Early burst	-356.6	0.53	0.25	0.75	**0.0004**
	Late burst	-252.8	0.57	0.17	0.77	**0.005**
4. Probability of sterile soma	Equal length	-617.7	0.58	0.11	0.73	**0.009**
	Early burst	-489.3	0.50	0.17	0.74	**0.005**
	Late burst	-328.1	0.44	0.09	0.77	**0.016**
5. % cells that are sterile	Equal length	-35.2	0.31	-0.59	0.81	0.34
	Early burst	-29.4	0.50	-0.46	0.82	0.28
	Late burst	-26.6	0.39	-0.62	0.87	0.36
6. Complex versus simple	Equal length	-625.2	0.26	-0.09	0.60	0.09
	Early burst	-485.2	0.40	0.10	0.69	**0.01**

	Late burst	-334.0	0.45	-0.009	0.78	0.05

Table S6. Wilcoxon Paired Rank Sum Tests of Differences in Social Traits across Independent Evolutionary Transitions between Clonal and Nonclonal States Identified Using Maximum Parsimony Ancestral Reconstruction

Clonal, black; nonclonal, red.

Maximum parsimony reconstruction of clonality	**Contrast**	**Probability of being obligate**		**Number of cells (log transformed)**		**Cell types**		**Probability of having sterile soma**		**% sterile cells**		**Probability of being complex**	
		Clonal	Non	Clonal	Non	Clonal	Non	Clonal	Non	Clonal	Non	Clonal	Non
	1	0	0	5.7	10.6	2	2	1	0.8	-	19.7	0	0
	2	1	0	4.95	6.2	4	1	1	0	-	0	0	0
	3	0.99	0	23.3	6.6	9.7	2	0.8	0	22.27	0	0.5	0
	4	0.78	0	4.6	11.5	1.6	1.8	0.6	0.8	0	0	0	0
	5	1	0	5.8	-	2.6	1	1	0	10	0	0	0
	6	0	0	-	-	2	2	1	0	-	0	0	0
	Average	0.63	0	8.9	8.8	3.6	1.6	0.88	0.26	10.76	3.28	0.08	0
	P value	0.07		0.88		0.14		0.07		0.18		0.32	