

**Current Biology, Volume 27**

**Supplemental Information**

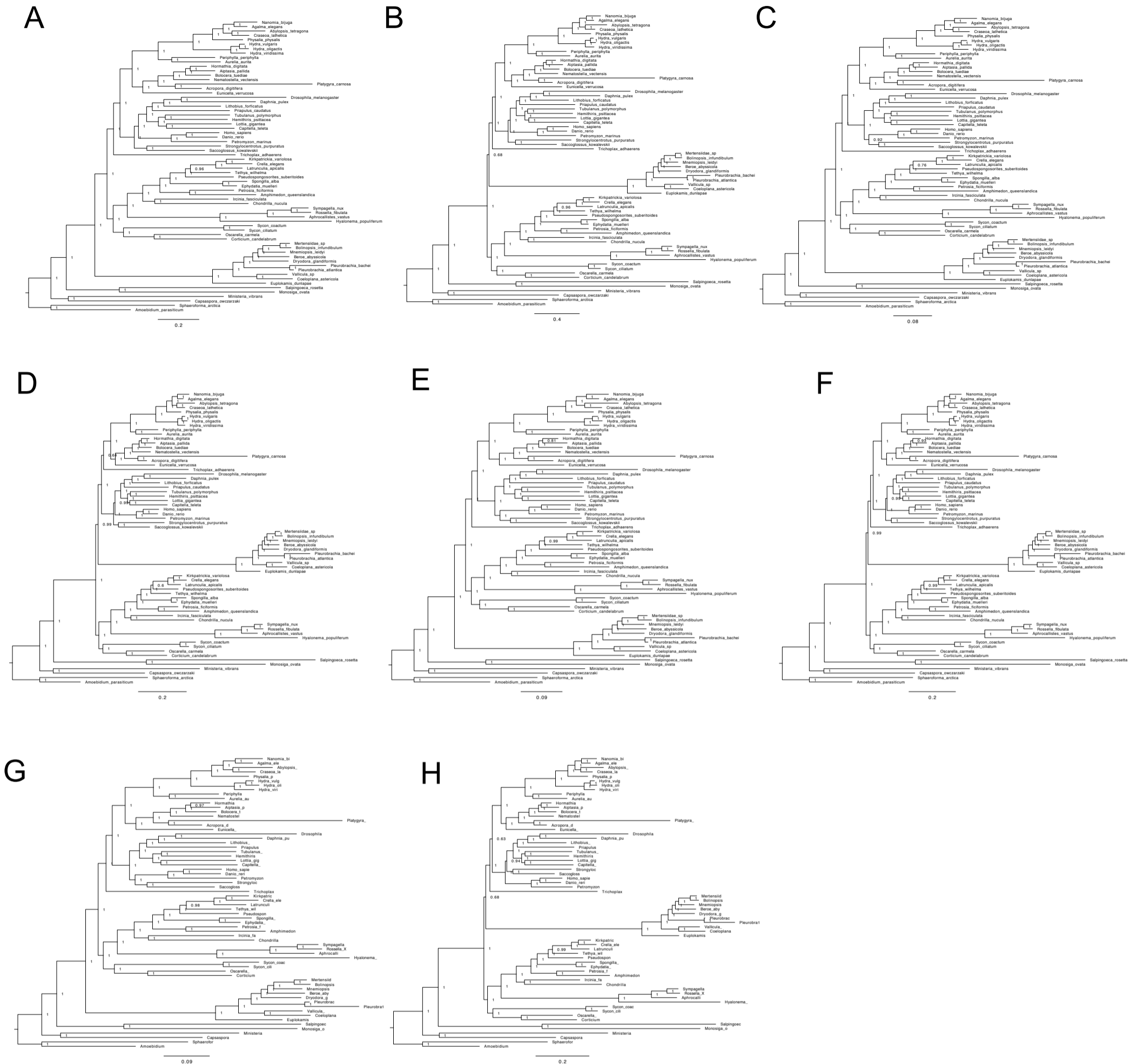
**Improved Modeling of Compositional Heterogeneity**

**Supports Sponges as Sister to All Other Animals**

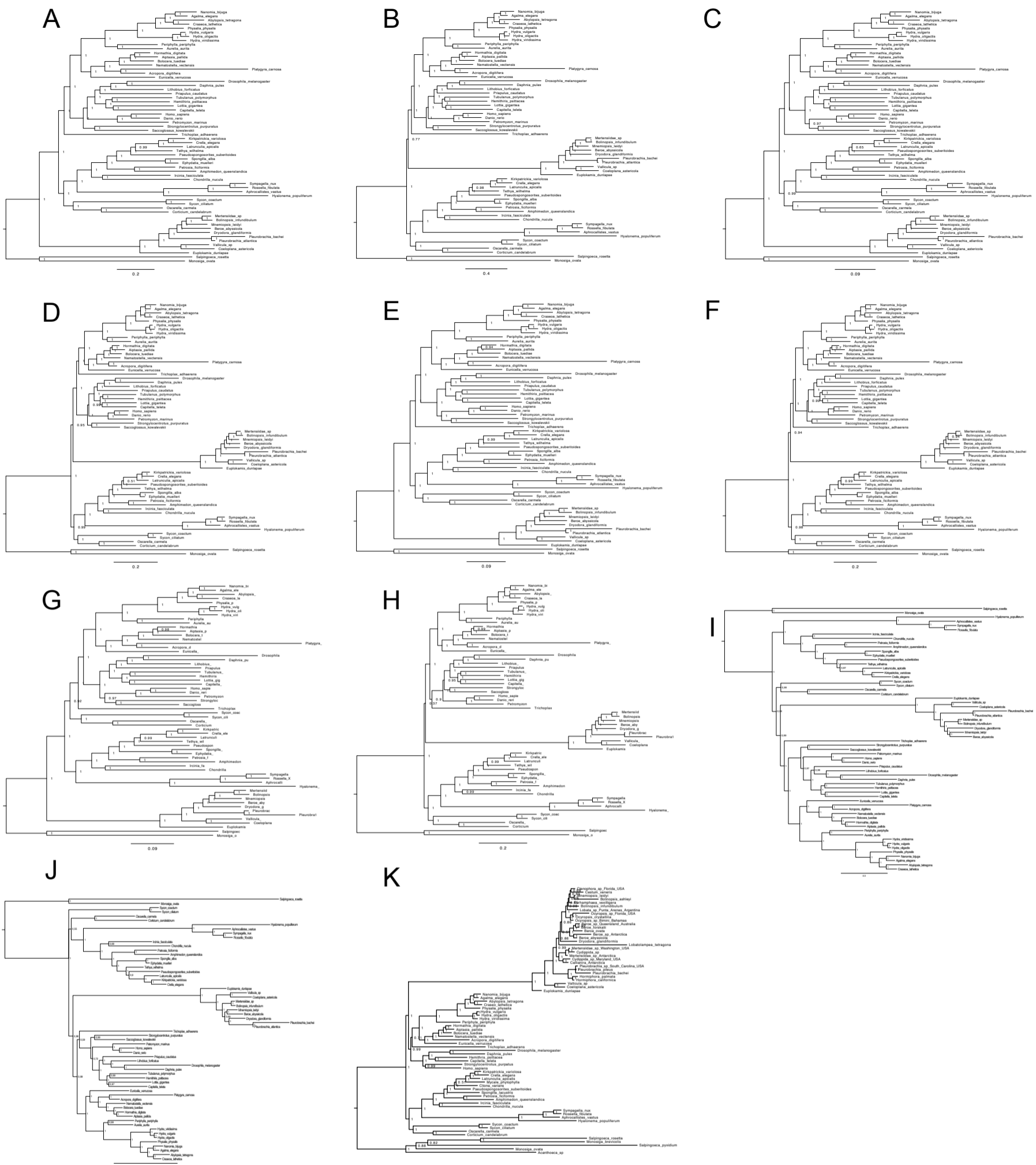
**Roberto Feuda, Martin Dohrmann, Walker Pett, Hervé Philippe, Omar Rota-Stabelli, Nicolas Lartillot, Gert Wörheide, and Davide Pisani**



**Figure S1. Results of phylogenetic analyses of WhelanD20-Opistho, related to Figure 1, Table 1 and Table 3.** Numbers at nodes are Bayesian posterior probabilities. **(A)** Under WAG+G: Burnin = 1000 samples; Total number of cycles = 2273; Bpcomp Maxdiff = 0; Tracecomp minimal overall effsize = 214; Maximal overall rel\_diff = 0.17. **(B)** Under LG+G: Burnin = 500 samples; Total number of cycles = 1740; Bpcomp Maxdiff = 0; Tracecomp minimal overall effsize = 97; Maximal overall rel\_diff = 0.147486. **(C)** Under GTR+G: Burnin: 500 samples; Total number of cycles = 1679; Bpcomp Maxdiff = 0; Tracecomp minimal overall effsize = 98; Maximal overall rel\_diff = 0.281729. **(D)** Under CAT-GTR+G: Burnin = 20000 samples; Total number of cycles = 46067; Bpcomp Maxdiff = 0.0153846; Tracecomp minimal overall effsize = 40; Maximal overall rel\_diff = 0.383492. **(E)** Under GTR+G with Dayhoff-6 recoding: Burnin = 7000 samples; Total number of cycles 17035; Bpcomp Maxdiff = 0.02; Tracecomp minimal overall effsize = 733; Maximal overall rel\_diff = 0.0482448. **(F)** Under CAT-GTR+G with Dayhoff-6 recoding: Burnin = 10000 samples; Total number of cycles = 32512; Bpcomp Maxdiff = 0.0488889; Tracecomp minimal overall effsize = 732; Maximal overall rel\_diff = 0.080975. **(G)** Under GTR+G with S&R-6 recoding: Burnin = 5000 samples; Total number of cycles = 15132; Bpcomp Maxdiff = 0.056; Tracecomp minimal overall effsize = 406; Maximal overall rel\_diff = 0.13. **(H)** Under CAT-GTR+G with S&R-6 recoding: Burnin = 6000 samples; Total number of cycles = 16087; Bpcomp Maxdiff = 0.21; Tracecomp minimal overall effsize = 468; Maximal overall rel\_diff = 0.13. **(I)** Under GTR+G with KGB-6 recoding: Burnin = 5000 samples; Total number of cycles = 14381; Bpcomp Maxdiff = 0.0042; Tracecomp minimal overall effsize = 2107; Maximal overall rel\_diff = 0.06. **(J)** Under CAT-GTR+G with KGB-6 recoding: Burnin = 8000 samples; Total number of cycles = 21145; Bpcomp Maxdiff = 0.03; Tracecomp minimal overall effsize = 1767; Maximal overall rel\_diff = 0.1.



**Figure S2. Results of phylogenetic analyses of WhelanD20-Holo, related to Table 3.** Numbers at nodes are Bayesian posterior probabilities. **(A)** Under GTR+G: Burnin = 1000 samples; Total number of cycles = 2630; Bpcomp Maxdiff = 0.0617284; Tracecomp minimal overall effsize = 57; Maximal overall rel\_diff = 0.176365. **(B)** Under CAT-GTR+G: Burnin = 40000 samples; Total number of cycles = 88851; Bpcomp Maxdiff = 0.067623; Tracecomp minimal overall effsize = 41; Maximal overall rel\_diff = 0.656317. **(C)** Under GTR+G with Dayhoff-6 recoding: Burnin = 300 samples; Total number of cycles = 748; Bpcomp Maxdiff = 0.159091; Tracecomp minimal overall effsize = 84; Maximal overall rel\_diff = 0.408643. **(D)** Under CAT-GTR+G with Dayhoff-6 recoding: Burnin = 10000 samples; Total number of cycles = 30038; Bpcomp Maxdiff = 0.0475; Tracecomp minimal overall effsize = 121; Maximal overall rel\_diff = 0.179874. **(E)** Under GTR+G with S&R-6 recoding: Burnin = 5000 samples; Total number of cycles = 11673; Bpcomp Maxdiff = 0.102; Tracecomp minimal overall effsize = 558; Maximal overall rel\_diff = 0.12. **(F)** Under CAT-GTR+G with S&R-6 recoding: Burnin = 6000 samples; Total number of cycles = 16904; Bpcomp Maxdiff = 0.014; Tracecomp minimal overall effsize = 618; Maximal overall rel\_diff = 0.14. **(G)** Under GTR+G with KGB-6 recoding: Burnin = 5000 samples; Total number of cycles = 15605; Bpcomp Maxdiff = 0.17; Tracecomp minimal overall effsize = 2549; Maximal overall rel\_diff = 0.04. **(H)** Under CAT-GTR+G with KGB-6 recoding: Burnin = 8000 samples; Total number of cycles = 16906; Bpcomp Maxdiff = 0.19; Tracecomp minimal overall effsize = 1273; Maximal overall rel\_diff = 0.15.



**Figure S3. Results of phylogenetic analyses: trees inferred from WhelanD20-Choano, WhelanD16-OutlierExcluded-Choano, and Whelan2017-all-outgroups. Related to Figure 1 and Table 3.** Numbers at nodes are Bayesian posterior probabilities. **(A)** Under GTR+G: Burnin = 1000 samples; Total number of cycles = 2832; Bpcomp Maxdiff = 0.010989; Tracecomp minimal overall effsize = 62; Maximal overall rel\_diff = 0.194923. **(B)** Under CAT-GTR+G: Burnin = 45000 samples; Total number of cycles = 87710; Bpcomp Maxdiff = 0.0491803; Tracecomp minimal overall effsize = 69; Maximal overall rel\_diff = 0.399076. **(C)** Under GTR+G with Dayhoff-6 recoding: Burnin = 1000 samples; Total number of cycles = 3338; Bpcomp Maxdiff = 0.0689655; Tracecomp minimal overall effsize = 176; Maximal overall rel\_diff = 0.244371. **(D)** Under CAT-GTR+G with Dayhoff-6 recoding: Burnin = 10000 samples; Total number of cycles = 21700; Bpcomp Maxdiff = 0.025641; Tracecomp minimal overall effsize = 227; Maximal overall rel\_diff = 0.12302. **(E)** Under GTR+G with S&R-6 recoding: Burnin = 5000 samples; Total number of cycles = 29256; Bpcomp Maxdiff = 0.03; Tracecomp minimal overall effsize = 2658; Maximal overall rel\_diff = 0.028. **(F)** Under CAT-GTR+G with S&R-6 recoding: Burnin = 6000 samples; Total number of cycles = 17184; Bpcomp Maxdiff = 0.073; Tracecomp minimal overall effsize = 613; Maximal overall rel\_diff = 0.1. **(G)** Under GTR+G with KGB-6 recoding: Burnin = 6000 samples; Total number of cycles = 16672; Bpcomp Maxdiff = 0.013; Tracecomp minimal overall effsize = 2347; Maximal overall rel\_diff = 0.04. **(H)** Under CAT-GTR+G with KGB-6 recoding: Burnin = 8000 samples; Total number of cycles = 22913; Bpcomp Maxdiff = 0.06; Tracecomp minimal overall effsize = 1666; Maximal overall rel\_diff = 0.147. **(I)** WhelanD16-OutlierExcluded-Choano, under CAT-GTR+G (see STAR Methods – Testing the distribution of the signal in favour of Porifera-sister). Burnin = 5000 samples; Total number of cycles = 17581; Bpcomp Maxdiff = 0.0516; Tracecomp minimal overall effsize = 87; Maximal overall rel\_diff = 0.414. **(J)** WhelanD16-OutlierExcluded-Choano, under CAT-GTR+G with Dayhoff-6 recoding (see STAR Methods – Testing the distribution of the signal in favour of Porifera-sister). Burnin = 12000 samples; Total number of cycles = 41550; Bpcomp Maxdiff = 0.0511; Tracecomp minimal overall effsize = 1703; Maximal overall rel\_diff = 0.068. **(K)** Whelan2017\_Metazoa\_Choano\_RCFV\_strict, under CAT-GTR+G with Dayhoff-6 recoding (see STAR Methods – Testing the effect of incrementing the number of ctenophoran lineages on the phylogenetic stability of recoded dataset). Burnin = 7000 samples; Total number of cycles = 25851; Bpcomp Maxdiff = 0.0076; Tracecomp minimal overall effsize = 1572; Maximal overall rel\_diff = 0.044.





Dayhoff-6	GTR	15.48	23.78	89.92
Dayhoff-6	CAT-GTR	10.54	13.56	43.34
S&R-6	GTR	20.19	31.10	74.17
S&R-6	CAT-GTR	10.79	12.18	34.43
KGB-6	GTR	18.57	21.03	86.42
KGB-6	CAT-GTR	10.70	6.63	36.62

---

	None	WAG+G	34.38	32.45	112.60
	None	LG+G	42.56	40.10	92.09
	None	GTR+G	38.13	41.63	100.26
	None	PF-scheme	39.67	36.86	61.04
Chang	None	CAT- GTR+G	12.58	11.70	53.66
	Dayhoff-6	GTR+G	22.85	19.52	59.02
	Dayhoff-6	CAT- GTR+G	17.72	15.28	7.94
	S&R-6	GTR+G	25.45	18.68	46.81
	S&R-6	CAT- GTR+G	18.75	16.05	8.59

KGB-6	GTR+G	22.87	12.11	25.96
KGB-6	CAT- GTR+G	16.79	8.31	7.60

---



**Table S2. Detailed results of PPA analyses (PPA-DIV and PPA-MAX), Related to Table 1.** This table illustrates the empirical heterogeneity observed directly from the data, the average posterior predictive mean and the dispersion around the mean for the mean amino acid diversity and maximal heterogeneity across taxa (the two measures used in Tables 1). These results illustrate that for all models, dispersion is comparable as long as the same recoding scheme is considered (as expected), and invariably tight (as expected given the large dimension of the datasets). Note: Observed, empirical, heterogeneity is not estimated but observed from the data. Accordingly, for each considered statistics and dataset this value does not change unless the recoding strategy changes. See [https://bitbucket.org/bzxdp/feuda\\_et\\_al\\_2017](https://bitbucket.org/bzxdp/feuda_et_al_2017) for the output files from the PPA analyses including the values summarised in this table.

Posterior predictive test		Mean amino acid diversity PPA-DIV			Maximal heterogeneity across the taxa PPA-MAX		
Dataset	Model & Recoding	Observed (empirical) Value	Posterior Predictive Mean	+/-	Observed (empirical) Value	Posterior Predictive Mean	+/-
WhelanD20	WAG	5.546	6.761	0.010	0.00316	0.00014	0.00004
	LG	5.546	6.612	0.010	0.00316	0.00014	0.00004
	PF SCHEME	5.546	6.572	0.010	0.00316	0.00022	0.00007
	GTR	5.546	6.562	0.011	0.00316	0.00015	0.00004
	CAT-GTR	5.546	5.604	0.009	0.00316	0.00030	0.00008
	Dayhoff-6 GTR	2.747	3.061	0.005	0.00254	0.00014	0.00007
	Dayhoff-6 CAT-GTR	2.747	2.742	0.005	0.00254	0.00030	0.00009

	S&R-6 GTR	2.893	3.222	0.005	0.00197	0.00014	0.00006
	S&R-6 CAT-GTR	2.893	2.887	0.005	0.00197	0.00061	0.00019
	KGB-6 GTR	2.724	3.013	0.005	0.00255	0.00014	0.00007
	KGB-6 CAT-GTR	2.724	2.723	0.004	0.00255	0.00038	0.00016
	WAG	4.080	5.569	0.008	0.00185	0.00021	0.00007
	LG	4.080	5.400	0.008	0.00185	0.00023	0.00008
	PF SCHEME	4.080	5.381	0.008	0.00185	0.00032	0.00010
	GTR	4.080	5.162	0.009	0.00185	0.00024	0.00009
	CAT-GTR	4.080	4.124	0.007	0.00185	0.00062	0.00010
	Dayhoff-6 GTR	2.153	2.454	0.004	0.00391	0.00020	0.00014
Chang	Dayhoff-6 CAT-GTR	2.153	2.147	0.004	0.00391	0.00273	0.00032
	S&R-6 GTR	2.247	2.573	0.005	0.00317	0.00021	0.00012
	S&R-6 CAT-GTR	2.247	2.241	0.004	0.00317	0.00216	0.00032
	KGB-6 GTR	2.145	2.449	0.004	0.00130	0.00019	0.00013
	KGB-6 CAT-GTR	2.145	2.142	0.003	0.00130	0.00102	0.00023