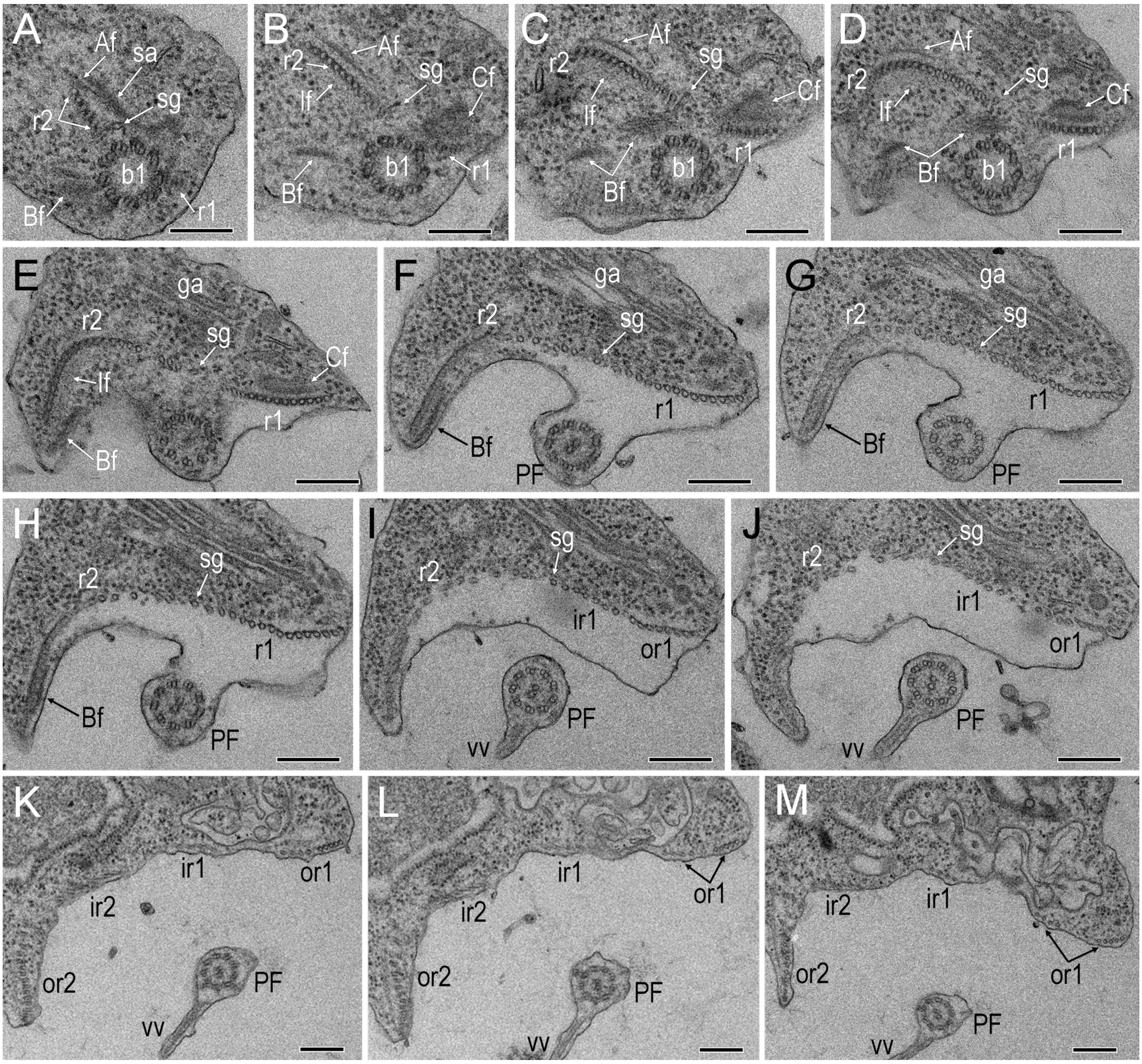

Figure S1. Transmission electron micrographs of *Agogonia voluta* gen et sp. nov. (A–J) Serial transverse sections of anterior part of the cell, showing origin and orientations of roots and fibres, with a more clearly visualized left root. (K–L) Serial transverse sections showing localization of the roots in the middle part of the cell. Scale bars: 200 nm. Af – 'A' fibre; b1 – basal body of the posterior flagellum; Bf – 'B' fibre; Cf – 'C' fibre; ga – Golgi apparatus; If – 'I' fibre; ir1 – inner portion of left root; ir2 – inner portion of right root; or1 – outer portion of left root; or2 – outer portion of right root; PF – posterior flagellum; r1 – left root; r2 – right root; sa – singlet-associated fibre; sg – singlet root; vv – ventral vane of posterior flagellum.

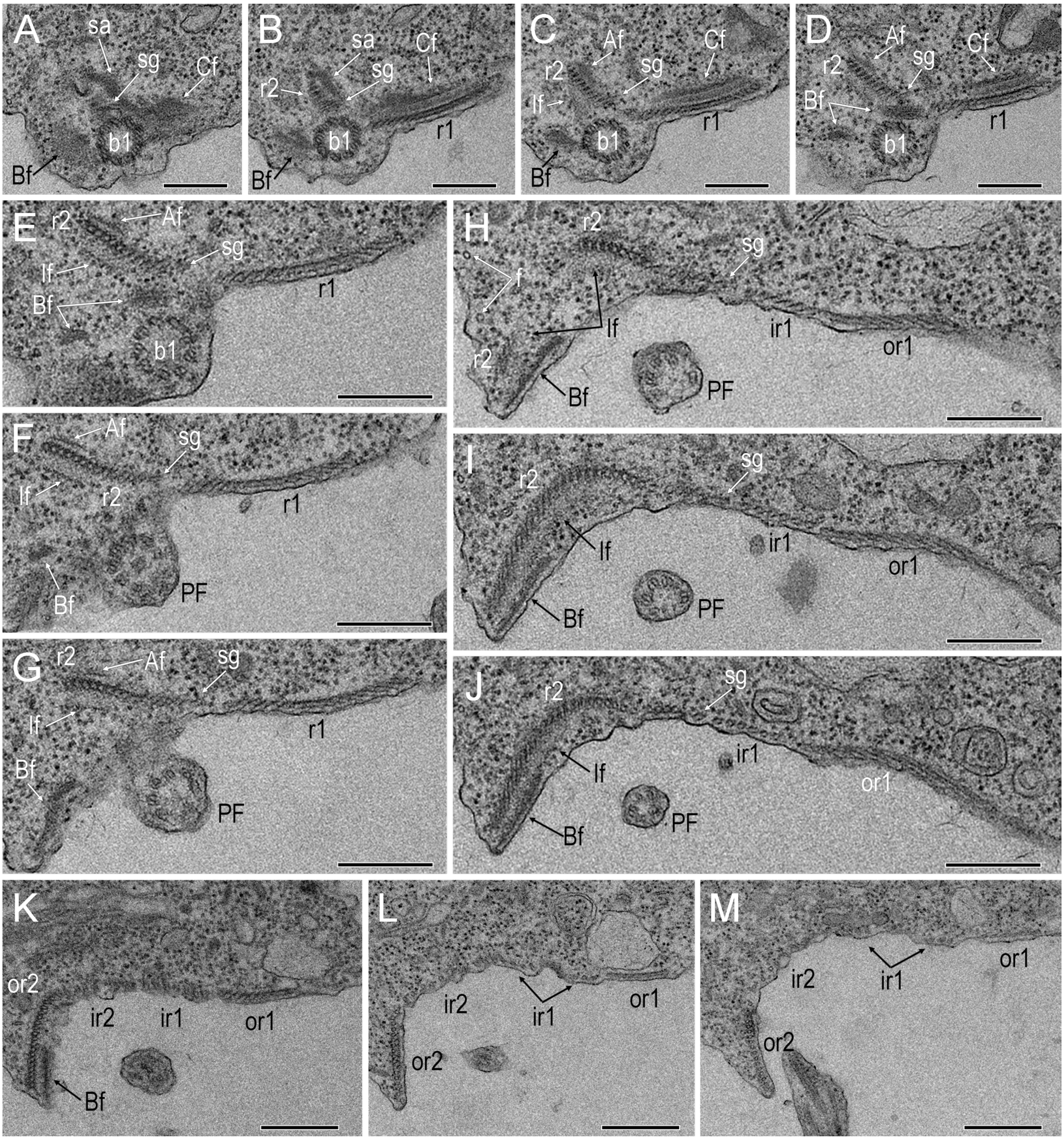

Figure S2. Transmission electron micrographs of *Agogonia voluta* gen et sp. nov. (A–J) Serial transverse sections of anterior part of the cell, showing origin and orientations of roots and fibres, with a more clearly visualized right root. (K–M) Serial transverse sections showing localization of the roots in the middle part of the cell. Scale bars: 300 nm. Af – 'A' fibre; b1 – basal body of the posterior flagellum; Bf – 'B' fibre; Cf – 'C' fibre; f – dorsal fan microtubules; If – 'I' fibre; ir1 – inner portion of left root; ir2 – inner portion of right root; or1 – outer portion of left root; or2 – outer portion of right root; PF – posterior flagellum; r1 – left root; r2 – right root; sa – singlet-associated fibre; sg – singlet root.

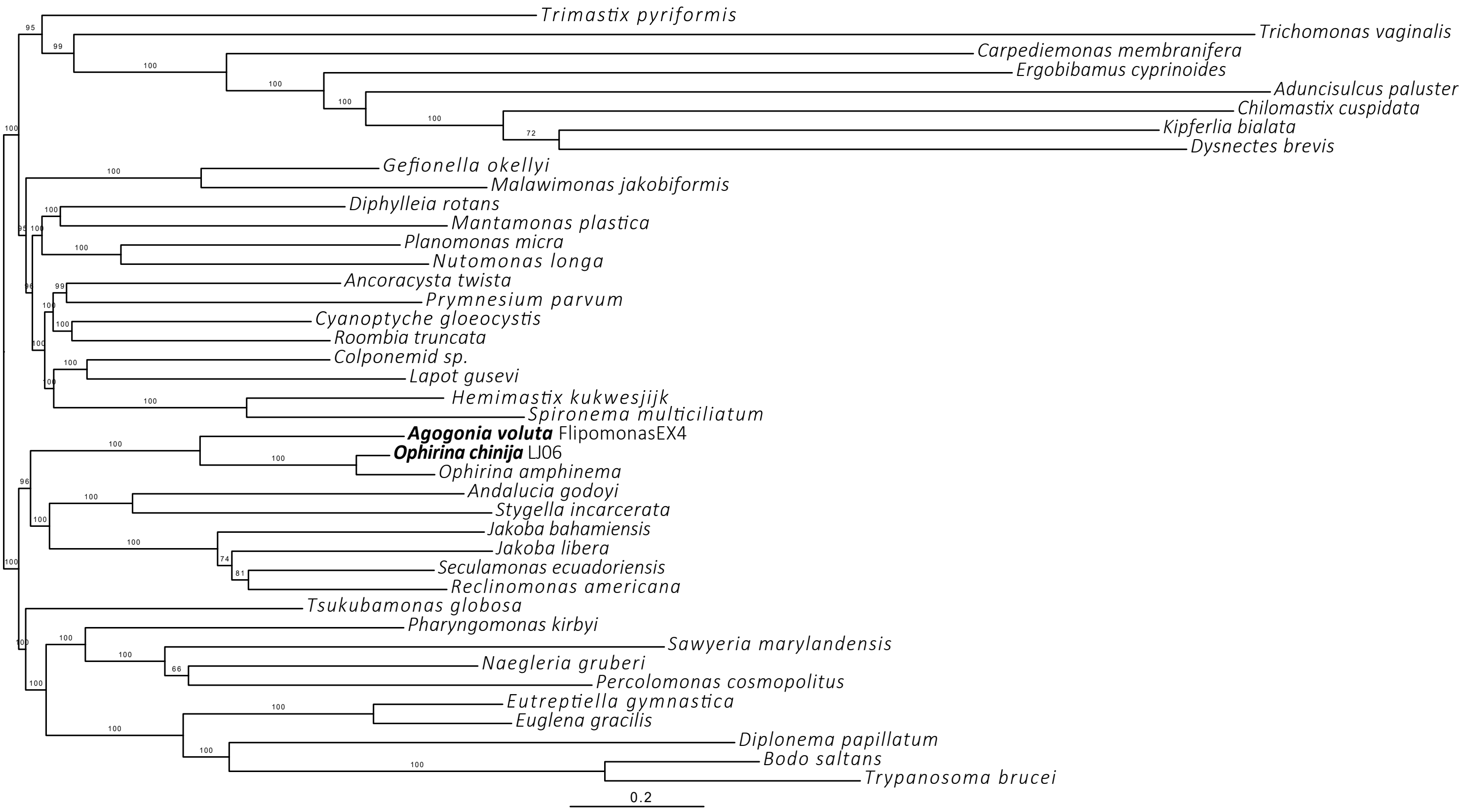

Figure S3. Maximum likelihood tree based on the dataset from Lax et al. 2018. The tree was reconstructed using 350 conserved proteins, 41 species, and 112,726 conserved amino acid positions, it was inferred with IQ-TREE using the PMSF approximation of the model LG+C60+F+R7 and ultrafast bootstrap as statistical support (shown over the branches).

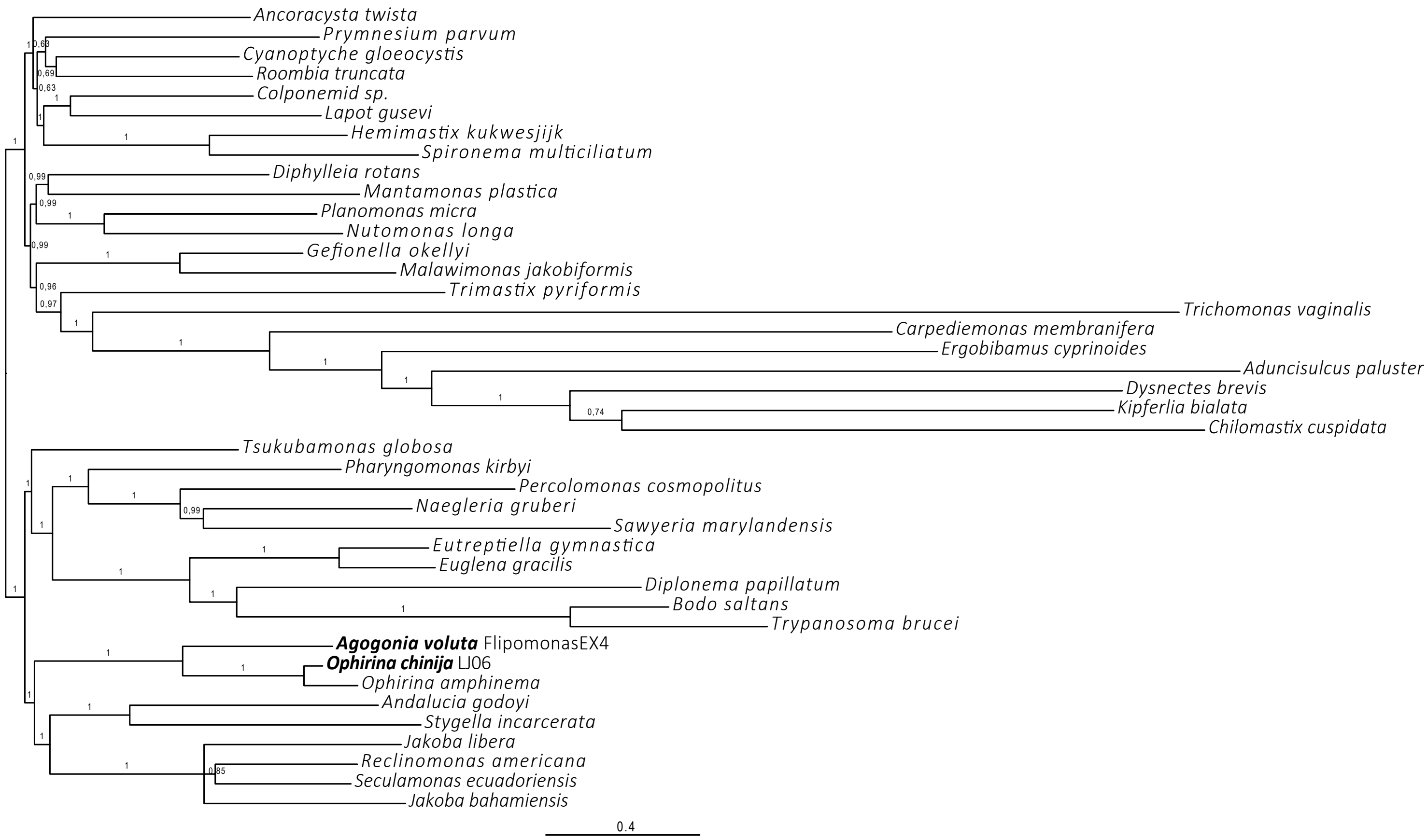

Figure S4. Bayesian inference tree based on the dataset from Lax et al. 2018. The tree was reconstructed using 350 conserved proteins, 41 species, and 112,726 conserved amino acid positions, it was inferred with Phylobayes using the CAT-GTR model and bayesian posterior probabilities as statistical support (shown over the branches).

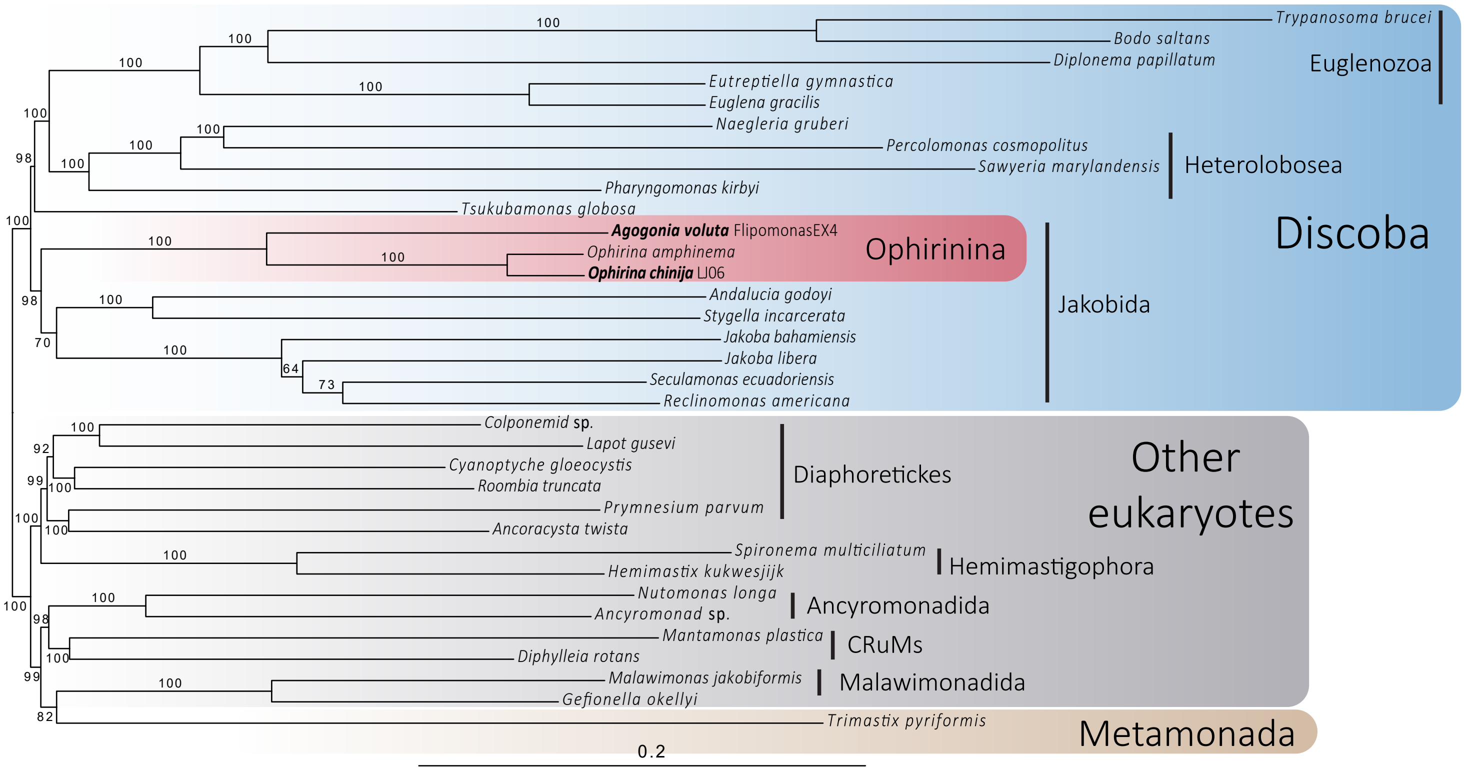

Figure S5. Maximum likelihood tree based on the dataset from Lax et al. 2018. The tree was reconstructed using 350 conserved proteins, 34 species (with the short-branching species Trimastix pyriformis as only representative of Metamonada), and 112,726 conserved amino acid positions. It was inferred with IG-TREE using the PMSF approximation of the model LG+C60+F+R7 and ultrafast bootstrap as statistical support (shown over the branches).

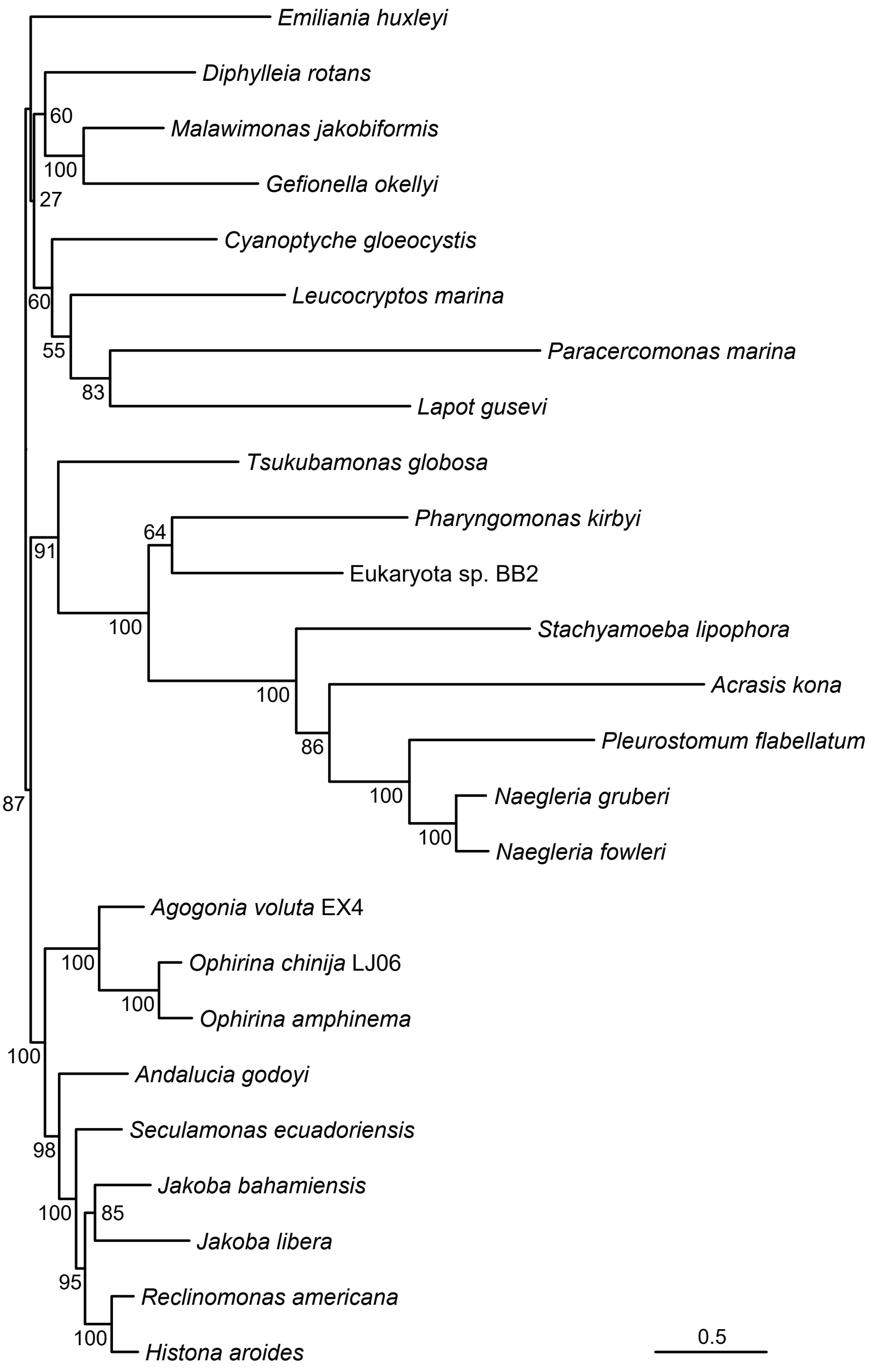

Figure S6. Maximum likelihood tree based on 14 conserved mitochondrion-encoded proteins. The tree was reconstructed using 25 species and 4,663 conserved amino acid positions. In was inferred with IQ-TREE using the PMSF approximation of the model LG+C60+F+R7 and ultrafast bootstrap as statistical support (shown next to the branches)

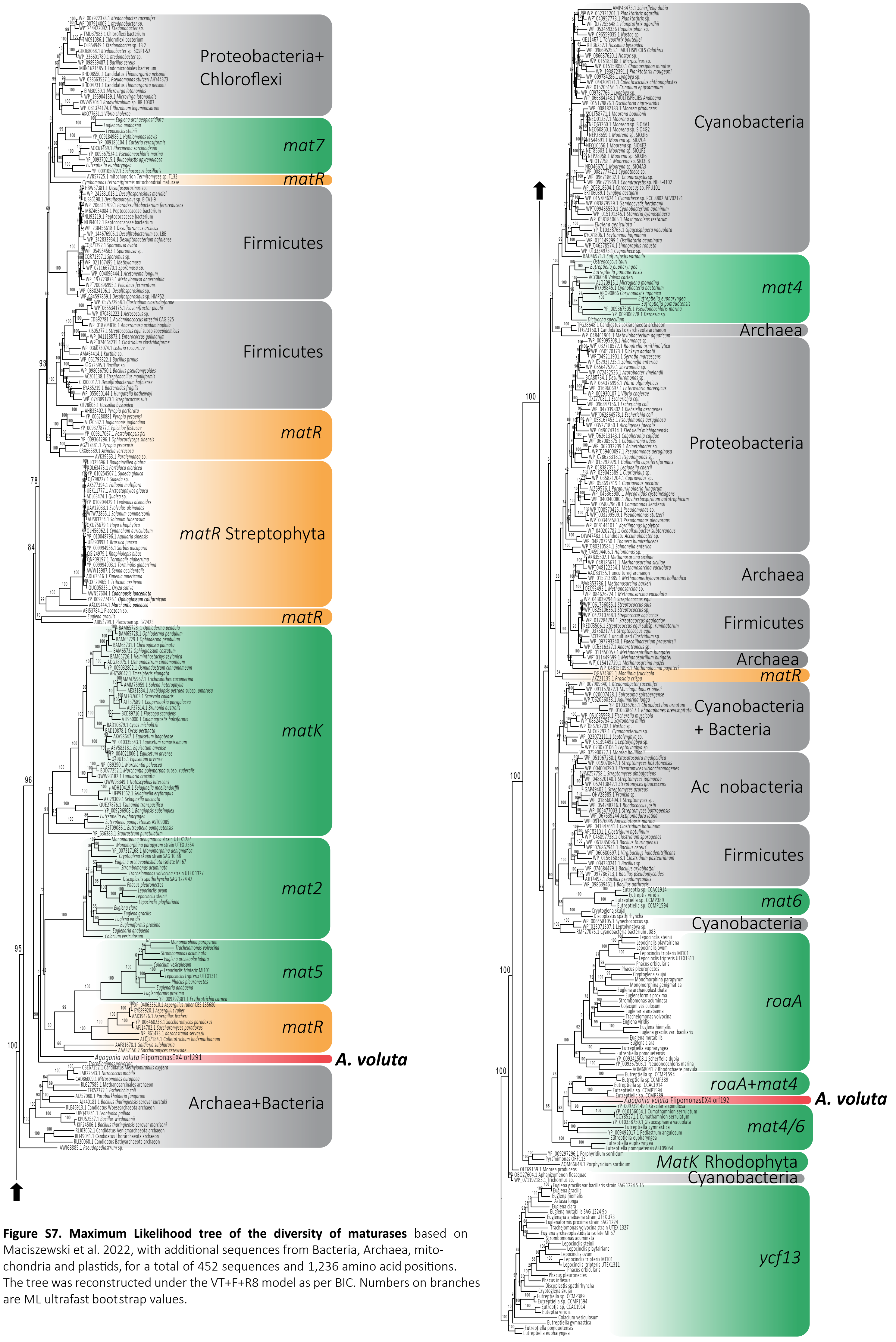

Figure S7. Maximum Likelihood tree of the diversity of maturases based on Maciszewski et al. 2022, with additional sequences from Bacteria, Archaea, mitochondria and plastids, for a total of 452 sequences and 1,236 amino acid positions. The tree was reconstructed under the VT+F+R8 model as per BIC. Numbers on branches are ML ultrafast bootstrap values.