**Figures S1 – S9.** Maximum likelihood phylogenetic trees for each specific EukRef database constructed in RAxML using GTRCAT model. Statistical support is derived from 100 bootstrap replicates. Euglenida and Glycomonada datasets were each split into two separate tree figures as they both contain very large number of sequences. Trees are outgroup rooted. Tip labels are the final classification in the databases.

**Figure S1**
**Preaxostyla**

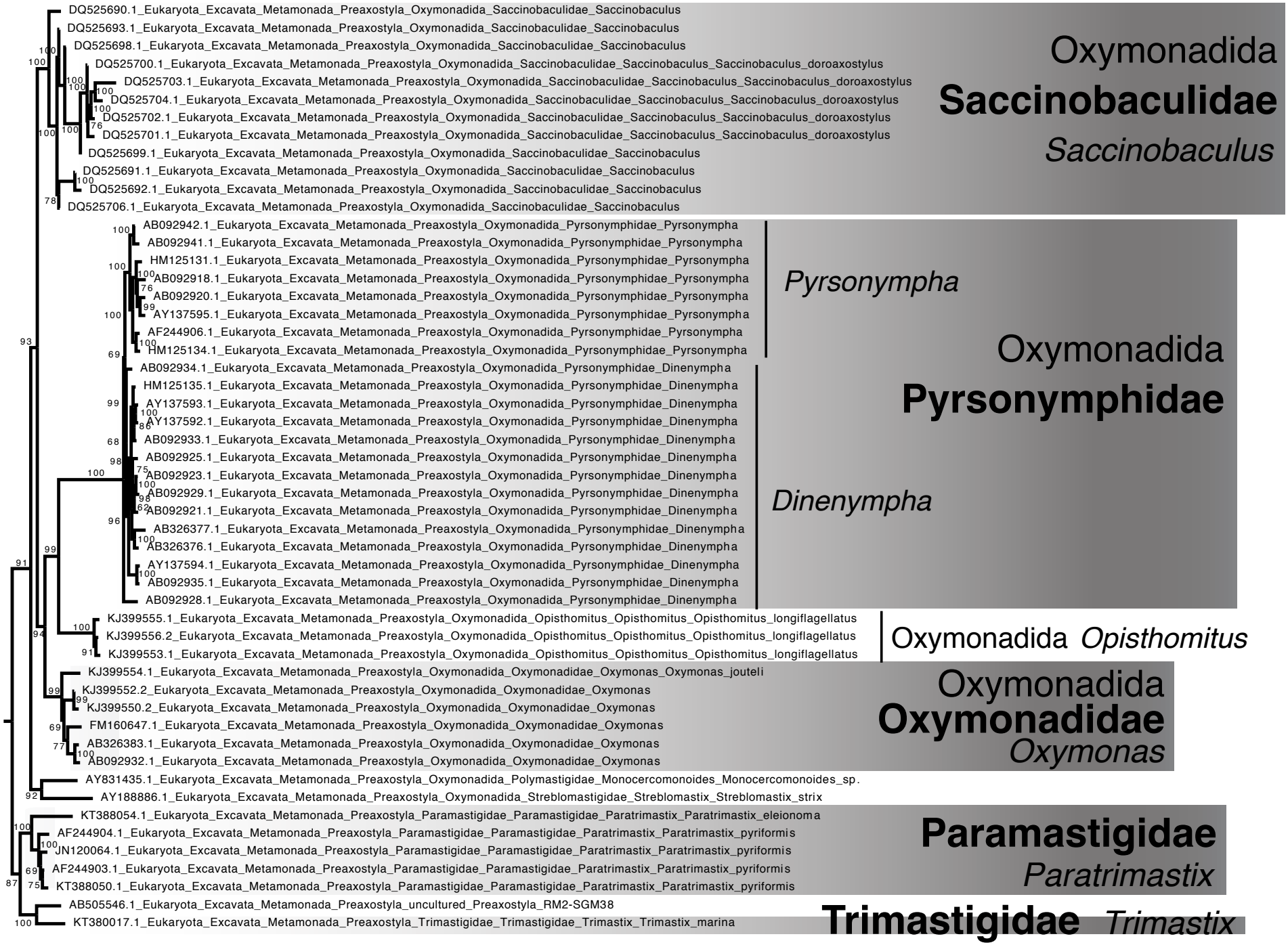


**Figure S2**
**Fornicata**

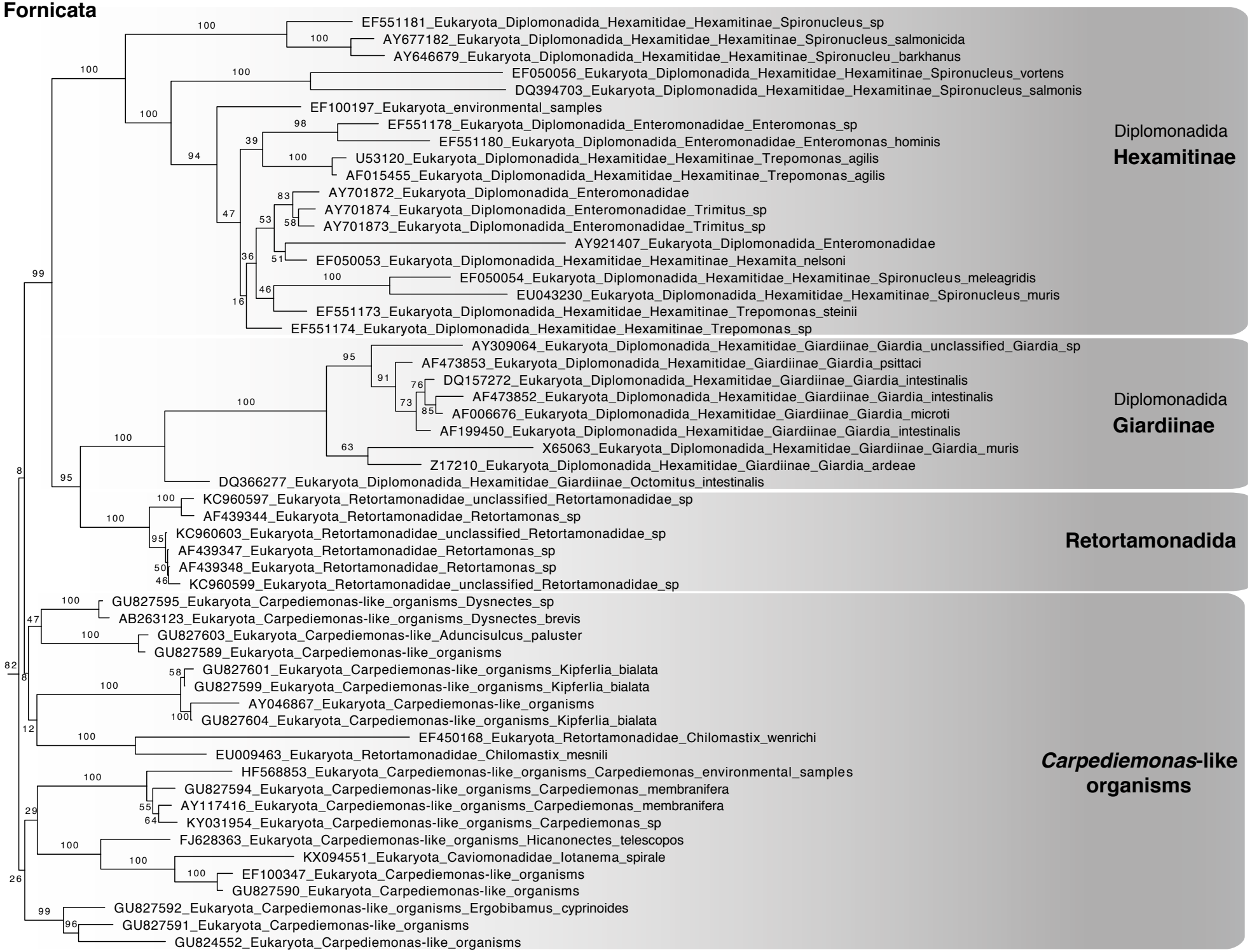


**Figure S3**
**Parabasalia**

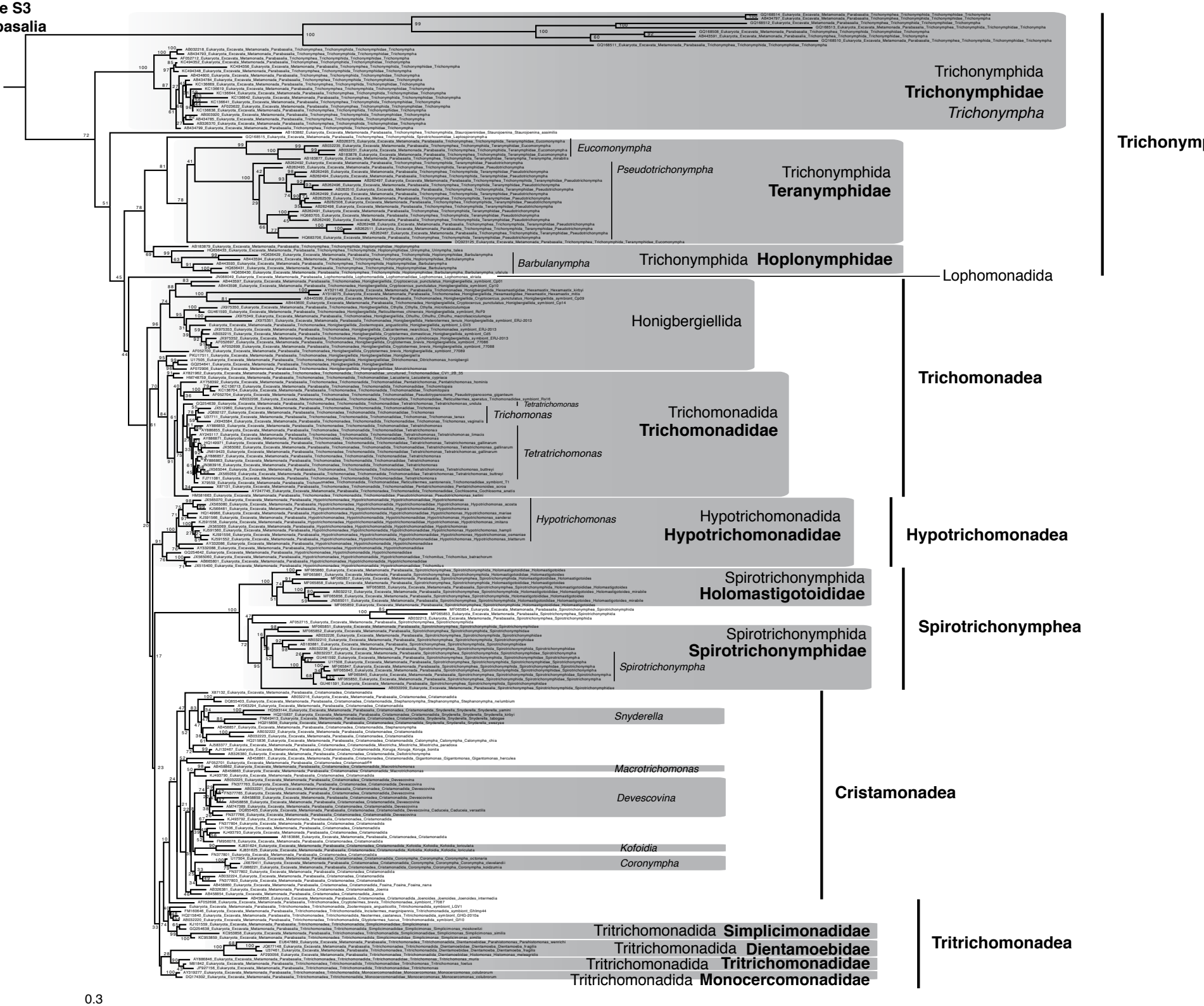
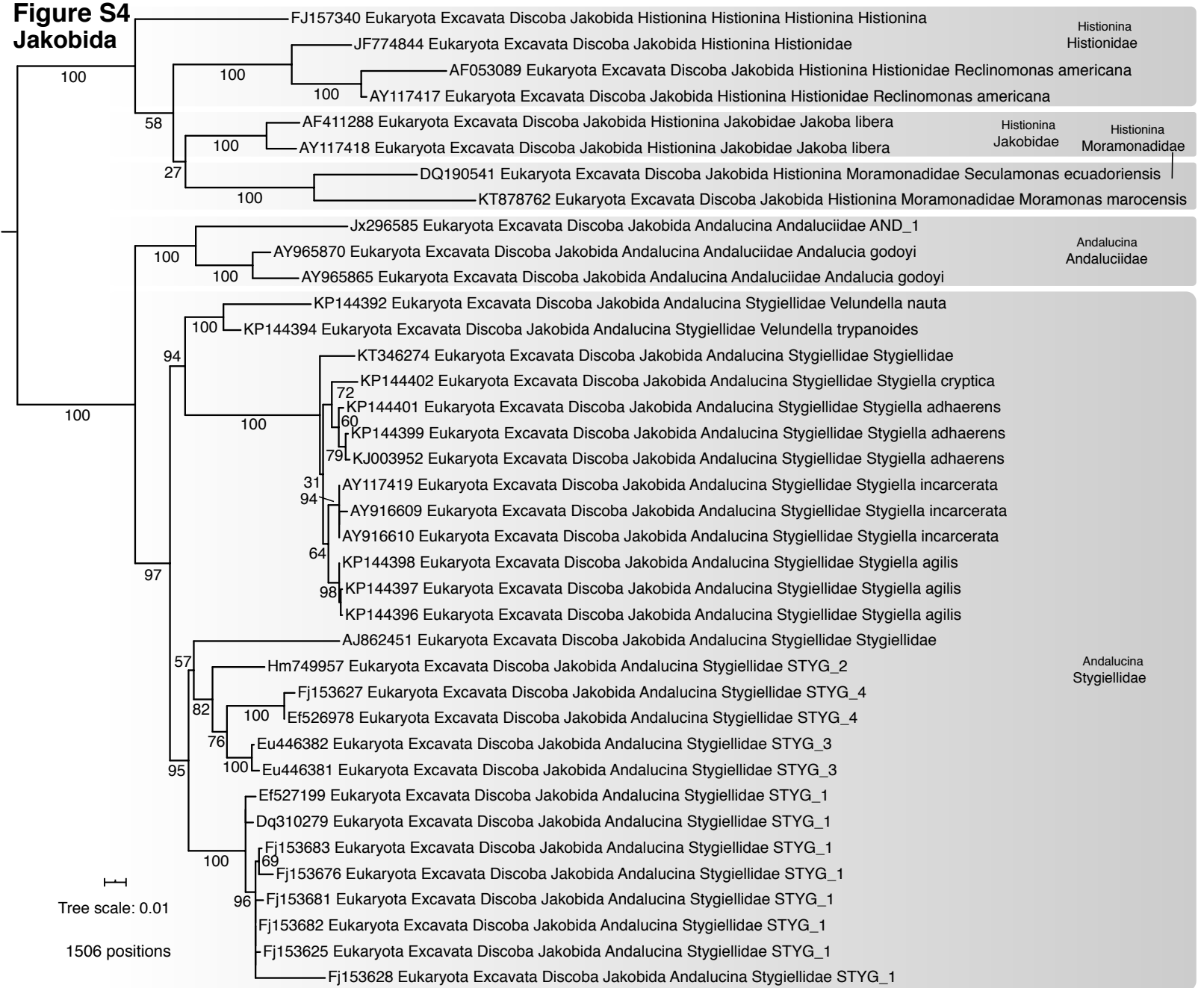


**Figure S4**
**Jakobida**



# Figure S5
# Heterolobosea

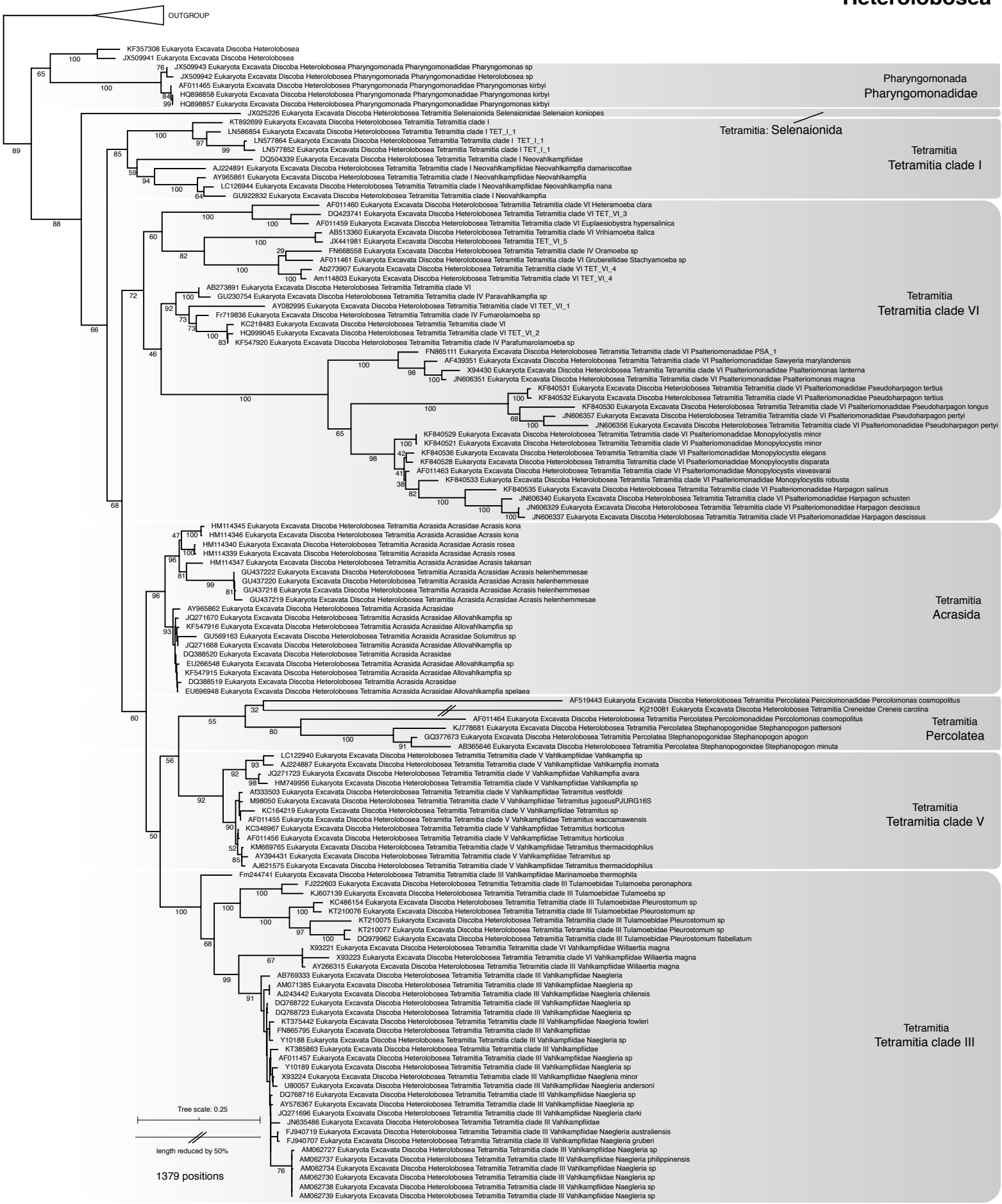


**Figure S6**
**Euglenida 1**

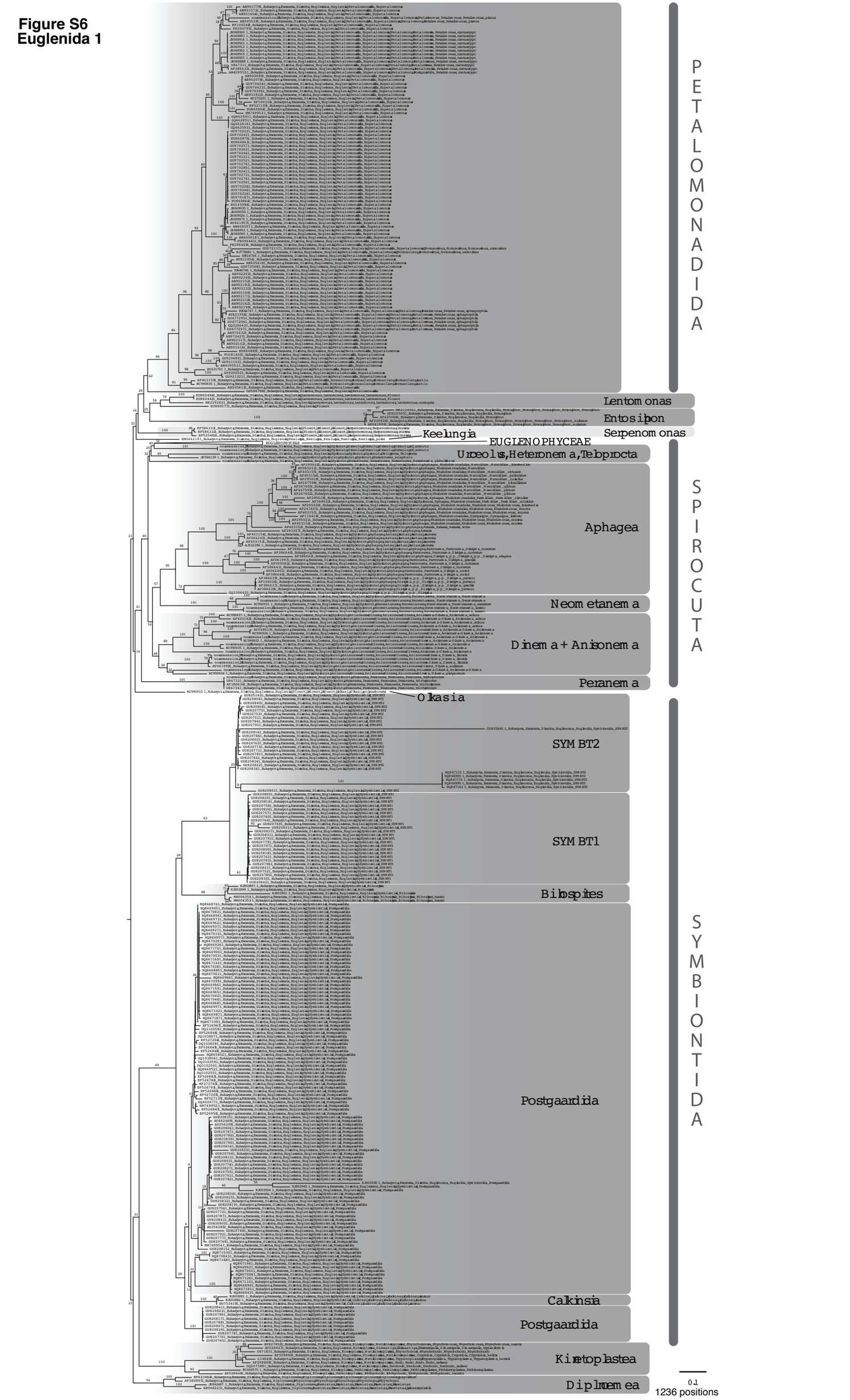


**Figure S7**
**Euglenida 2**

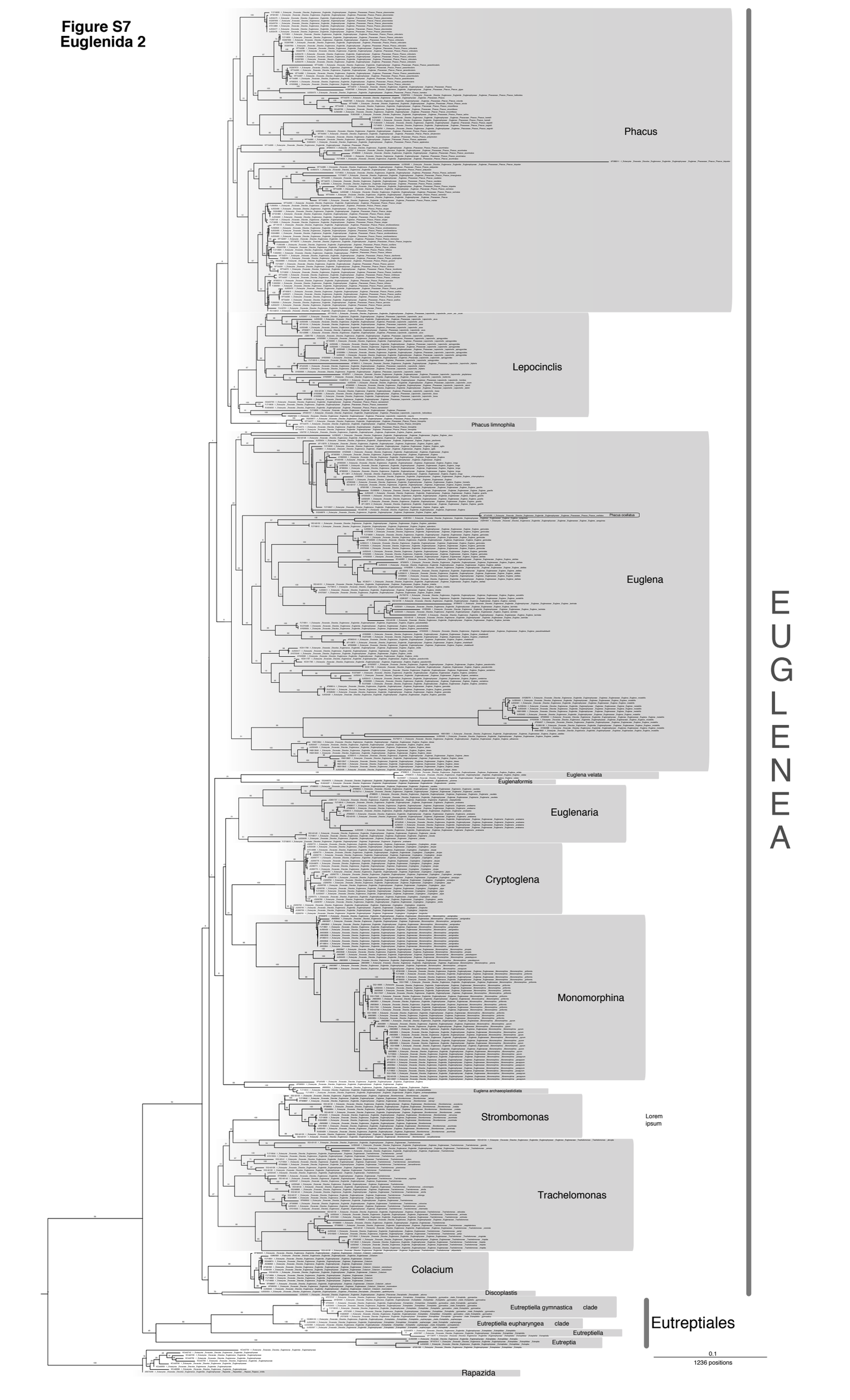


**Figure S8**

**Diplonemea**

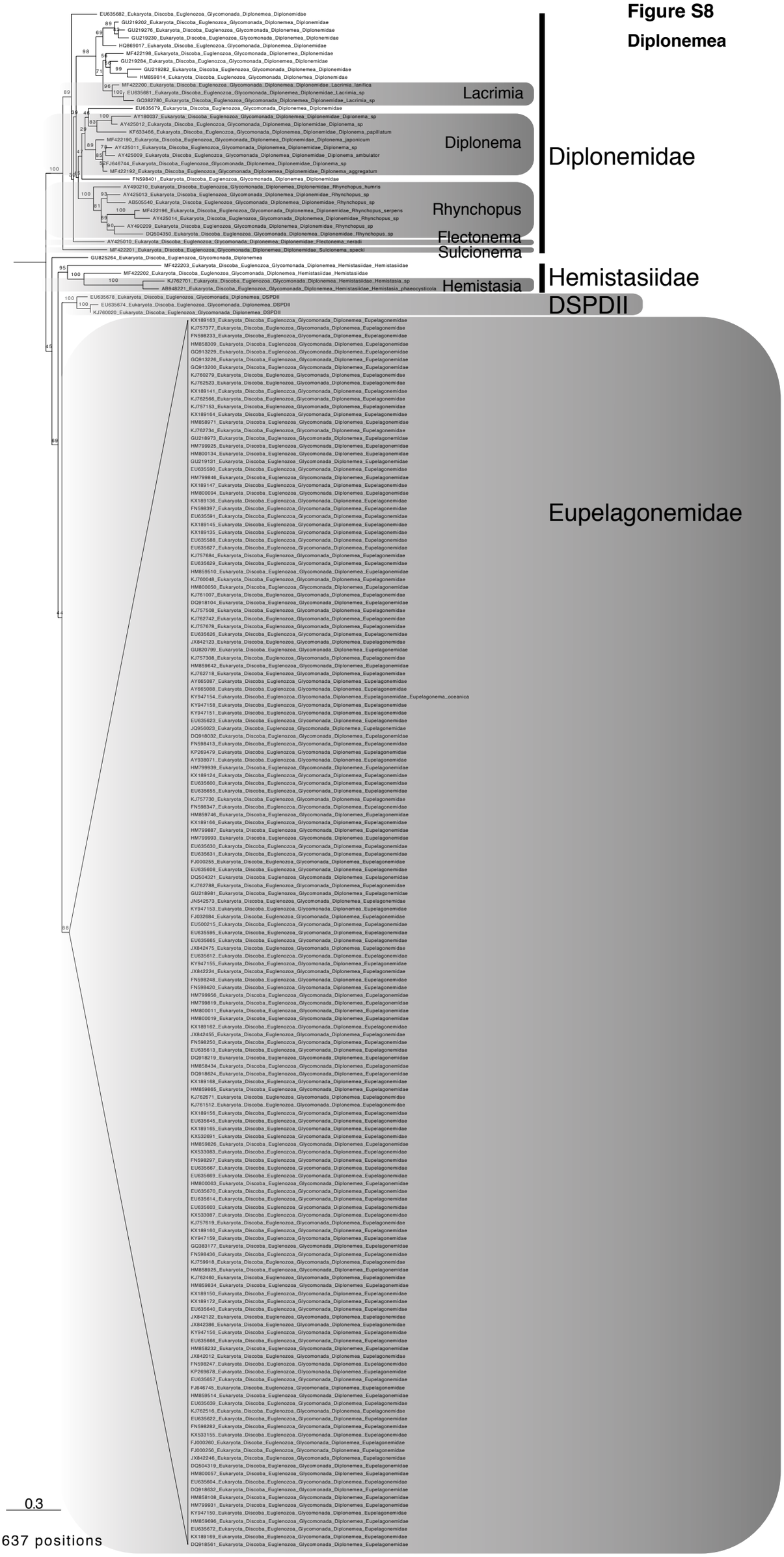


**Figure S9**
**Kinetoplastea**

KIN1

Prokinetoplastina

PRO1

Ichthyobodo

Perkinsela

PRO2

Metakinetoplastina

Neobodo

Neobodo

MET1

Azumiobodo

Cruzella

MET2

MET3

MET4

MET5

MET6

MET7

Rhynchomonas

Dimastigella

MET8

Rhynchobodo

Klosteria

MET9

MET10

Eubodonida

Parabodonida

Trypanosomatida

Parartrypanosoma

Wallacemonas

Lafontella

Herpetomonas

Phytomonas

Leishmaniinae

Blechomonas

Sergeia

Blastocrithidia

TRY1

Angomonas

Strigomonas

Kentomonas

Trypanosoma

1637 positions
0.3