SUPPLEMENTARY MATERIAL 3

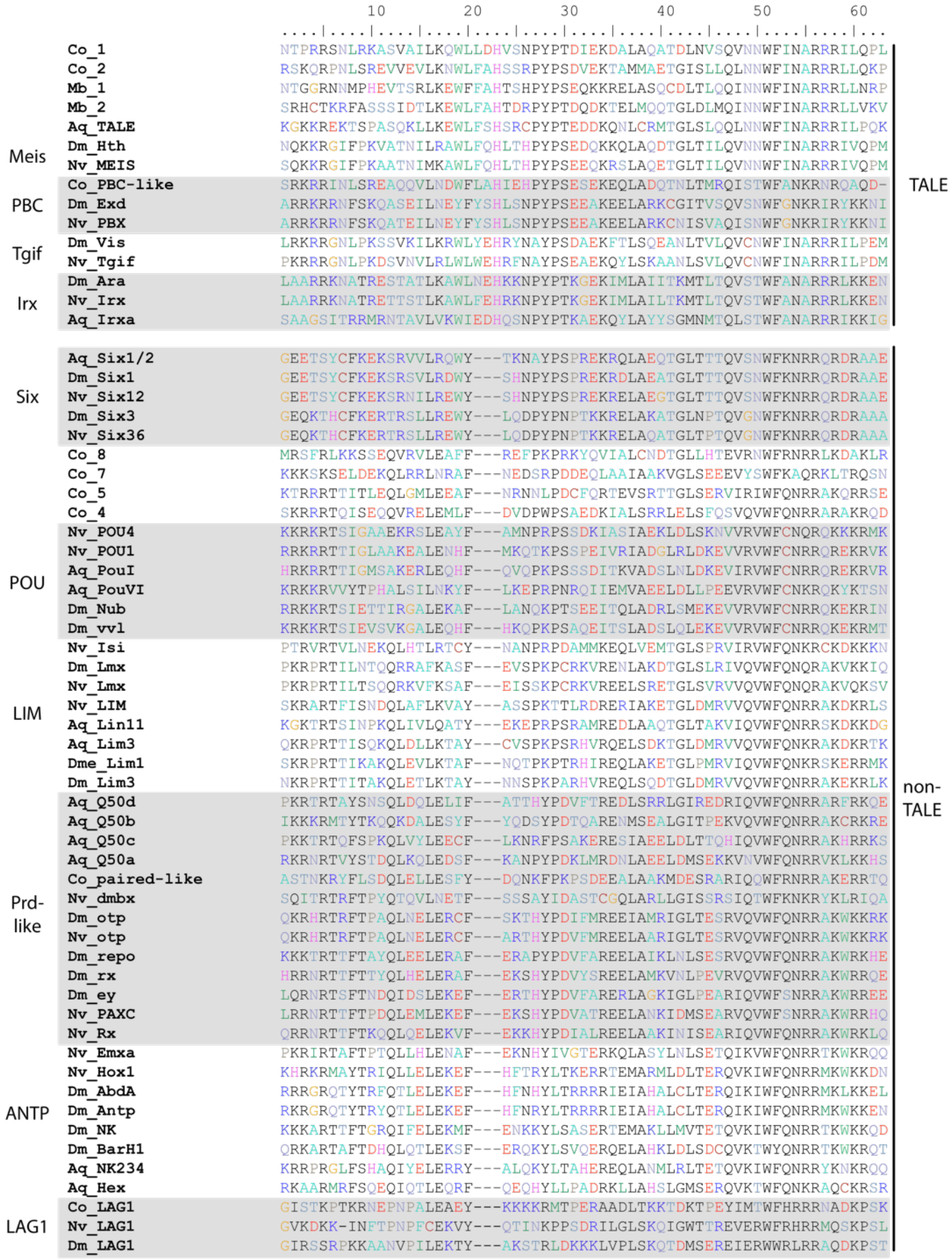

Figure S18. An illustrative homeobox alignment, showing the *Capsaspora* and *Monosiga* homeobox domains in contrast to other metazoan families. Taxa used includes Nv (*Nematostella vectensis*), Aq (*Amphimedon queenslandica*), Co (*Capsaspora owczarzaki*), Dm (*Drosophila melanogaster*) and Mb (*Monosiga brevicollis*).

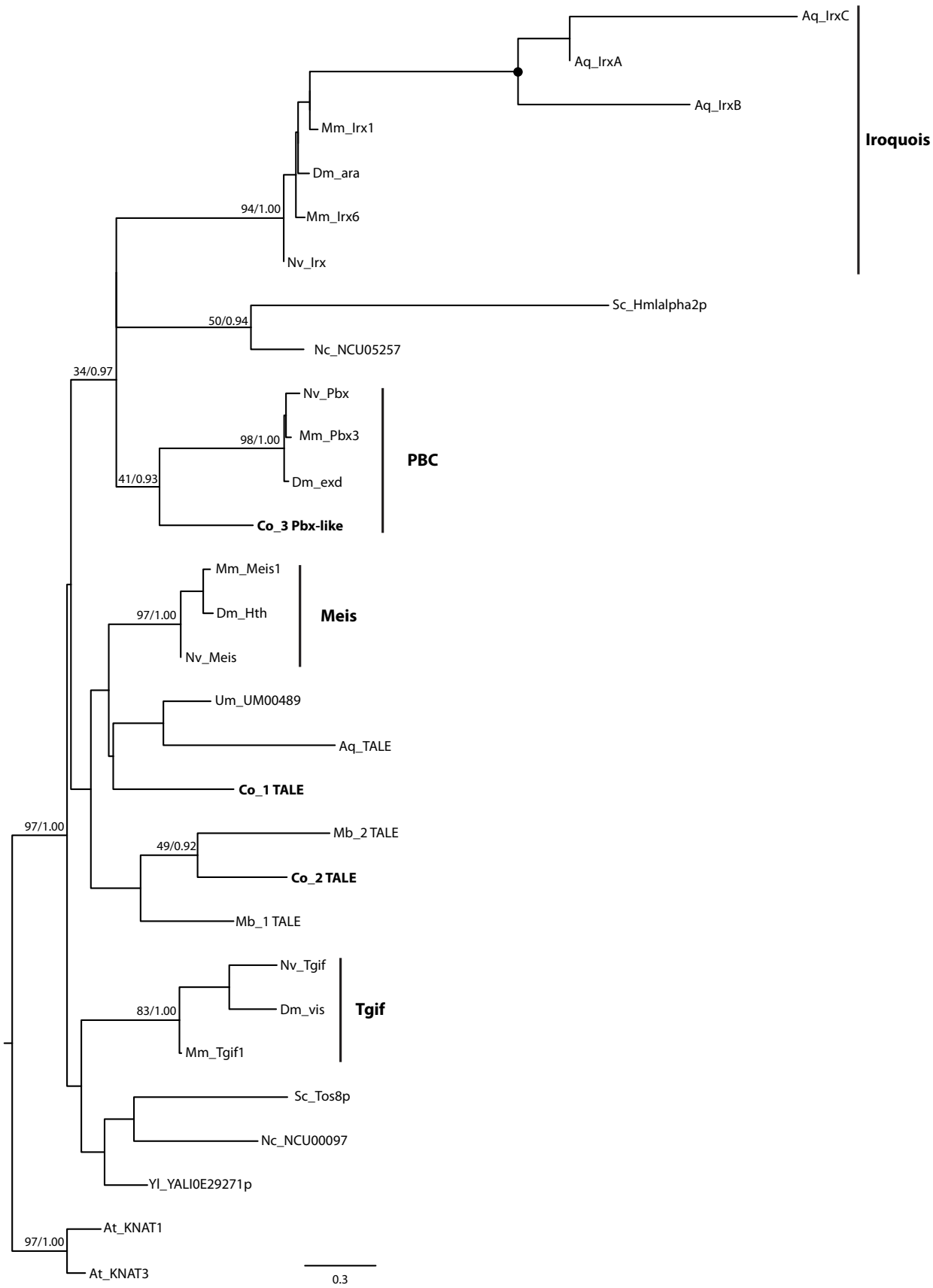

Figure S19. Maximum likelihood tree from the homeobox domain of TALE HDs. The tree is rooted using *Arabidopsis thaliana* sequences as outgroup. Statistical support was obtained by RAxML with 100-bootstrap replicates (bootstrap value, BV) and Bayesian Posterior Probabilities (BPP). Both values are shown on key branches. A black dot indicates BV > 90% and BPP > 0.95. Taxa used adapted from Larroux et al. Nv (*Nematostella vectensis*), Aq (*Amphimedon queenslandica*), Mb (*Monosiga brevicollis*), Co (*Capsaspora owczarzaki*), Dm (*Drosophila melanogaster*), Mm (*Mus musculus*), Sc (*Saccharomyces cerevisiae*), Um (*Ustilago maydis*), At (*Arabidopsis thaliana*), Nc (*Neurospora crassa*) and Yl (*Yarrowia lypolitica*).

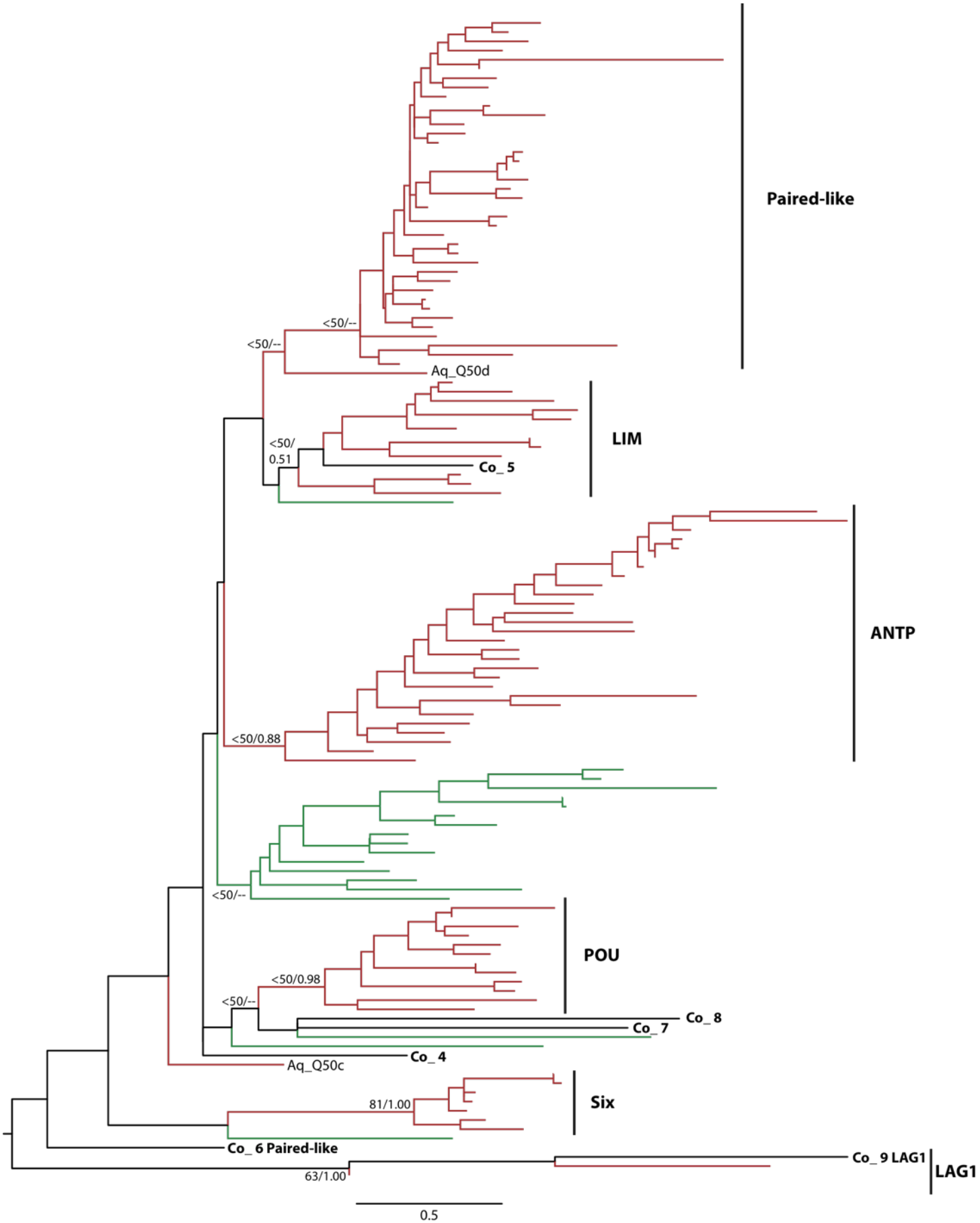

Figure S20. Maximum likelihood tree of non-TALE homeobox domains including all *Capsaspora* non-TALE homologs. The tree is rooted using the midpoint-rooted tree option. Statistical support was obtained by RAxML with 100-bootstrap replicates (BV) and Bayesian Posterior Probabilities (BPP). Both values are shown on key branches. Mb (*Monosiga brevicollis*), Co (*Capsaspora owczarzaki*). Metazoan branches depicted in red and fungal branches in green.

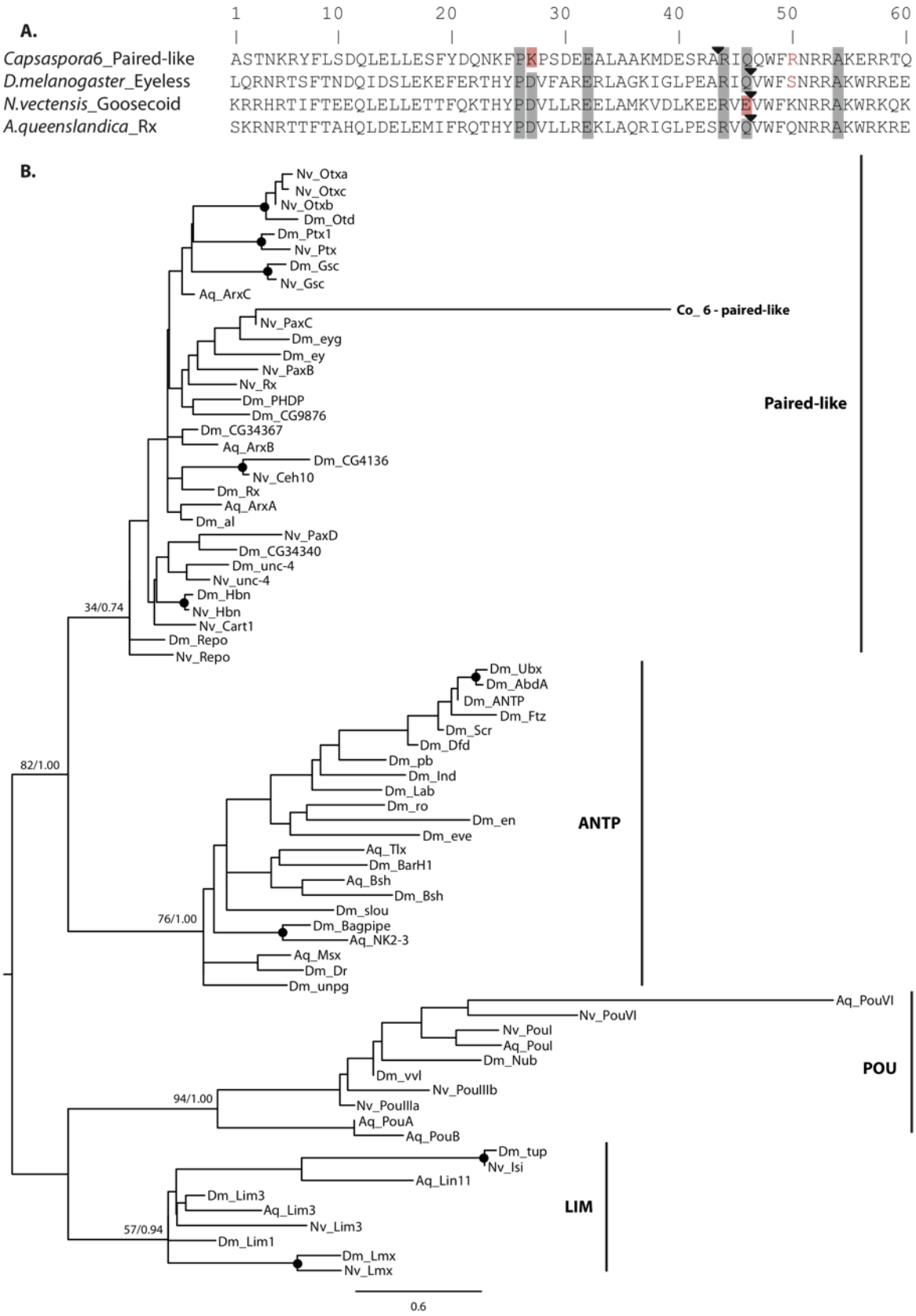

Figure S21. A) An illustrative homeobox alignment, showing the Capsaspora6 with some other archetypical Prd-like genes from different species. In grey the aminoacids that define Prd-like class (Galliot et al. 1999). Black triangle shows the intron position. **B)** Maximum likelihood tree from the homeobox domain of non-TALE HDs from ANTP, Prd-Like, POU and LIM-HD classes. The tree is rooted using the midpoint-rooted tree option. Statistical support was obtained by RAxML with 100-bootstrap replicates (BV) and Bayesian Posterior Probabilities (BPP). Both values are shown on key branches. A black dot indicates BV > 90% and BPP > 0.95. Taxa used Nv (*Nematostella vectensis*), Aq (*Amphimedon queenslandica*), Co (*Capsaspora owczarzaki*) and Dm (*Drosophila melanogaster*).

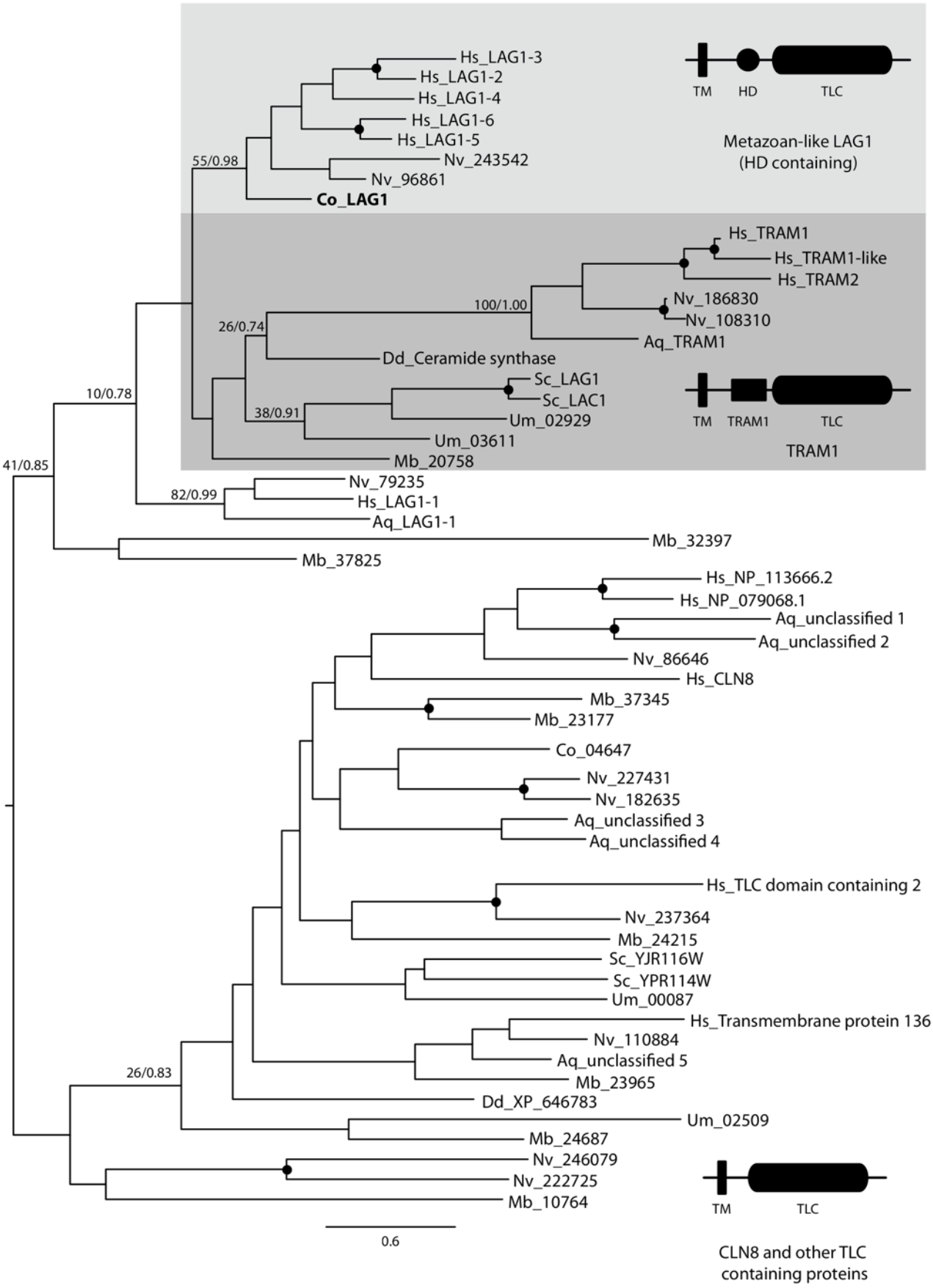

Figure S22. Maximum likelihood tree from the TLC domain. The tree is rooted using the midpoint-rooted tree option. Statistical support was obtained by RAxML with 100-bootstrap replicates (BV) and Bayesian Posterior Probabilities (BPP). Both values are shown on key branches. A black dot indicates BV > 90% and BPP > 0.95. Taxa used Nv (*Nematostella vectensis*), Aq (*Amphimedon queenslandica*), Mb (*Monosiga brevicollis*), Co (*Capsaspora owczarzaki*), Hs (*Homo sapiens*), Sc (*Saccharomyces cerevisiae*), Um (*Ustilago maydis*) and Dd (*Dictyostelium discoideum*). PFAM domain architecture displayed in the different classes.

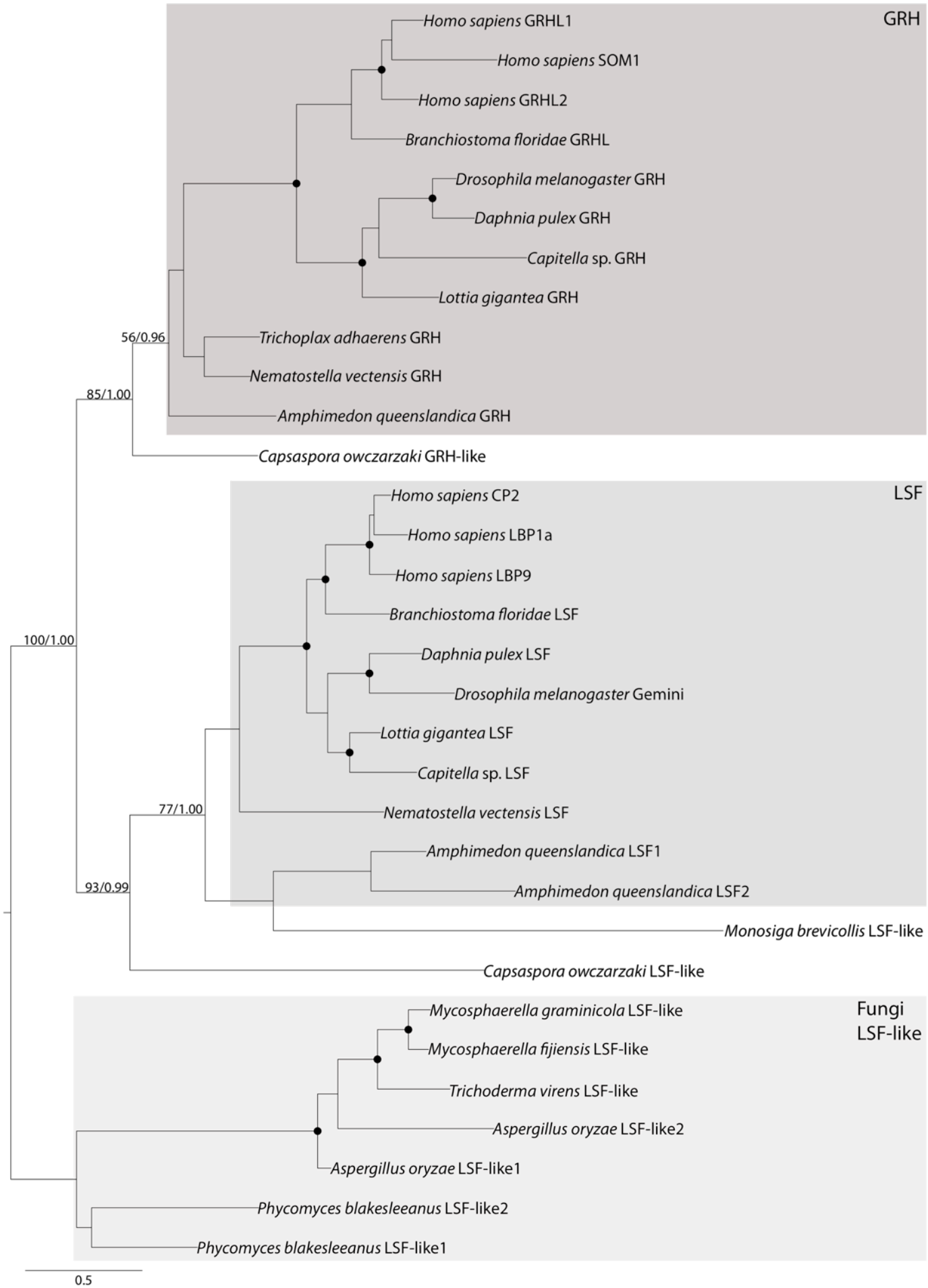

Figure S23. Maximum likelihood tree of CP2 domain containing genes. The tree is rooted using fungi. Statistical support was obtained by RAxML with 100-bootstrap replicates (BV) and Bayesian Posterior Probabilities (BPP). Both values are shown on key branches. A black dot indicates BV > 90% and BPP > 0.95.