



**Figure S2.** Phylogenetic tree of RBCS protein sequences. The maximum-likelihood tree was inferred with RAxML using the LG+ $\Gamma$  substitution model. The bootstrap support values and posterior probabilities (from PhyloBayes) are indicated at branches when higher than 50% and 0.95, respectively. Clades with organisms possessing secondary plastids are boxed. *E. longa* is in bold.