



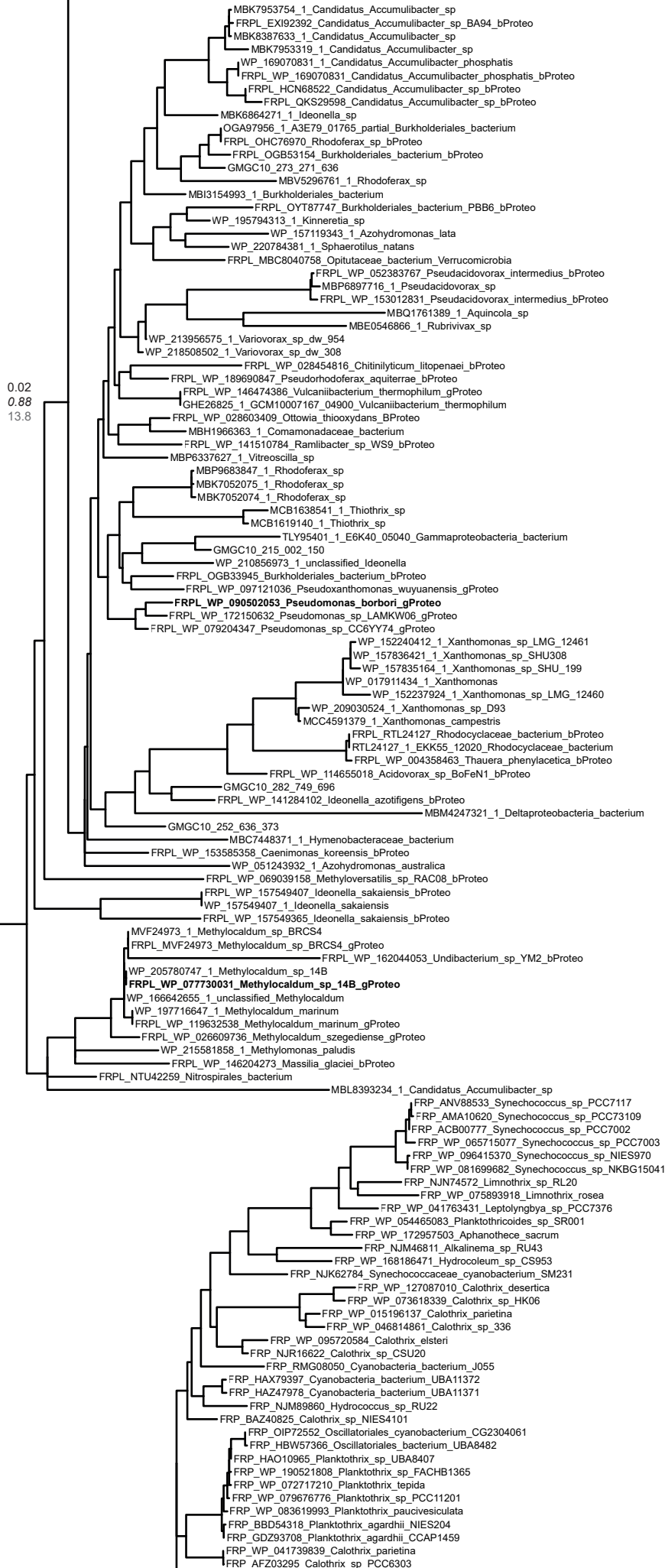
# Fortuitously compatible protein surfaces primed allosteric control in cyanobacterial photoprotection

---

In the format provided by the authors and unedited

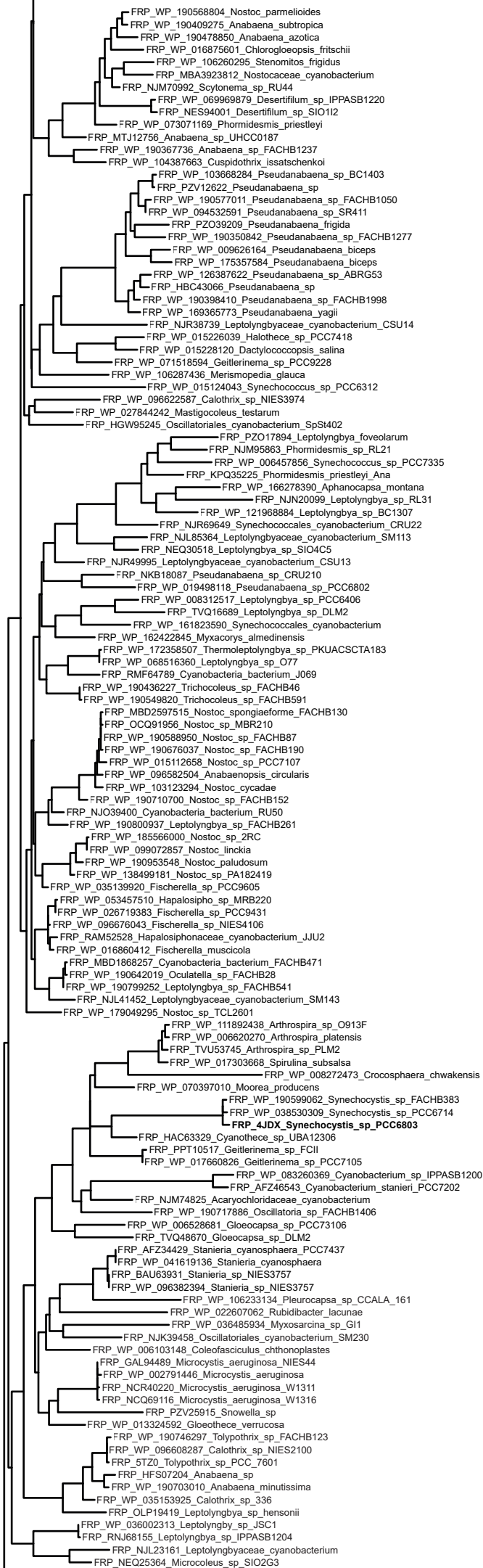


alphaproteobacteria



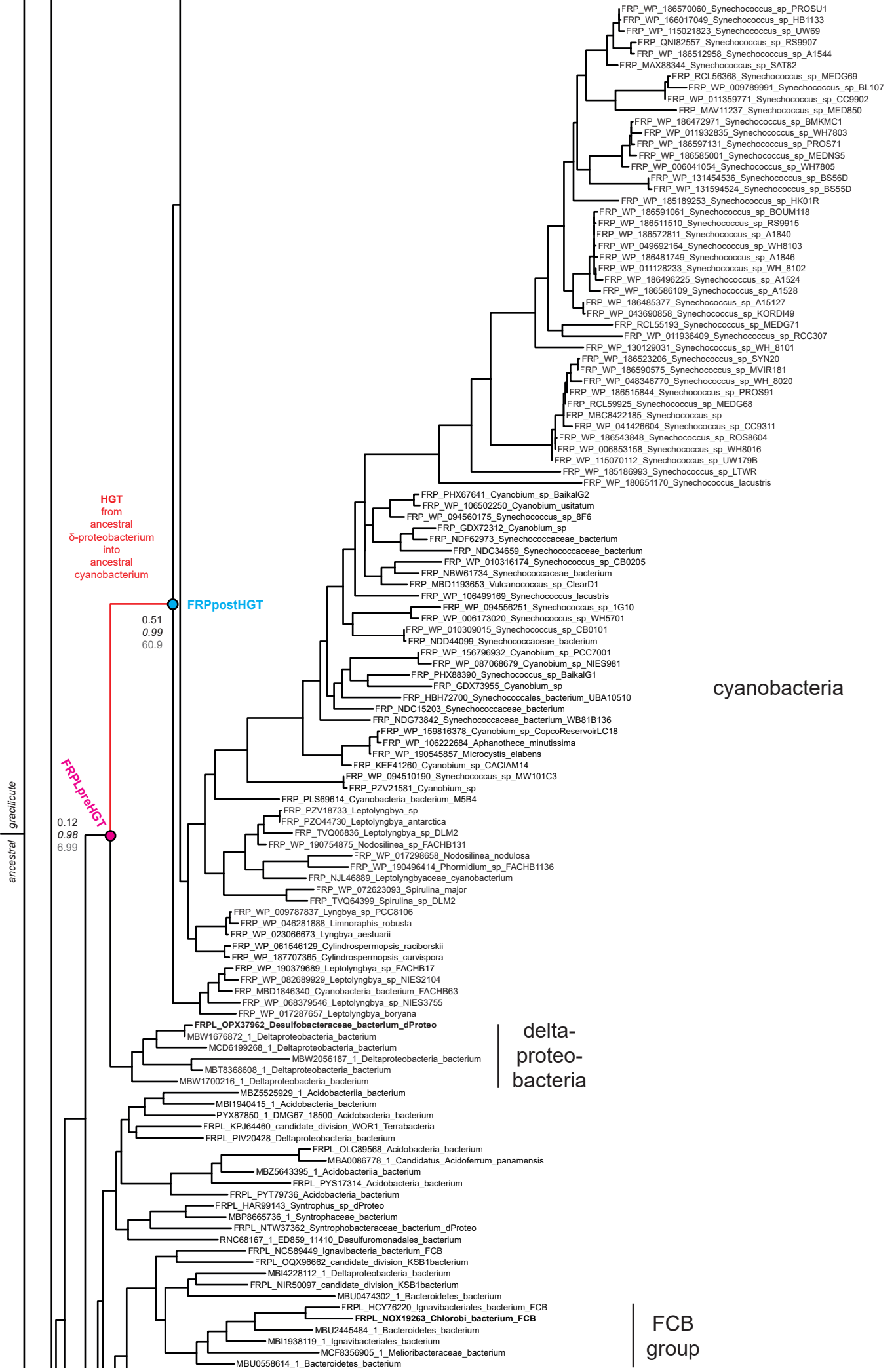
betaproteobacteria

cyanobacteria

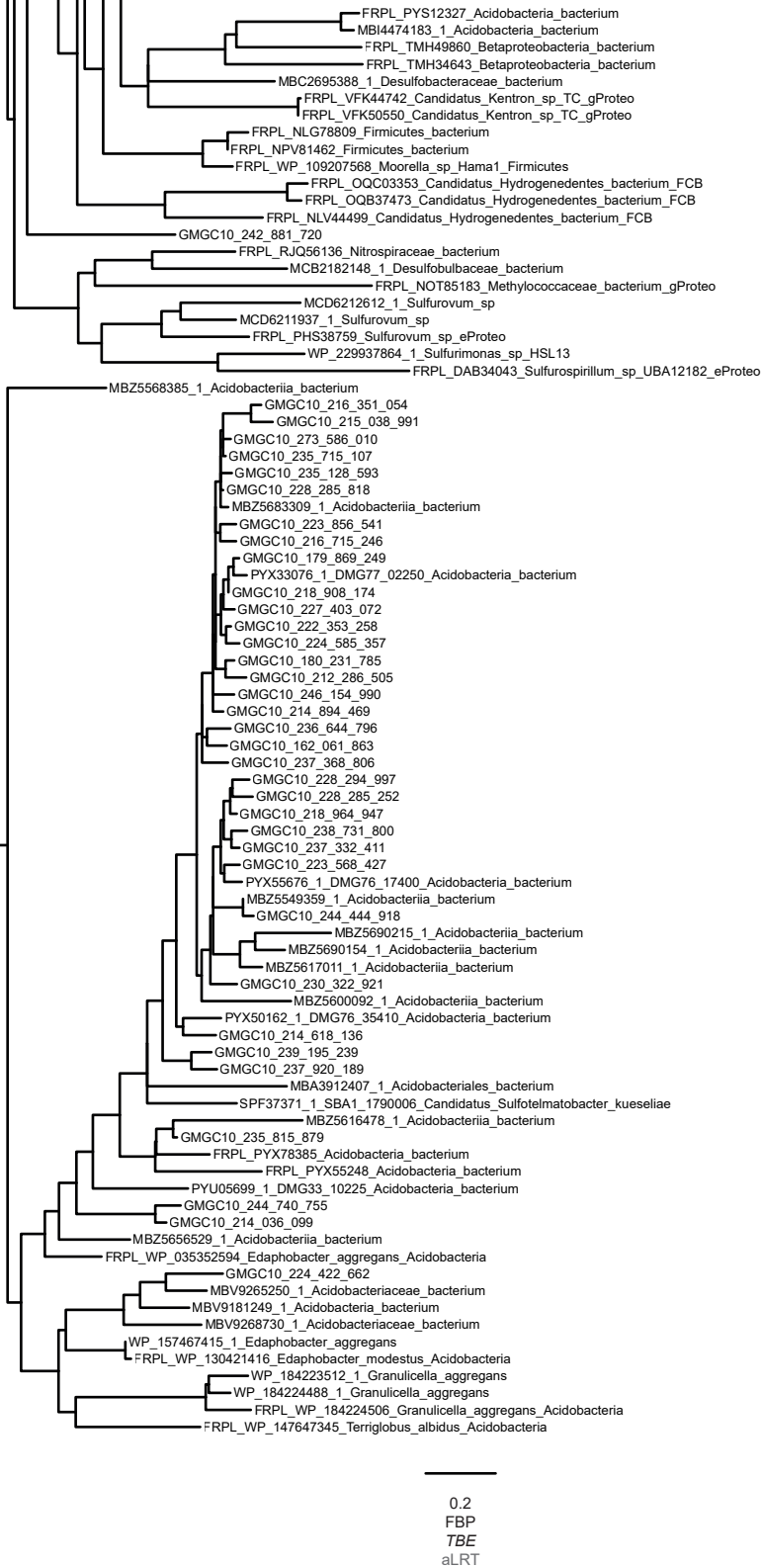


cyanobacteria

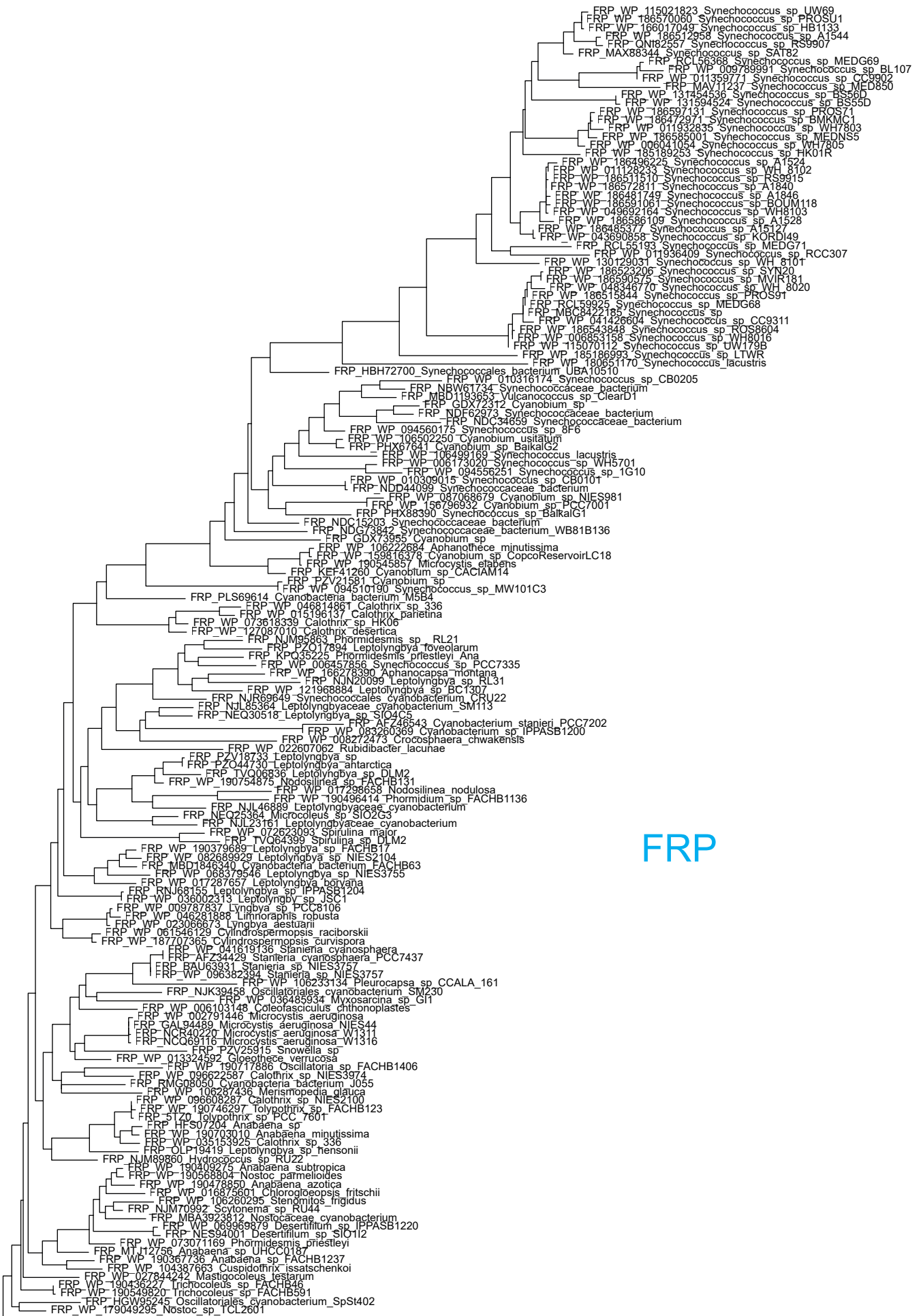
0.60  
0.97  
48.4



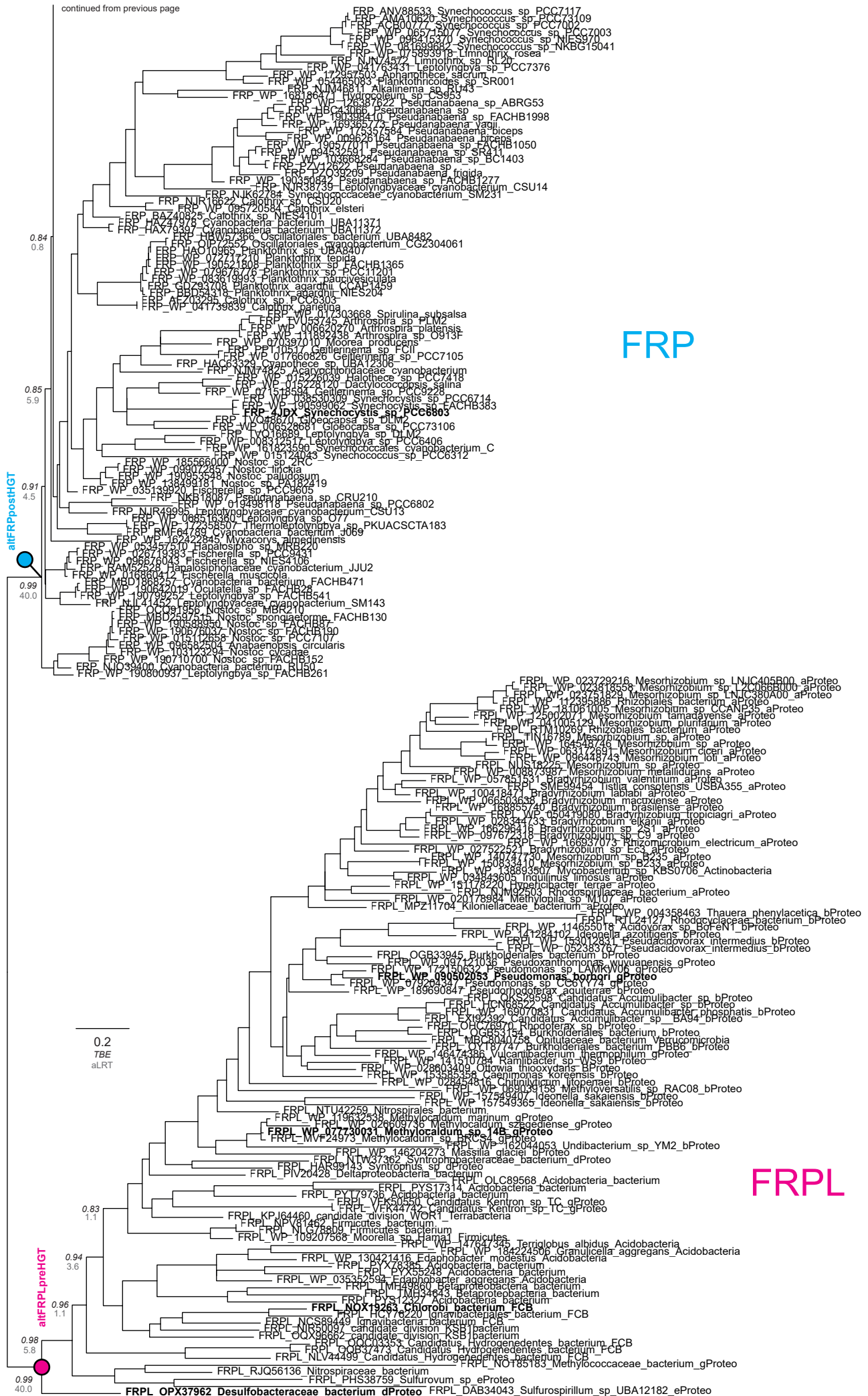
continued from previous page



**Supplementary Fig. 1 | Complete phylogeny of FRP and FRPL proteins.** Uncollapsed ML phylogeny of data shown in Fig. 2a with reconstructed ancestral proteins that existed before (FRPLpreHGT) and after the horizontal gene transfer event (FRPpostHGT) at labeled nodes. Felsenstein Bootstrap Probabilities (FBP) and Transfer Bootstrap Expectations (*TBE*) of 100 replicates are given at important nodes. Grey numbers are approximated likelihood ratio test values (aLRT). Branch-lengths represent average substitutions per site. Examined proteins in this study are highlighted in bold. Underlying multiple sequence alignment in Supplementary Data 1.



FRP

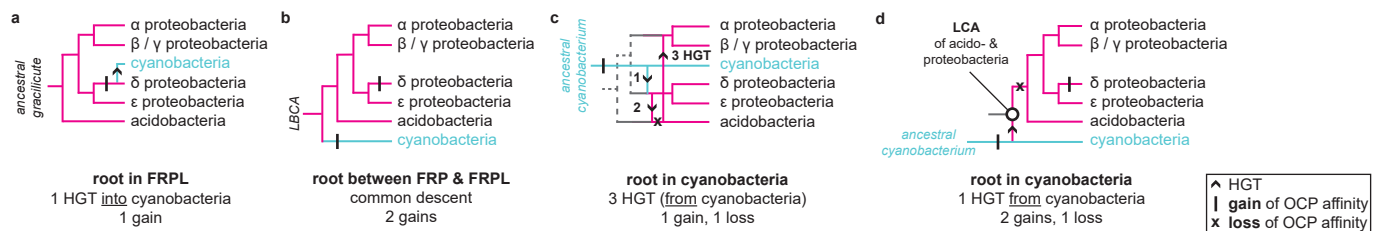


FRP

FRPL



**Supplementary Fig. 2| Complete unrooted initial alternative phylogeny of FRP and FRPL proteins.** Uncollapsed ML phylogeny of data shown in Extended Data Fig. 8e with reconstructed alternative ancestral proteins that existed before (altFRPLpreHGT) and after the horizontal gene transfer event (altFRPpostHGT) at labeled nodes. Transfer Bootstrap Expectations (*TBE*) of 100 replicates are given in italics at important nodes. Grey numbers are approximated likelihood ratio test values (aLRT). Branch-lengths represent average substitutions per site. Examined proteins in this study are highlighted in bold. Underlying multiple sequence alignment in Supplementary Data 1.



**Supplementary Fig. 3| Scenarios of FRP and FRPL evolution.** **a**, Simplified scheme of the evolution of FRP and FRPL proteins, according to the most parsimonious scenario shown in Fig. 2a. **b-d**, Simplified schemes of the alternative, but less parsimonious scenarios. *LBCA*, last bacterial common ancestor. HGT, horizontal gene transfer.

### Supplementary Tab. 1| Utilized oligonucleotides in this study

oligonucleotide	sequence (5' to 3')
primer for linearization of pLIC vectors, forward	CTCGAGCACCACCACCAC
primer for linearization of pLIC vectors, reverse	ATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGG
primer for amplification of genes to clone into pLIC vectors, forward	ATAATTTTGTTTAACTTTAAGAAGGAGATATACAT
primer for amplification of genes to clone into pLIC vectors, reverse	GCCGGATCTCAGTGGTGG

## **Supplementary Discussion**

### **The fortuitous ability of FRPL to act on OCP is agnostic about an HGT event**

We postulate that cyanobacteria acquired FRP via a horizontal gene transfer (HGT) from delta-proteobacteria early in their history after OCP first formed in an ancestral cyanobacterium (Fig. 2a). This is the most parsimonious explanation given the available sequence data our phylogenetic tree is based on, but also implies a large number of gene losses in different kinds of non-cyanobacterial groups (Supplementary Fig. 3a).

A root between FRPs and FRPLs would place the evolution of FRPLs near the last common ancestor (LCA) of all bacteria<sup>23</sup>, which would imply even more gene losses in virtually all major bacterial groups (Supplementary Fig. 3b). However, even in this scenario, the intrinsic ability of FRPLs to act on OCP would have evolved in a non-cyanobacterium by chance without direct selective pressure. In fact, FRPLs would have randomly drifted in and out of the sequence space that enables the interaction with OCP during its evolution, and happened to be capable of the interaction when OCP first evolved in cyanobacteria.

We have also considered the possibility that the root may lie within the FRP clade, implying HGT to proteobacteria and acidobacteria. However, we consider this scenario very implausible: it would require a transfer from a relatively modern cyanobacterium with FRP into at least the LCA of all proteobacteria and acidobacteria. This is not only temporally implausible, but further incongruent with the topology of our gene tree: when rooted inside FRPs, our phylogeny does not place proteobacteria and acidobacteria sister to all FRPs. This means our gene tree would require additional HGT between different kinds of proteobacteria to explain the distribution of FRPL (Supplementary Fig. 3c). In addition, this scenario implies that FRPLs in delta-proteobacteria fortuitously retained their ability to interact with OCP since around the time of the LCA of all living cyanobacteria.

Another possibility is that our gene tree is simply incorrect, perhaps owing to the short length of the FRP and FRPL proteins. For example, if the true tree in the FRPL clade actually follows the species phylogeny of acidobacteria and proteobacteria, we could root the tree inside the FRP clade and explain the tree with a single horizontal transfer (Supplementary Fig. 3d). FRP's function would then be ancestral. But this would also imply that the ability to bind OCP was lost many times independently in FRPLs, and was regained in only a small set of delta-proteobacteria. We also consider this very implausible.

Further, a reconciliation of 100 bootstrap trees with a species phylogeny using amalgamated likelihood estimation (ALE)<sup>45</sup> found no root between FRPs and FRPLs, but 72 tree topologies featuring a root within the FRPLs.

Taken together, the OCP-matching interface in FRPLs evolved without direct selective pressure mainly by chance even if we remain agnostic about the horizontal transfer event, that we think is still the most likely scenario here.