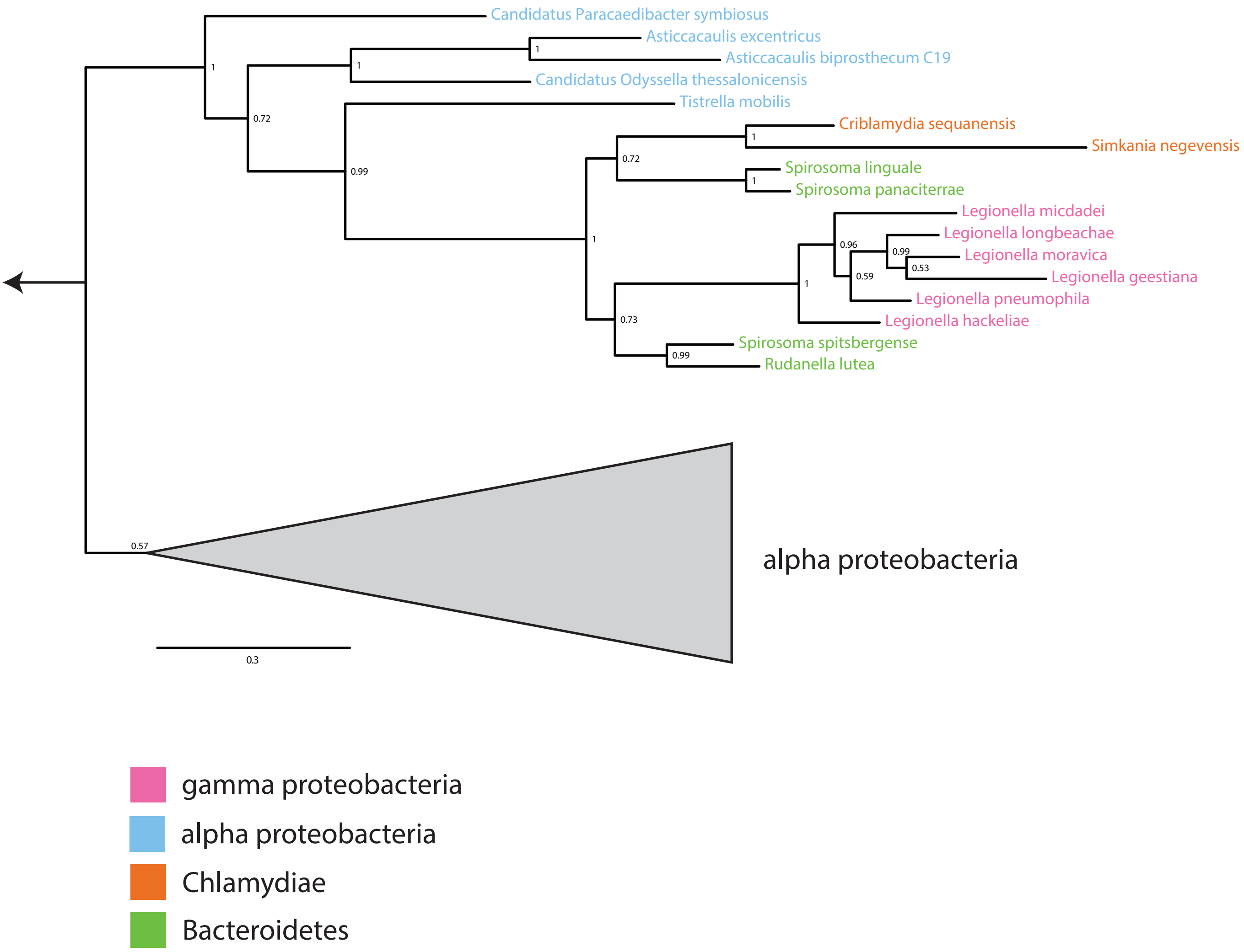
Following the footsteps of chlamydial gene regulation

Domman and Horn

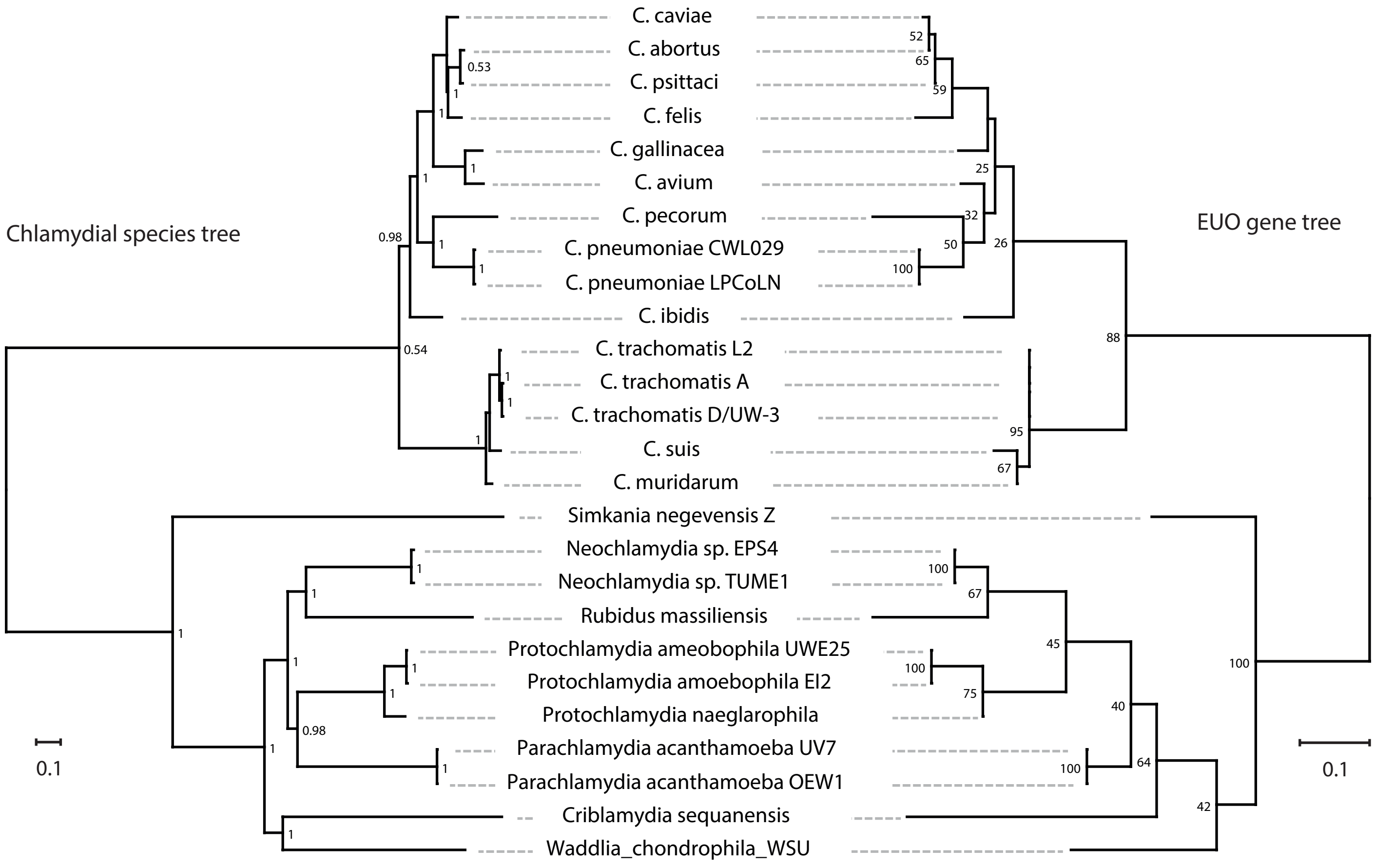
Supplementary Information

Table S2. **Comparison of operon predictions.**

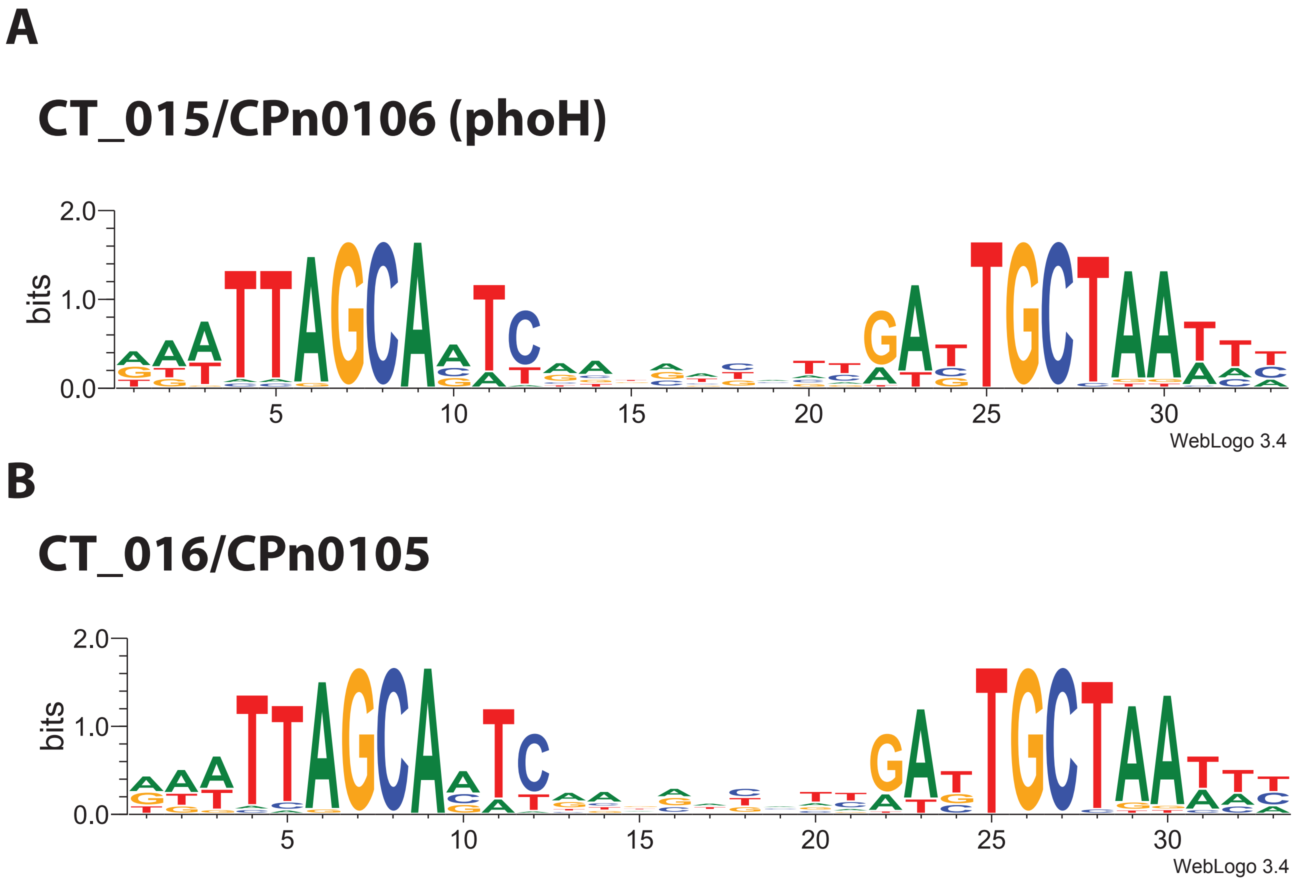
	***Chlamydia trachomatis* 434/Bu**	***Chlamydia pneumoniae* CWL029**	***Protochlamydia amoebophila* UWE25**
Total genes	874	1033	2031
Genes in RSAT operon prediction	537	634	947
Genes in DOOR operon prediction	575	691	995
RSAT predictions in DOOR predictions	531	624	917
Genes in non-operon network	311	385	384
Genes in operon network	644	733	710



Supplementary Figure S1. **Acquisition of transcription factors for members of the *Chlamydiae*.** The transcription factor *pecT* is part of the general LysR transcriptional regulator family, and seems to have been aquired by *Simkania negevensis* (F8L9G7)and *Criblamydia sequanensis* (A0A090E2E0) from members of the *Baceteroidetes*. The gene history is complicated, as the gene was transfered first from alpha-proteobacteria into the clade with the *Chlamydiae* and *Bacteroidetes*. There was a subsequent transfer of this gene to members of the *Legionella*, which are known to infect free living ameoba.



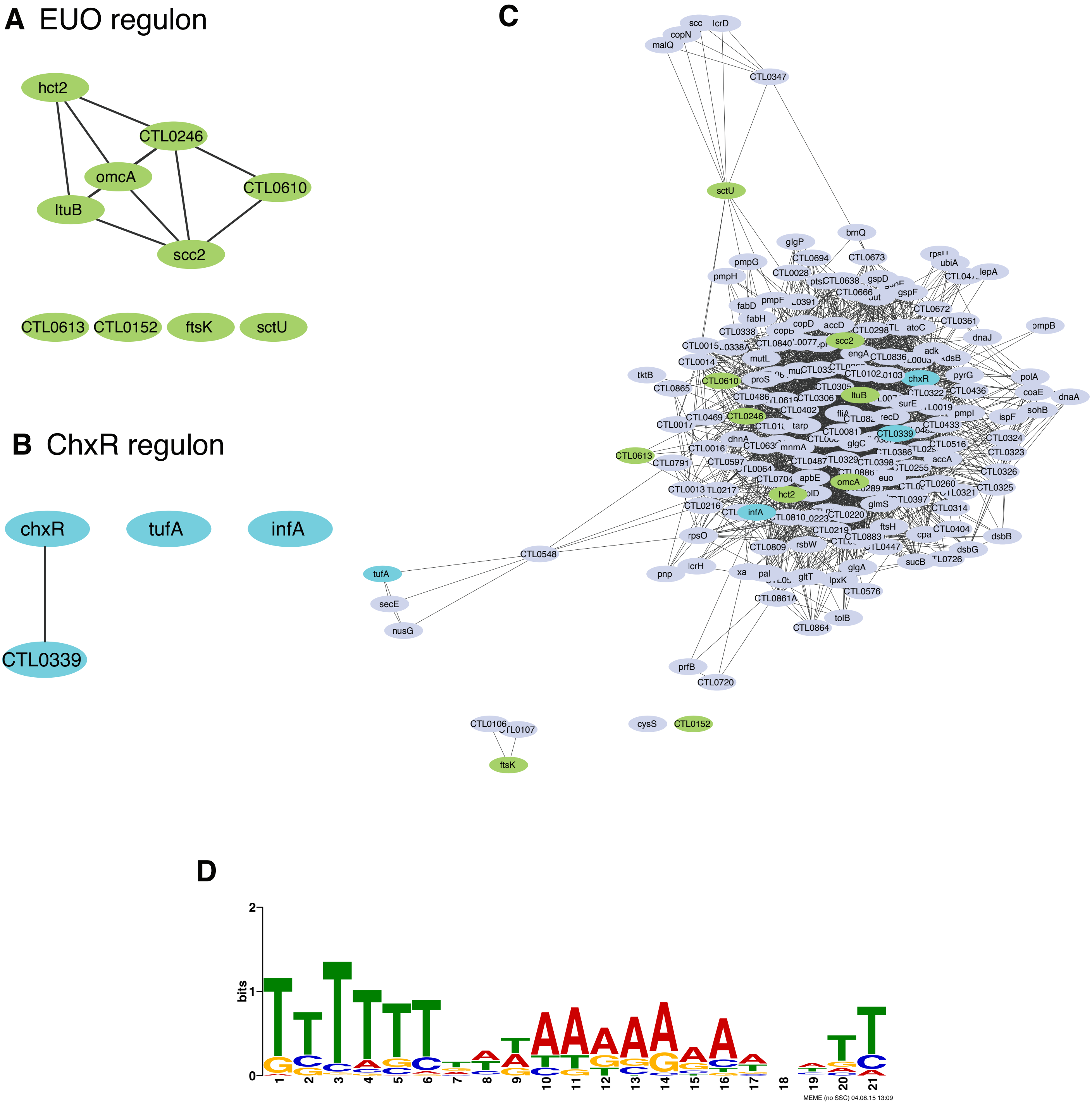
Supplementary Figure 2. **Gene trees of global regulators, such as EUO, largely follow the species tree.** Shown is a comparison of the species phylogeny (based on 33 concatenated marker proteins) to the gene tree of the late gene regulator EUO. Congruence between species and gene trees indicates vertical inheritance of this global regulator throughout chlamydial evolution. The EUO gene tree was calculated via RAxML under the LG +gamma+F model with 1000 bootstraps.



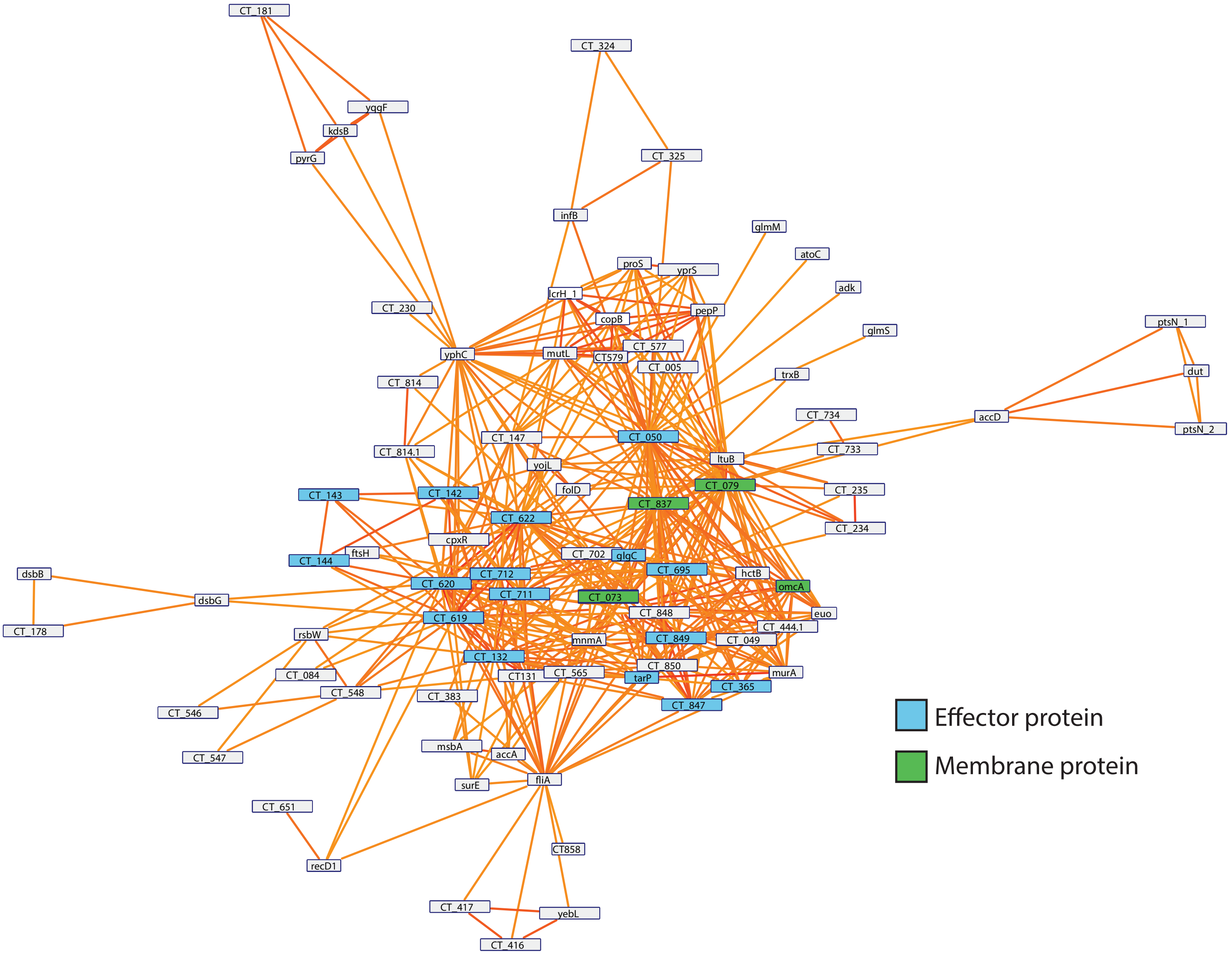
Consensus CIRCE

TTAGCACTC-N$_9$-GAGTGCTAA

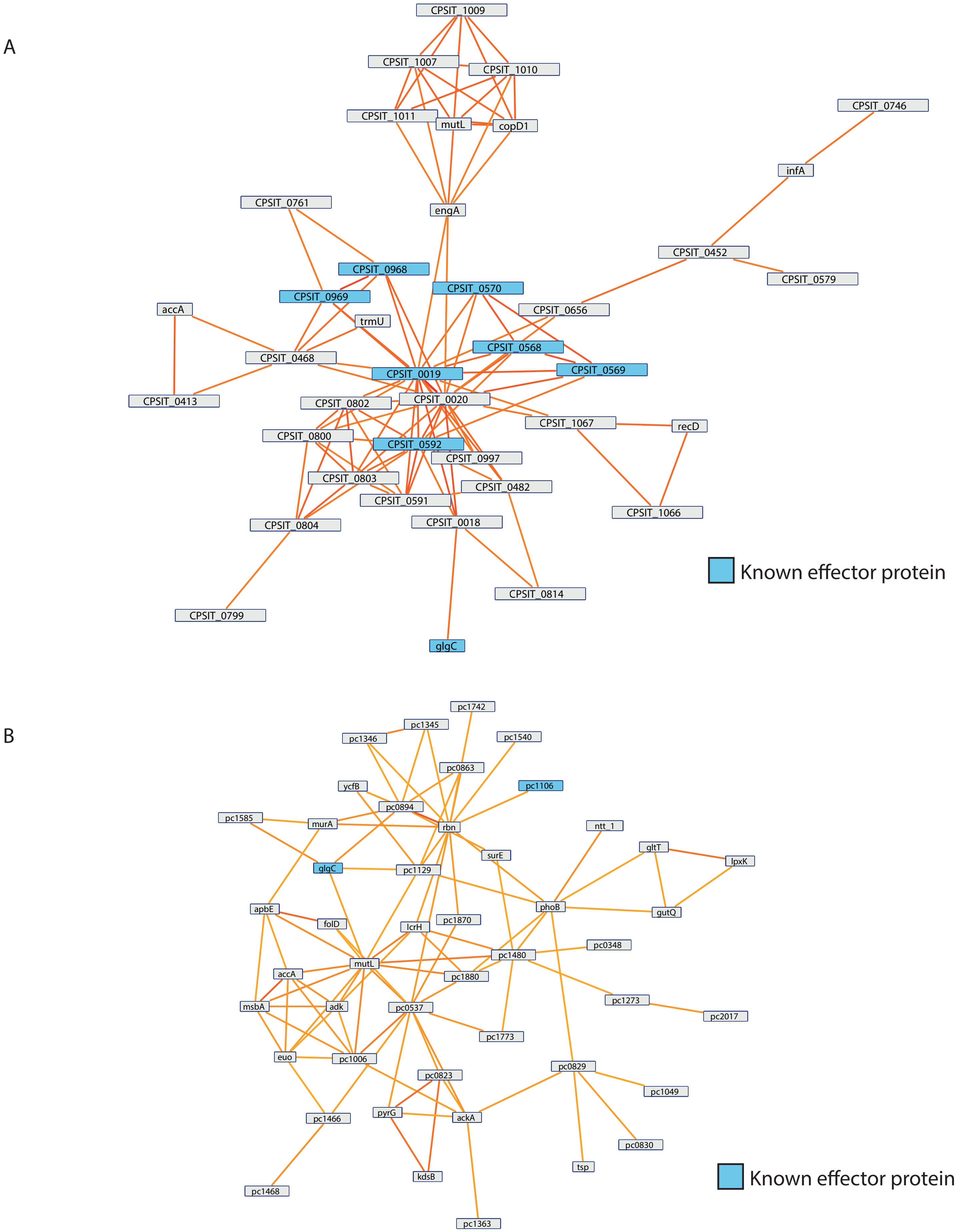
Supplementary Figure S3. **Conserved motifs for predicted novel HrcA regulon members.** The discovered motifs for both *phoH* (A) and CPn0105 (B) are near perfect matchers to the consensus CIRCE element reported in Wilson and Tan (2004), strongly suggesting that these genes are under the control of HrcA.



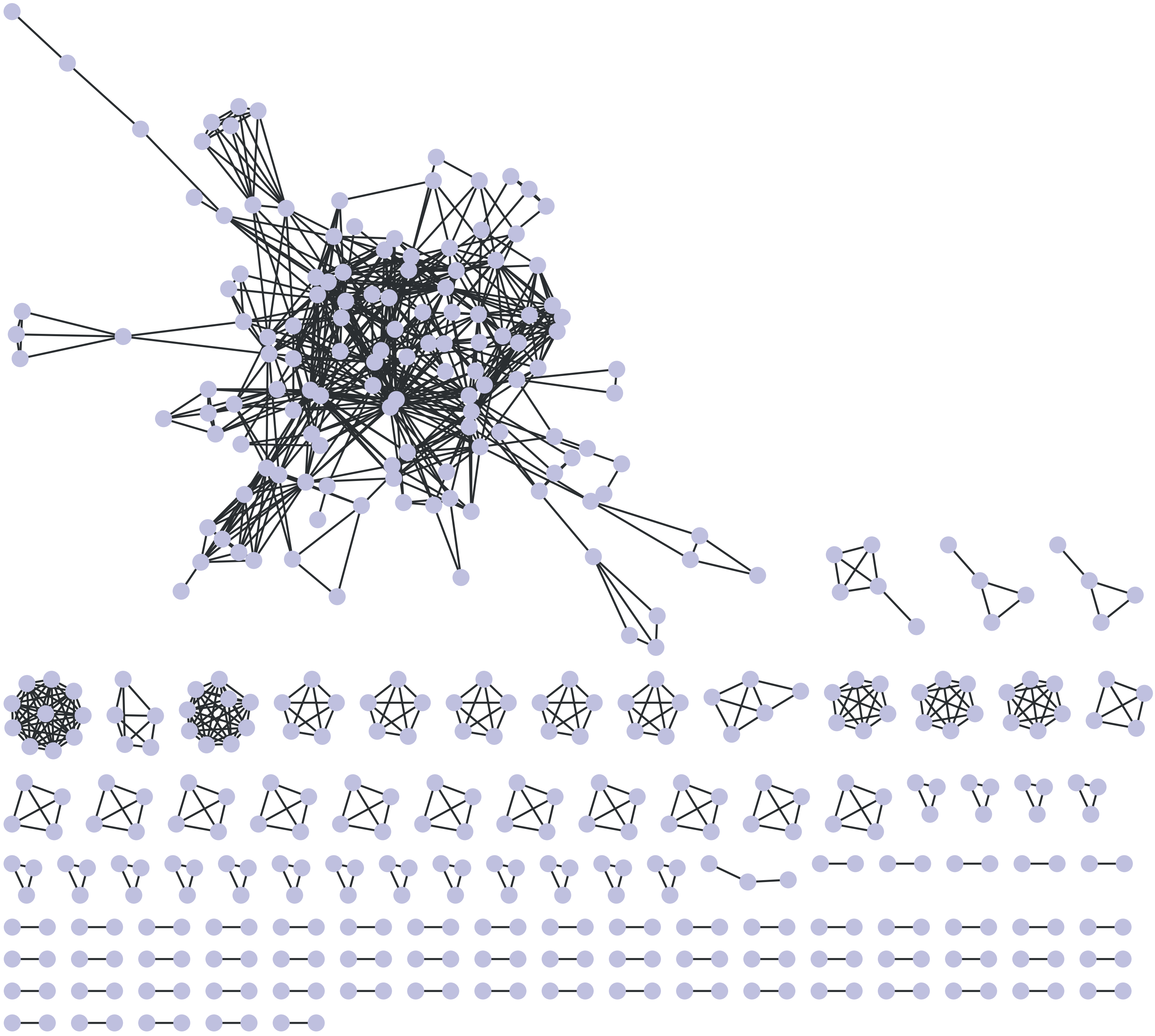
Supplementary Figure S4. **EUO and ChxR regulons found in predicted networks.** The co-regulatory network predictions from *Chlamydia trachomatis* 434/Bu of the known members of the (A) EUO regulon and (B) ChxR regulon. The known EUO regulon currently consists of 15 members, while the known ChxR regulon is comprised of five. EUO members are colored in green and ChxR are colored in blue. Six members of the EUO network have direct connections with each other in the network. Only chxR and CTL0339 (CT_084) are linked among the ChxR regulon. (C) Extracting the nearest neighbors from the complete network for each known member of both the EUO and ChxR regulons, one observes a tightly connected subnetwork of 178 nodes. This tight clustering is strongly indicative of these genes being part of the same regulon. It may be that some members are duely regulated, and thus the EUO and ChxR networks are linked together. (D) Logo plot of the motif detected using MEME in the promoter regions of genes in (C).



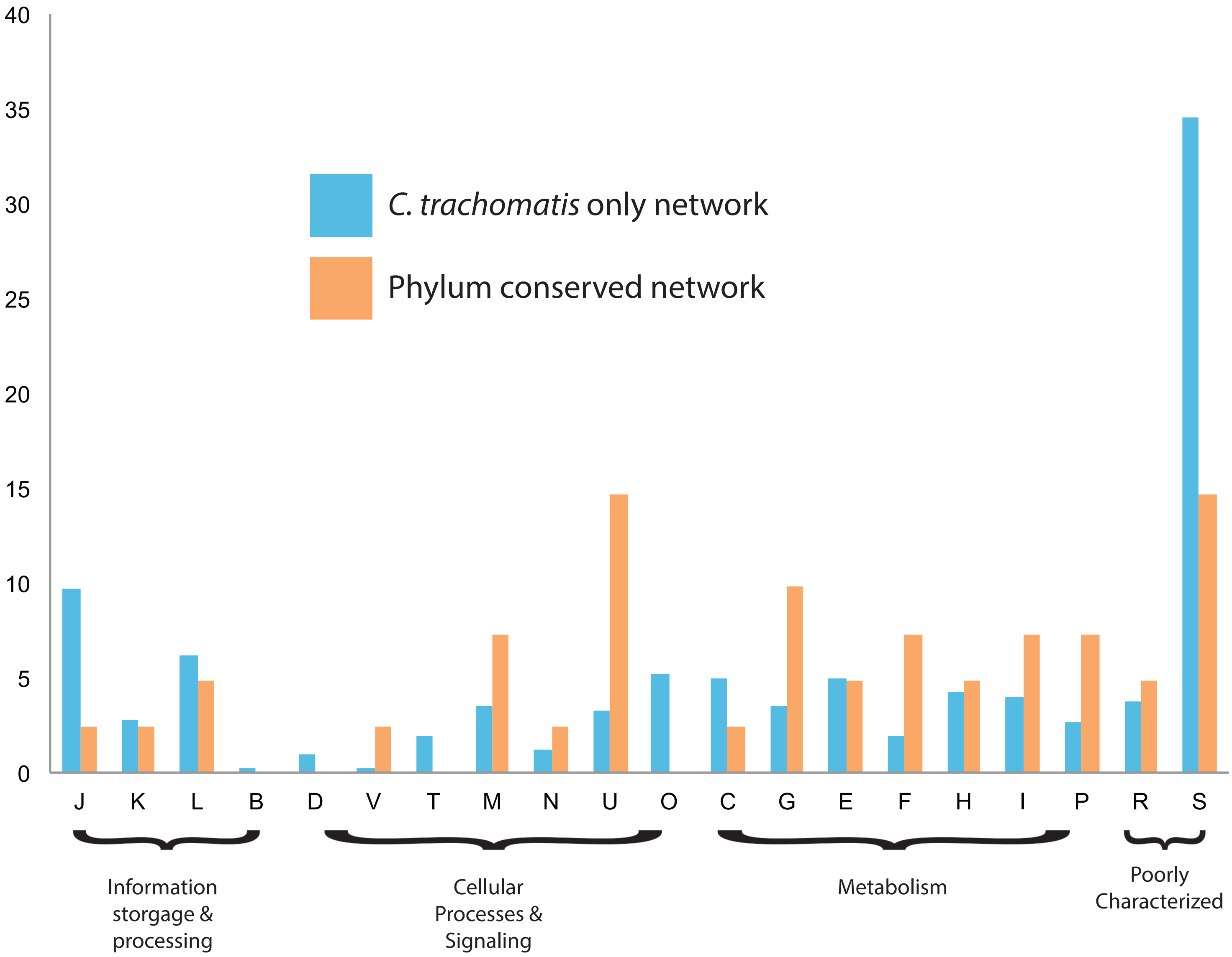
Supplementary Figure 5. **Predicted virulence co-regulation sub-network of *C. trachomatis.*** The original network was filtered for edges with a DPbits score >= 5 and the putative virulence regulon extraced. This network is comprised of 89 nodes, including many known Type III secretion effector proteins and other genes involved in virulence. The strength of the predictions are given by the weight and color of the edges, where the more red and thicker equals better prediction.



Supplementary Figure S6. **Predicted virulence co-regulatory networks for (A) *Chlamydia psittaci* 6BC and (B) *Protochlamydia amoebophila UWE25.*** The network for C. psittaci was filtered at a DPbits score >=6, while the the Protochlamdyia networks was filtered at DPbits score >=4. Orthologs to known effector proteins are colored in blue. Many of the effector proteins found in *C. trachomatis* are still highly connected in the *C. psittaci* network, indicating genus level conservation of virulence genes. Protochlamydia still share some known effectors, but many of the genes linked here are of unknown function, possibly revelaing a different set of virulence factors for this organism.



Supplementary Figure S7. **Consensus co-regulatory network from members of the *Chlamydiaceae*.** This network was created by only incorporating edges that are present in all six *Chlamydiaceae* genomes analyzed. The network contains 443 nodes and 975 edges.



Supplementary Figure S8. **Comparison of eggNOG functional categories for the *C. trachomatis* and phylum conserved sub-networks.** The phylum conserved network is enriched in eggNOG functional categories that are involved in metabolism and cellular proccesses and signaling versus that of the *C. trachomatis* network. Note the dramatic reduction in the poorly characterized proteins, as this largely represents the Type III effector proteins used for host cell entry, rather than those used by chlamydia to exploit the intracellular niche. eggNOG functional categries are descbied here: ftp://eggnog.embl.de/version_4.0.beta/data/downloads/eggnogv4.funccats.txt.