

Reconstructing Ancestral Gene Blocks in Bacteria

Huy Nguyen¹ and Iddo Friedberg²

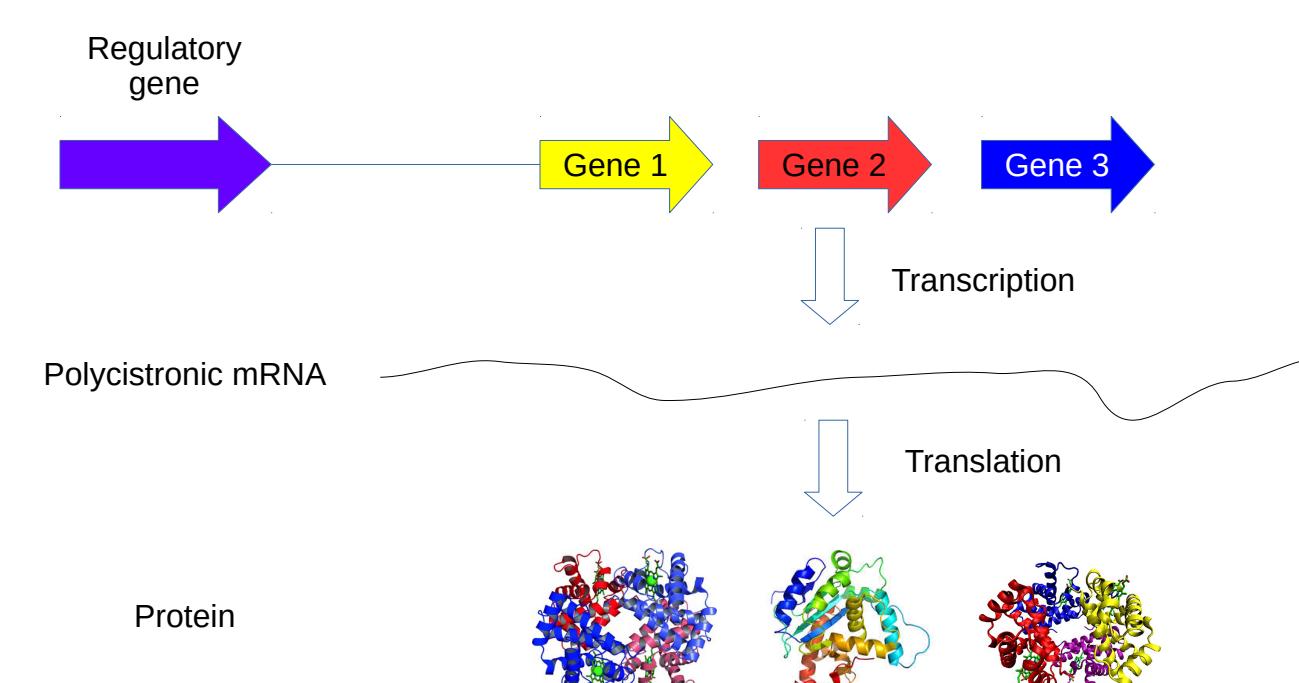
¹Program in Computer Science and ²Department of Veterinary Microbiology and Preventive Medicine
Iowa State University, Ames, IA, USA

How do orthologous gene blocks evolve in bacteria?

Gene blocks are genes co-located and co-regulated usually sharing the same pathway or protein complex



The operon is a special case of gene block where co-transcription occurs on a polycistronic mRNA under common regulation



Describe the evolution of gene blocks using event-driven method

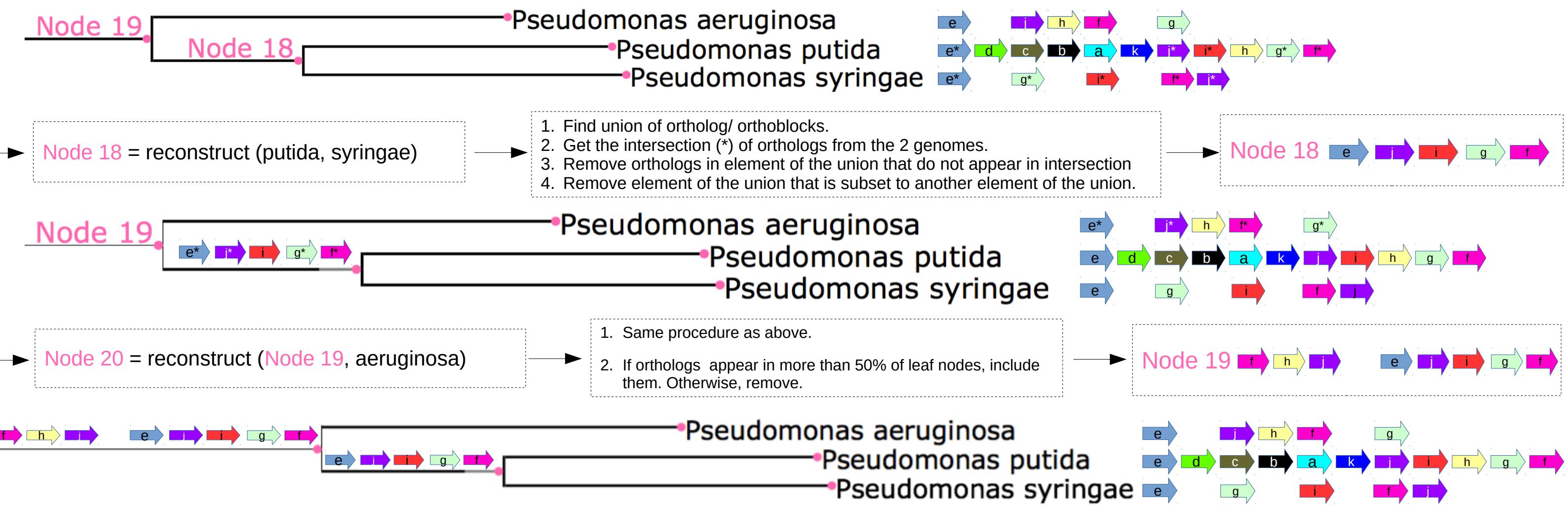
We define 3 events:

- Split:
- Deletion:
- Duplication:

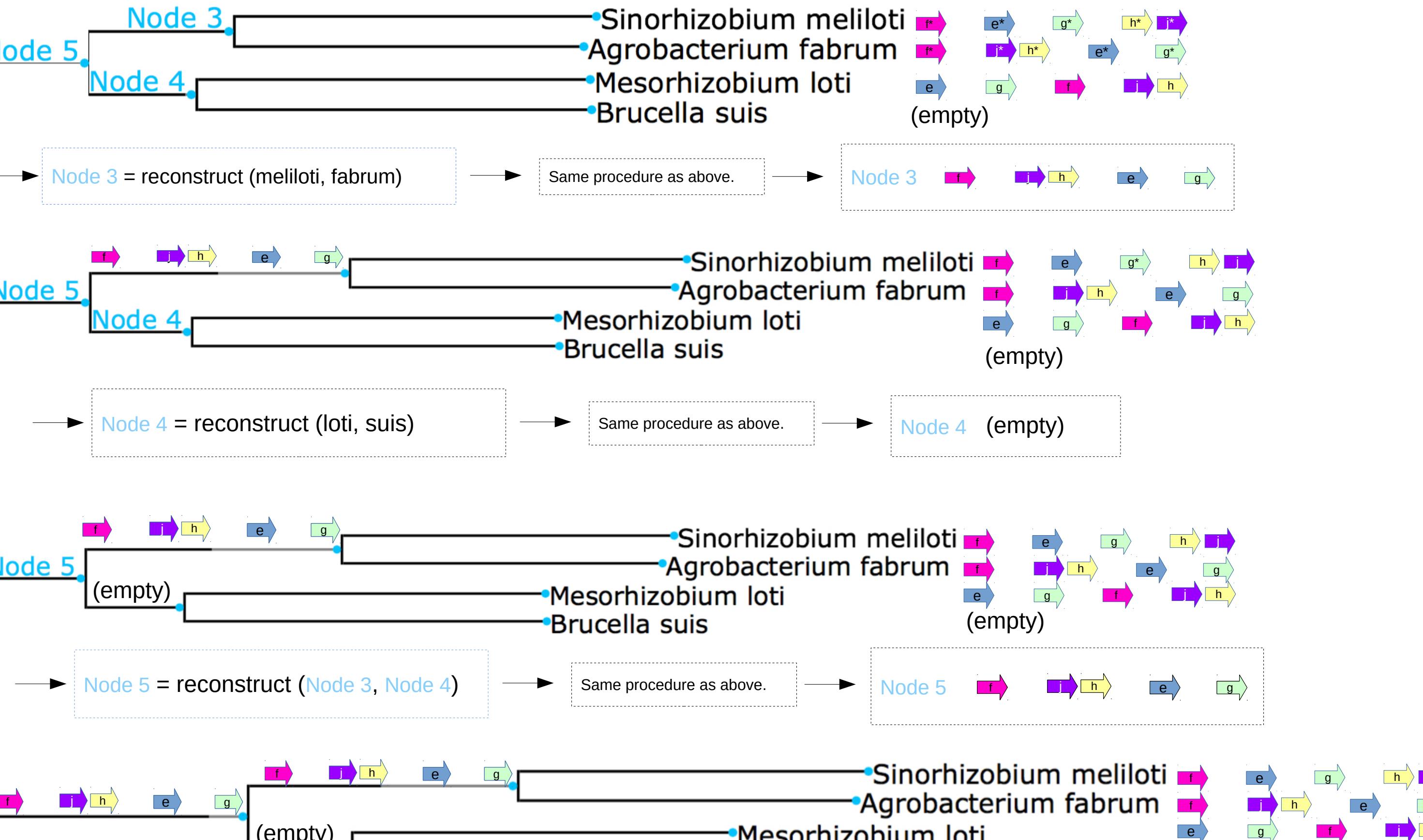
The method describes evolutionary events that occur between gene blocks that are homologous between different bacterial species.

Reconstruct ancestral gene blocks using maximum parsimony

Example 1: A sub tree for operon paaABCDEFGHIJK



Example 2: Another sub tree for operon paaABCDEFGHIJK



Conclusions

- We have developed a method to reconstruct ancestral gene blocks using a simple distance measure and maximum parsimony. The scoring for the distance measure is generated based on the corpus of extant orthoblocks in the studied species.
- We surveyed several proteobacteria for conservation of orthoblocks. Here we study two examples for conservation and evolution.
- The rplKJL-rpoB operon codes for proteins K, A, J and L of the large ribosomal subunit, as well as proteins B (beta subunit) and C (beta' subunit) of RNA polymerase. As such it is perfectly conserved in proteobacteria up to the last common ancestor.
- In contrast, the paaABCDEFGHIJK operon is only found in two species in full, and is not well conserved. The paa operon genes degrade phenylacetic acid in *E. coli*. This is a highly specific operon, and in our set the full complement of genes are only found in two orthoblocks. It is therefore interesting to see how this operon may have evolved by vertical evolution, or (as we suspect) may have been the result of more than one horizontal gene transfer.

Examine the relative conservation of orthologous gene blocks: orthoblocks

Figure 1. rplKJL-rpoBC: gene block diagram, deletions, duplications, splits

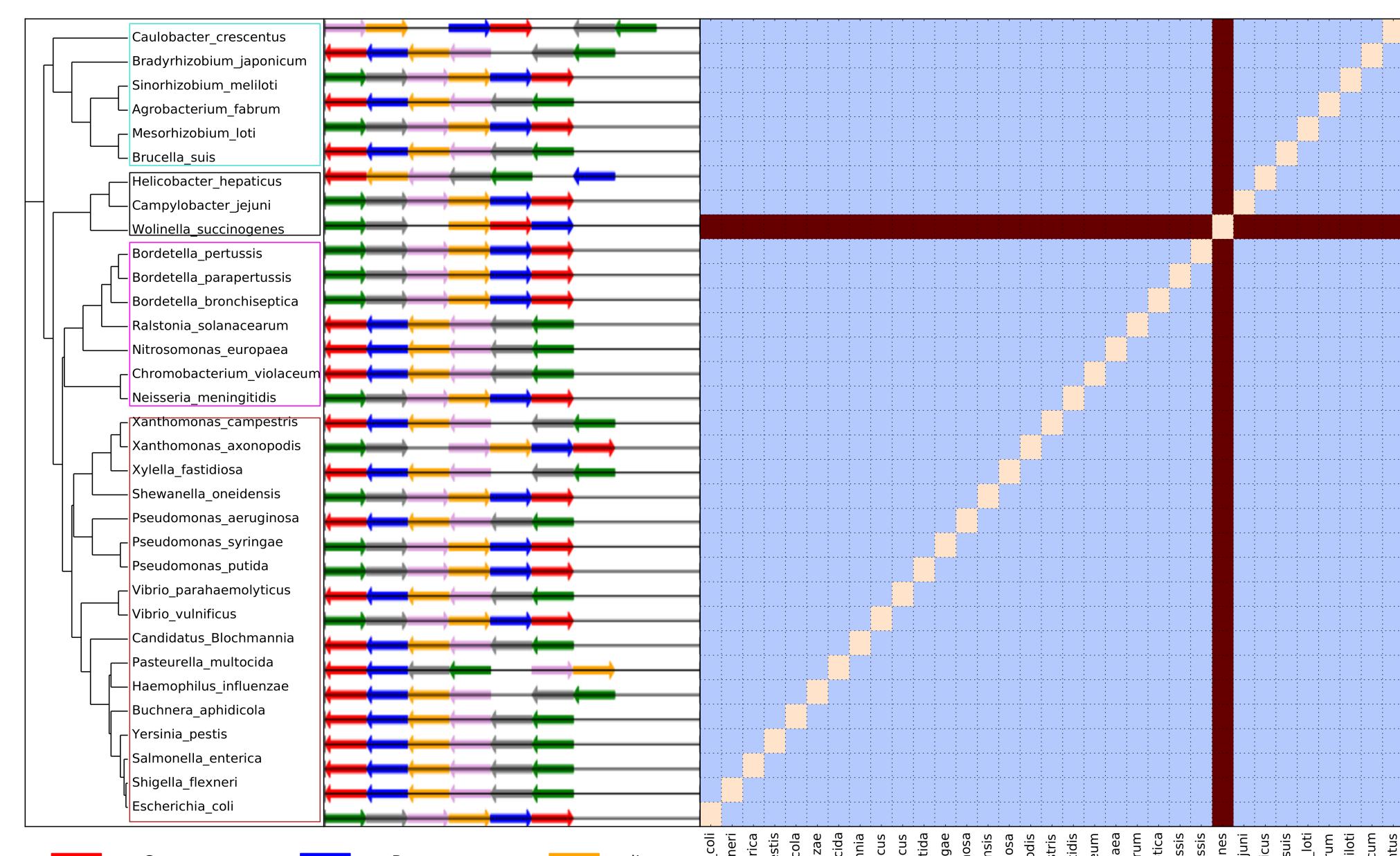


Figure 2. paaABCDEFGHIJK: gene block diagram, deletions, duplications, splits

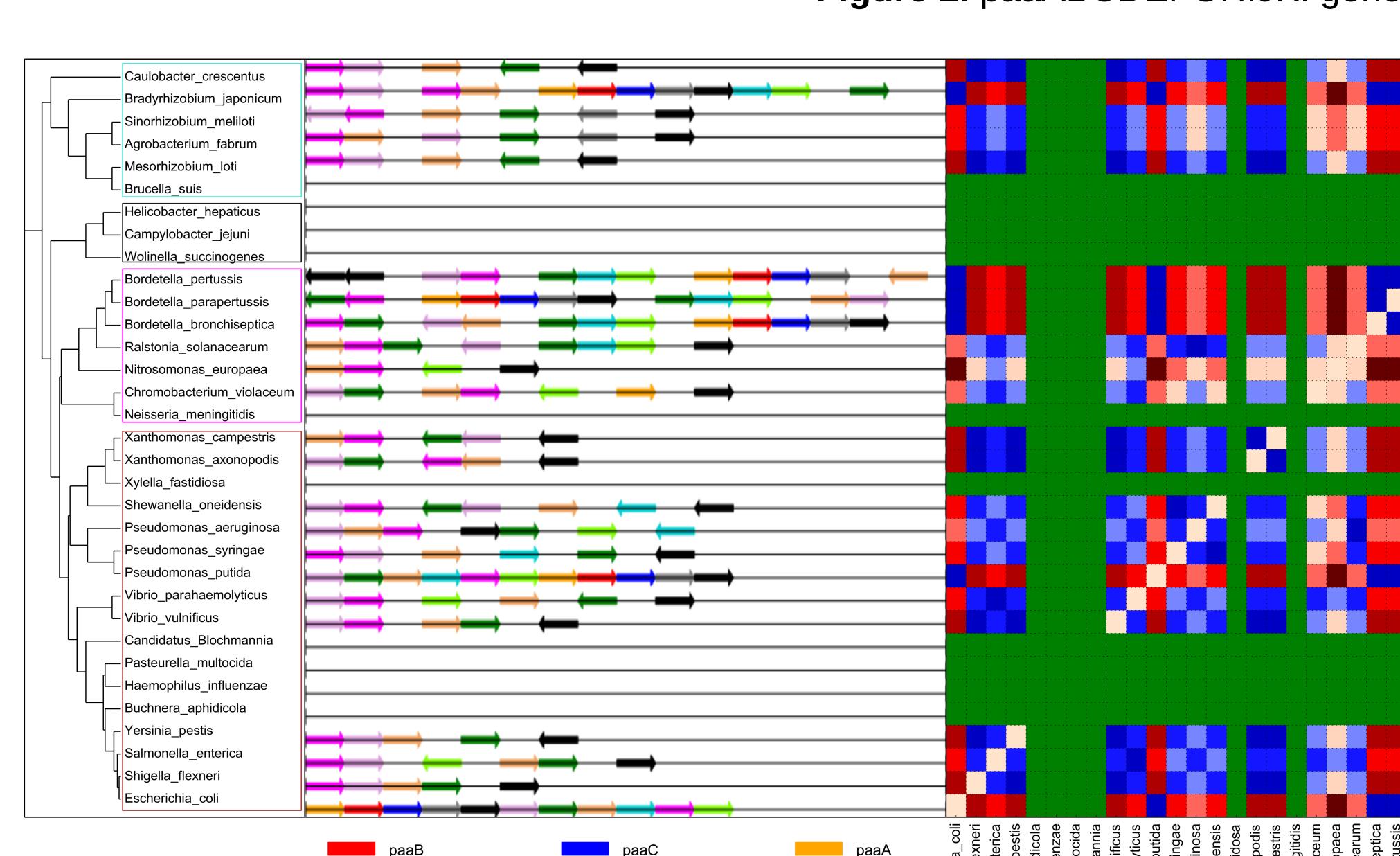


Figure 1:

- No gene duplications or deletions were detected. Erroneously call a deletion of rplJ in *Wolinella*.
- Splits are mostly between rplKA and rplJL-rpoBC.
- Other studies suggest some evidence for the strong conservation of rplJL-rpoBC.

Figure 2:

- Genes paaA, B, C, D always co-occur in an operon when they are found.
- Only full gene block in *E. coli* and *P. putida*
- Genes paaF, G co-occur in 12 out of 23 species.

Results

Figure 3. Reconstruction of rplKJL-rpoBC (highly conserved orthoblocks)

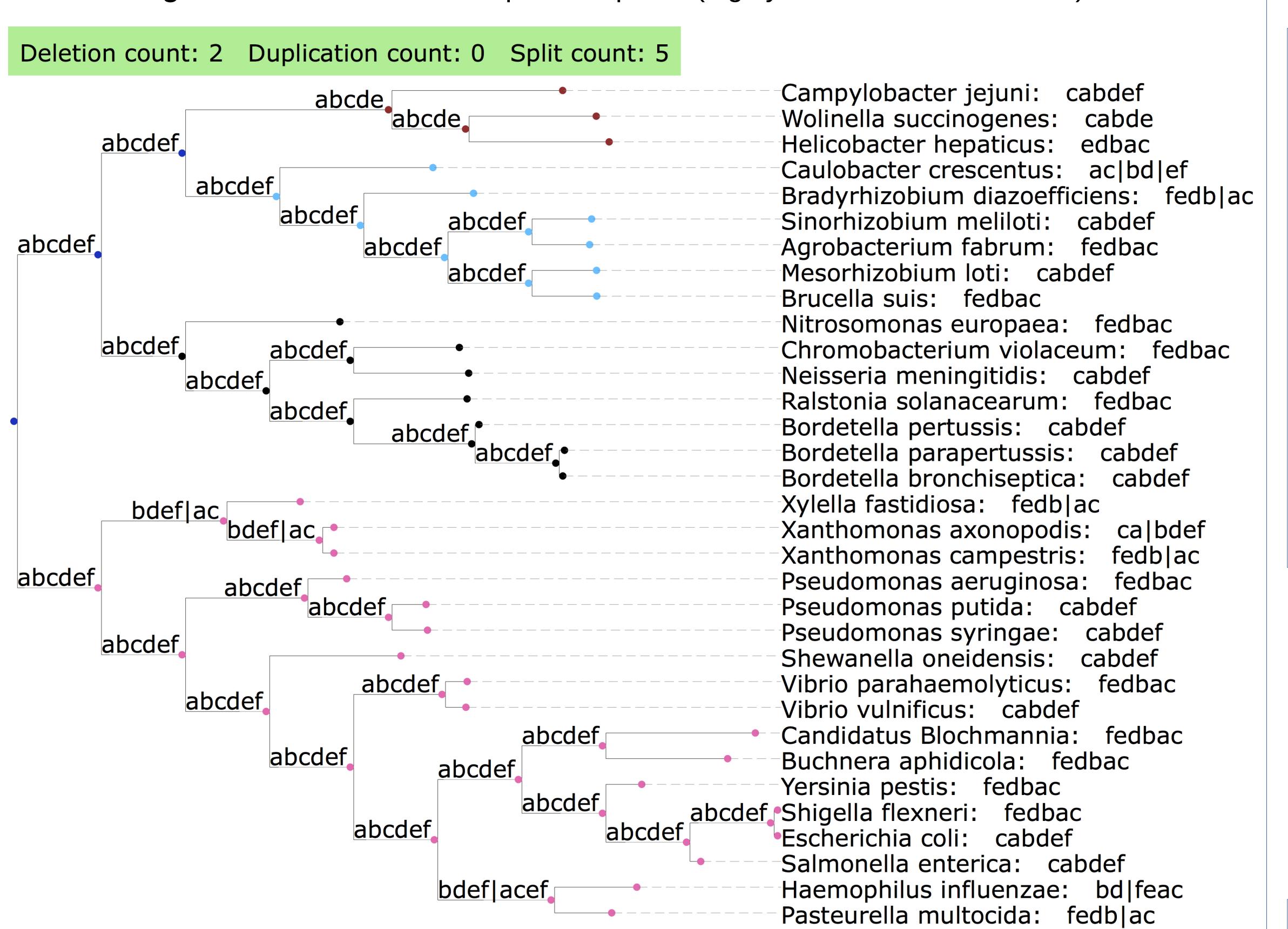
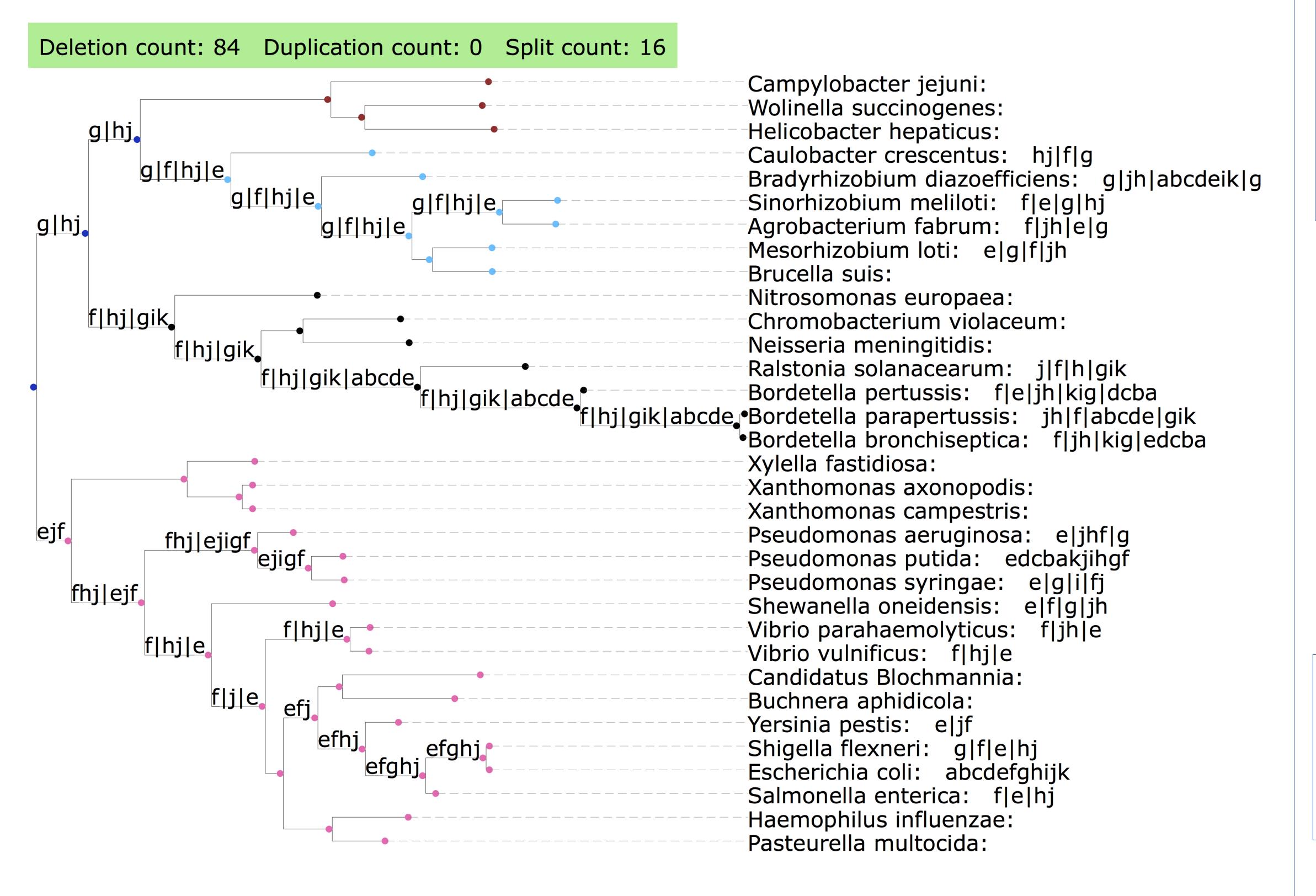


Figure 4. Reconstruction of paaABCDEFGHIJK (less conserved orthoblocks)



References

- David C. Ream, Asma R. Bankapur and Iddo Friedberg An event-driven approach for studying gene block evolution in bacteria (2015) *Bioinformatics* 31(13) 2075-83
- Arjun Bhutkar , William M Gelbart and Temple F Smith Inferring genome-scale rearrangement phylogeny and ancestral gene order: a *Drosophila* case study (2007) *Genome Biology* 8:R236

Paper: An Event-driven approach for studying gene block evolution



Code used in this study:

