Phylogenetic investigation of Gammaproteobacteria proteins involved in exogenous long-chain fatty acid acquisition and assimilation

Saksham Saksena[1], Kwame Forbes[2], Nipun Rajan[3], David Giles[4]

[1]College of Arts and Sciences, Vanderbilt University, Nashville, TN, USA; [2]Curriculum in Bioinformatics and Computational Biology, Department of Genetics, University of North Carolina Chapel Hill; [3]Houston High School; [4]Department of Biology, Geology and Environmental Science, The University of Tennessee at Chattanooga

Background

- Fatty Acids (FA) can be sourced exogenously or *de novo*.
- Exogenous FA incorporation into the cell membrane causes increased membrane permeability to hydrophobic compounds and altered resistance to certain antibiotics, ultimately modulating virulence.

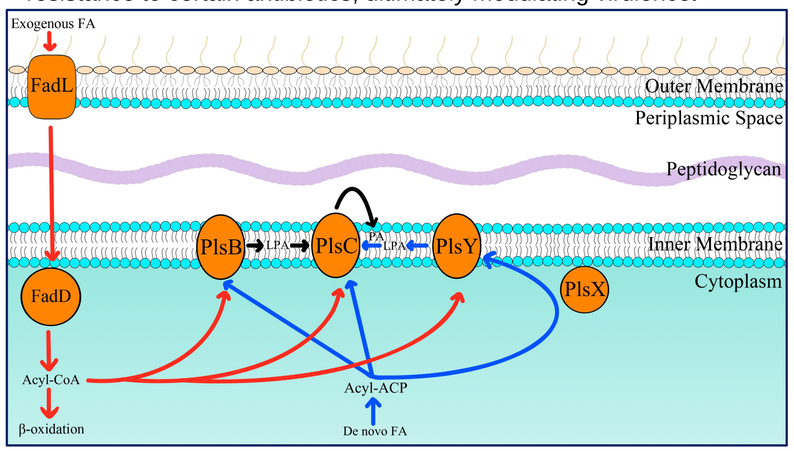

Fig. 1. Fatty Acid Metabolism (From [1]). Two pathways for FA cell membrane incorporation exist: PlsB/PlsC and PlsX/PlsY/PlsC. Many orders of Gammaproteobacteria host both pathways.

- Multiple homologs of FadL and FadD are encoded by numerous species.

Study Aims

- **This study aims to address the dynamics of structure-function, genomic context, and ecology/evolutionary perspective behind the multiple homologs phenomena.**
- Develop phylograms for FadL, FadD, PlsB, PlsC, PlsX, and PlsY across Gammaproteobacteria species to visualize relationships
- Investigate operon arrangement and potential FA preferences for FadL

Methods

Obtained amino acid & nucleotide sequences from NCBI BLAST with base sequences from *Vibrio cholerae* and *Escherichia coli*

↓

Identified suitable outgroups by making test phylograms

↓

Created full phylograms using Mr. Bayes and RAxML algorithms for both amino acid and nucleotide sequences for all investigated proteins

↓

Investigated potential operon arrangements of FadL homolog genes

↓

Identified 17 representative FadL homologs and modeled them via Swiss Model using previously crystalized *E. coli* and *Pseudomonas aeruginosa*

↓

Docked predicted structures with 5 fatty acids via AutoDock Vina: Linoleic Acid, Arachidonic Acid, Docosahexaenoic Acid, LDAO, and Oleic Acid

↓

Classified docking into 4 nodes as identified by Turgeson *et al.* [2]

Phylogenetic Analyses

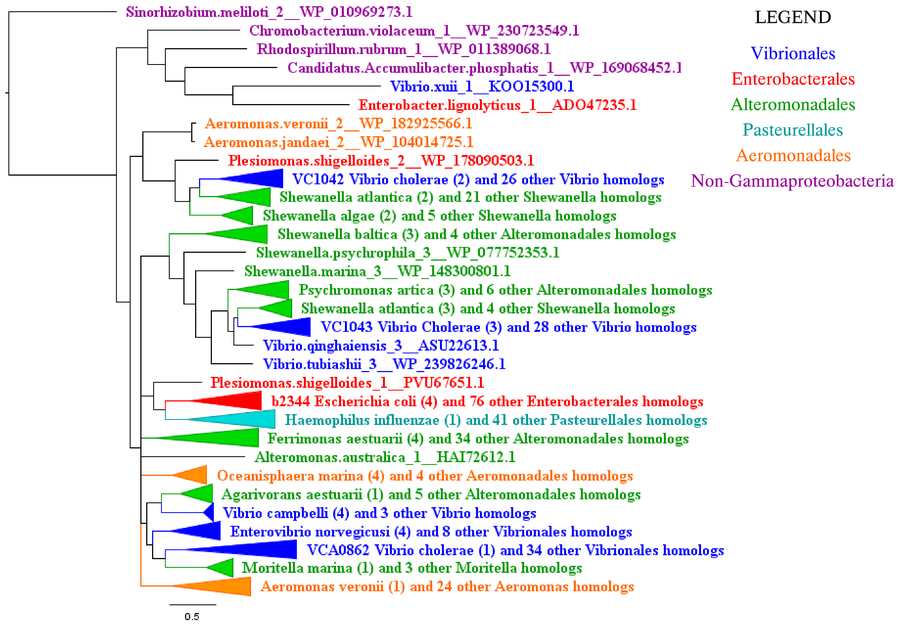

Fig. 2. FadL Condensed Phylogram (adapted from [1]). FadL Phylogram made using amino acid sequences and algorithm Mr. Bayes condensed by collapsing mostly homogenous branches in terms of order. Please see Saksena *et al.* [1] for all the full phylograms from this study.

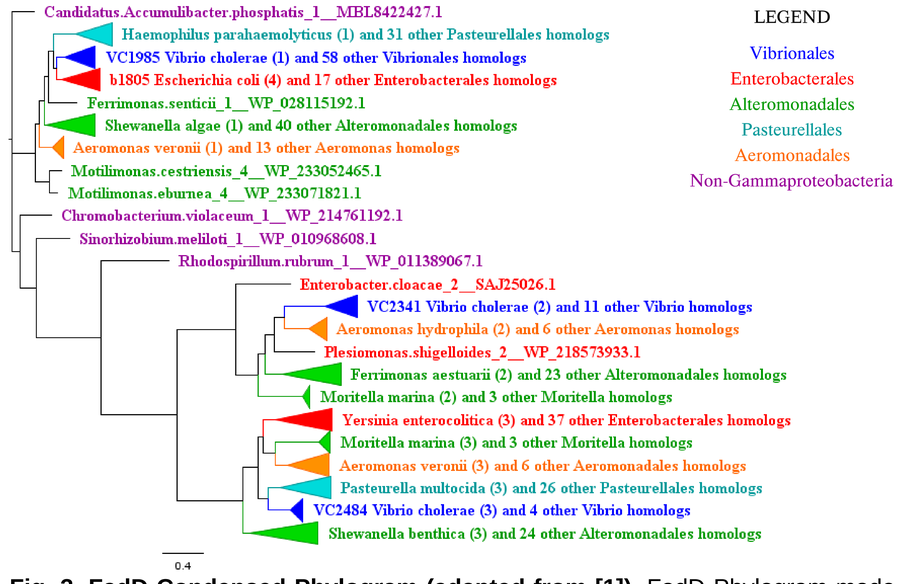

Fig. 3. FadD Condensed Phylogram (adapted from [1]). FadD Phylogram made using amino acid sequences and algorithm Mr. Bayes condensed by collapsing mostly homogenous branches in terms of order. Please see Saksena *et al.* [1] for all the full phylograms from this study.

- Homologs seem to cluster based on either order or which base sequence they were blasted from.
- Seven potential operon structures were identified.
- Operon structures mostly match branches of the phylogram.

Molecular Binding Investigation

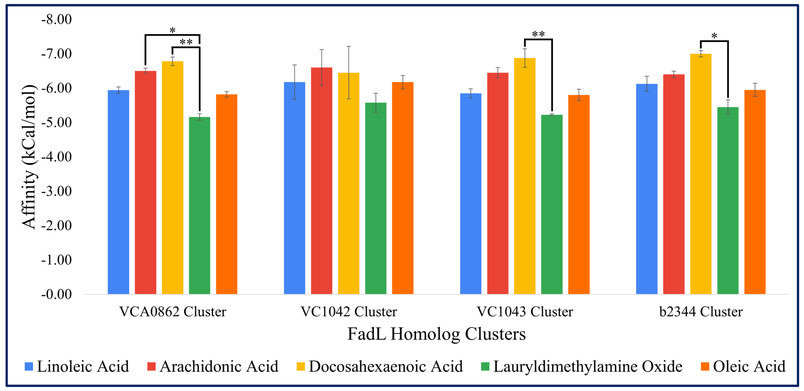

Fig. 4. Mean Binding Affinities at S3 Kink for each Homolog Cluster (from [1]). The brackets above the bars represent significant differences in the mean affinities. If the p-value is less than 0.05, there is one asterisk above. If the p-value is less than 0.01, there are two asterisks above.

- S3 kink was the only node with consistent fatty acid binding.
- FadL homolog cluster does not affect binding affinity ($p = 0.6531$).
- Fatty acids exhibit unique binding patterns ($p < 0.001$).

Conclusions

- The tendency of most homologs to cluster with the base sequences they were blasted from rather than order suggests a more ancestral origin of FadL and FadD.
- The potential operon structure analysis validates phylogenetic findings as homologs clustered based on operon structure more than other factors
- Similar binding patterns between the different FadL homologs suggests that the role of FadL is critical, necessitating multiple homologs to ensure function.

Future Work

- Verify that potential operon structures are indeed operons to corroborate findings from our operon interrogation
- Crystalize more FadL homologs to validate predicted models and improve accuracy of future FadL docking analyses
- Conduct a broader docking analyses with more FadL homologs to determine nuances in FA preferences
- Extend docking analyses to FadD

Acknowledgements

This work was supported by the National Science Foundation (Award 1852042, 2149956).

We would like to thank Fernando Alda (Assistant Professor, Department of Biology, Geology, and Environmental Science, University of Tennessee at Chattanooga) and Bradley Harris (Associate Professor, Civil and Chemical Engineering Department, University of Tennessee at Chattanooga) for reviewing our work and providing expert feedback. We also thank the internal support from Hong Qin (Professor, College of Engineering and Computer Science, University of Tennessee at Chattanooga), the Department of Computer Science and Engineering, and the Office of the Vice Chancellor of Research at the University of Tennessee at Chattanooga.

References

1. Saksena SR, Forbes K, Rajan N, Giles D. Phylogenetic investigation of Gammaproteobacteria proteins involved in exogenous long-chain fatty acid acquisition and assimilation. *Biochemistry and Biophysics Rep.* 2023;35:101504. Published 2023 Jul 4. doi:10.1016/j.bbrep.2023.101504
2. Turgeson A, Morley L, Giles D, Harris B. Simulated docking predicts putative channels for the transport of long-chain fatty acids in *Vibrio cholerae*. *Biomolecules*. 2022;12(9):1269. Published 2022 Sep 9. doi:10.3390/biom12091269