

Letter to the Editor

The Histidine Operon Is Ancient

Morgan N. Price,^{1,2} Eric J. Alm,^{2,5} Adam P. Arkin^{1–4}

¹ Lawrence Berkeley National Lab, Berkeley, CA 94720, USA

² Virtual Institute for Microbial Stress and Survival, Berkeley, CA 94720, USA

³ Department of Bioengineering, University of California at Berkeley, Berkeley CA 94720, USA

⁴ Howard Hughes Medical Institute of Berkeley, Berkeley, CA 94720, USA

⁵ Departments of Biological Engineering and Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Fani et al. (2005) studied the organization of the histidine synthesis genes across the Proteobacteria. They reported that the unified operon *hisGDC(NB)HAF(IE)* found in *Escherichia coli*, where *hisNB* and *hisIE* are gene fusions, is broken up in many Proteobacteria. They considered that the common ancestor of the Proteobacteria may have had this unified organization and it was then broken up in several lineages, or that the unified operon may have evolved by joining several pieces. They favored the joining model, with the unified operon forming in the ancestor of the β/γ -Proteobacteria, because they thought it unlikely that the operon would independently break up multiple times. However, they restricted their analysis to the Proteobacteria and considered other genomes as being outside the scope of their study.

As can be seen with the MicrobesOnline comparative genome browser (e.g., <http://microbesonline.org/hisoperon>), bacteria across many phyla have a compact histidine operon, including high-GC Gram positives (e.g., *Mycobacterium tuberculosis* has *hisDCBH-impA-hisFI*), low-GC Gram positives (*Bacillus subtilis* has *hisZGDBHAFI*), *Thermotoga maritima* (*hisGDCBHAFI*), Euryarchaeota (*Pyrococcus furiosus* has *hisZGDBHAF(IE)C*), and Crenarchaeota (*Sulfolobus solfataricus* has *hisCGABFDEHI*). This broad phylogenetic distribution suggests that the unified histidine operon is ancient. Consistent with

the hypothesis of a single origin, *hisDBHAFI* are in the same relative order in all of these groups except the Crenarchaeota. Also consistent with this hypothesis, some δ -Proteobacteria have a unified histidine operon (e.g., *Geobacter metallireducens* has *hisGDCBHAFI*). The δ -Proteobacteria branched early in the evolution of the Proteobacteria (probably after the ϵ 's and definitely before the α 's diverged from the β/γ 's) and were not included in Fani et al.'s study (the data may not have been available).

Could the β/γ -Proteobacteria have formed the unified histidine operon and then transported it to other lineages? To test this hypothesis we built a concatenated protein tree from the eight genes *hisGDCBHAFI*. These eight genes have easily identifiable orthologues (from bidirectional best hits or single-copy COGs) in most genomes. Given 122 genomes that contain all 8 genes, we built 8 alignments with MUSCLE, concatenated them, removed positions with gaps or adjacent to gaps, and built a consensus neighbor-joining tree with PHYLIP NEIGHBOR and CONSENSE. We used 100 bootstraps and gamma-distributed rates in eight categories with a coefficient of variation of 1.14 (estimated using tree-puzzle).

As shown in Fig. 1, the unified histidine operon is present across the tree and not only in sequences affiliated with the β/γ -Proteobacteria. The low-GC Gram positives, the high-GC Gram positives, and diverse thermophiles form well-supported clades containing the unified operon. It also appears that the β -Proteobacteria have rather distinct sequences from most (but not all) of the γ -Proteobacteria, so

Correspondence to: Eric J. Alm, Massachusetts Institute of Technology, Bldg 48-317, 77 Massachusetts Ave., Cambridge, MA 02139 USA; email: ejalm@mit.edu

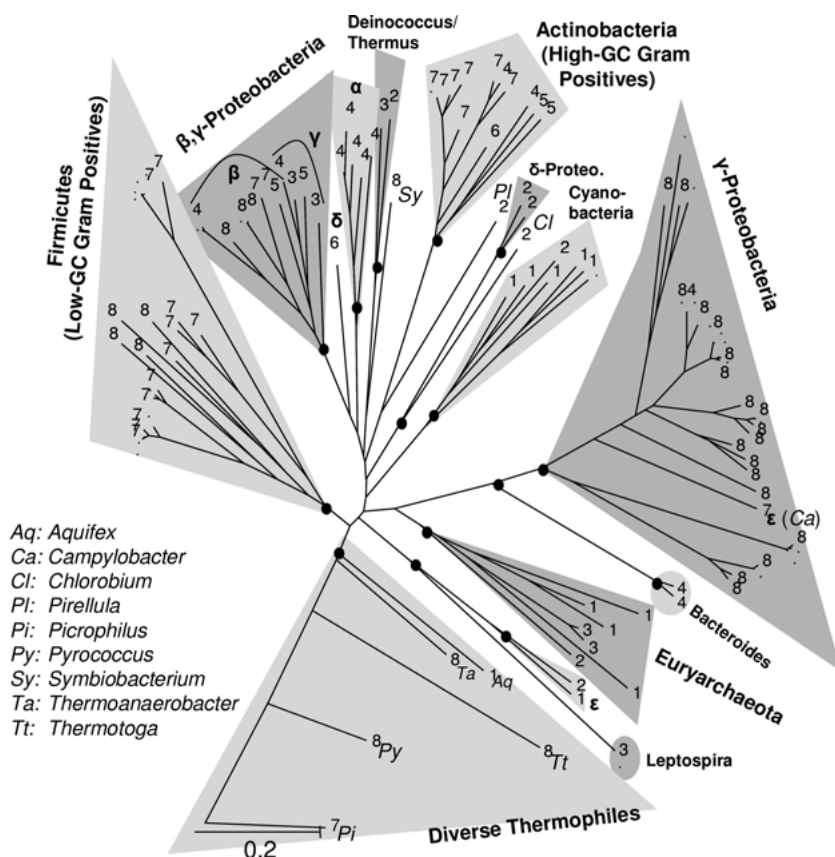


Fig. 1. A concatenated protein distance tree of *hisGDCBHAFI* across 122 prokaryotes. The number on each leaf (species) indicates how many of these eight genes are present in a single operon: 7 and 8 indicate unified operons and 1 indicates that the genes are completely scattered. Additional members of a species with the same operon organization are marked with a period. The heavy circles mark well-supported major clades present in more than 90 of 100 bootstraps. The vast majority of the branches within each major clade have similarly high bootstraps. None of the larger groupings of major clades have bootstraps over 80.

that the major γ -Proteobacterial group on the right in Fig. 1 (which includes *E. coli* and other γ 's with the unified operon) may have arisen by horizontal transfer. The paraphyly of Proteobacterial *hisG*, *hisD*, and *hisC* was previously noted by Bond and Francklyn (2000).

Thus, the unified *his* operon is indeed ancient. Because horizontal transfer is common in bacteria, we argue that it is essential to examine as many genomes as possible when drawing conclusions about the origins of a character such as operon structure. Furthermore, operon structure evolves rapidly relative to the age of the Bacteria, or even the age of the Proteobacteria: about half of *E. coli* operons are broken apart in non- γ -Proteobacteria.

Reconstructions of operon evolution from parsimony may be misleading when used on large timescales.

Acknowledgment. This work was supported by a grant from the DOE Genomics:GTL program (DE-AC03-76SF00098).

References

- Bond JP, Francklyn CF (2000) Proteobacterial histidine-biosynthetic pathways are paraphyletic. *J Mol Evol* 50:339–347
 Fani R, Brilli M, Lio P (2005) The origin and evolution of operons: the piecemeal building of the proteobacterial histidine operon. *J Mol Evol* 60:370–390