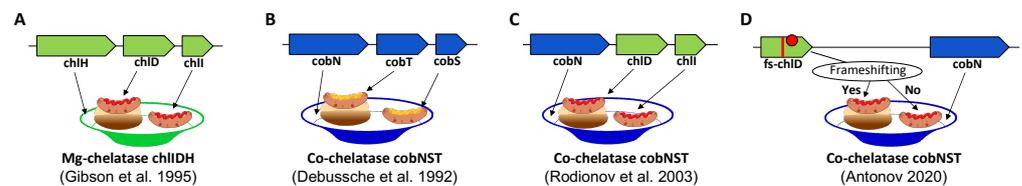


Translational frameshifting in the *chlD* gene gives a clue to the co-evolution of the chlorophyll and cobalamin biosyntheses

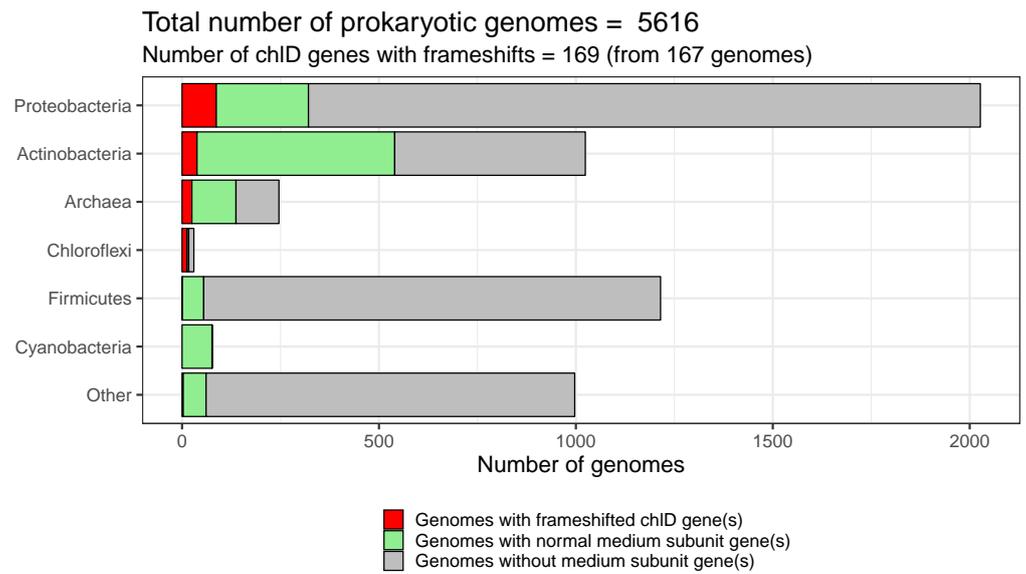
Stepan Kuznetsov¹, Alexander Milenkin¹ and Ivan Antonov^{2,3,*} 

- ¹ Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia
² Institute of Bioengineering, Research Center of Biotechnology, Russian Academy of Science, Moscow, Russia;
³ Laboratory of Bioinformatics, Faculty of Computer Science, National Research University Higher School of Economics, Moscow, Russia

Supplementary Figures



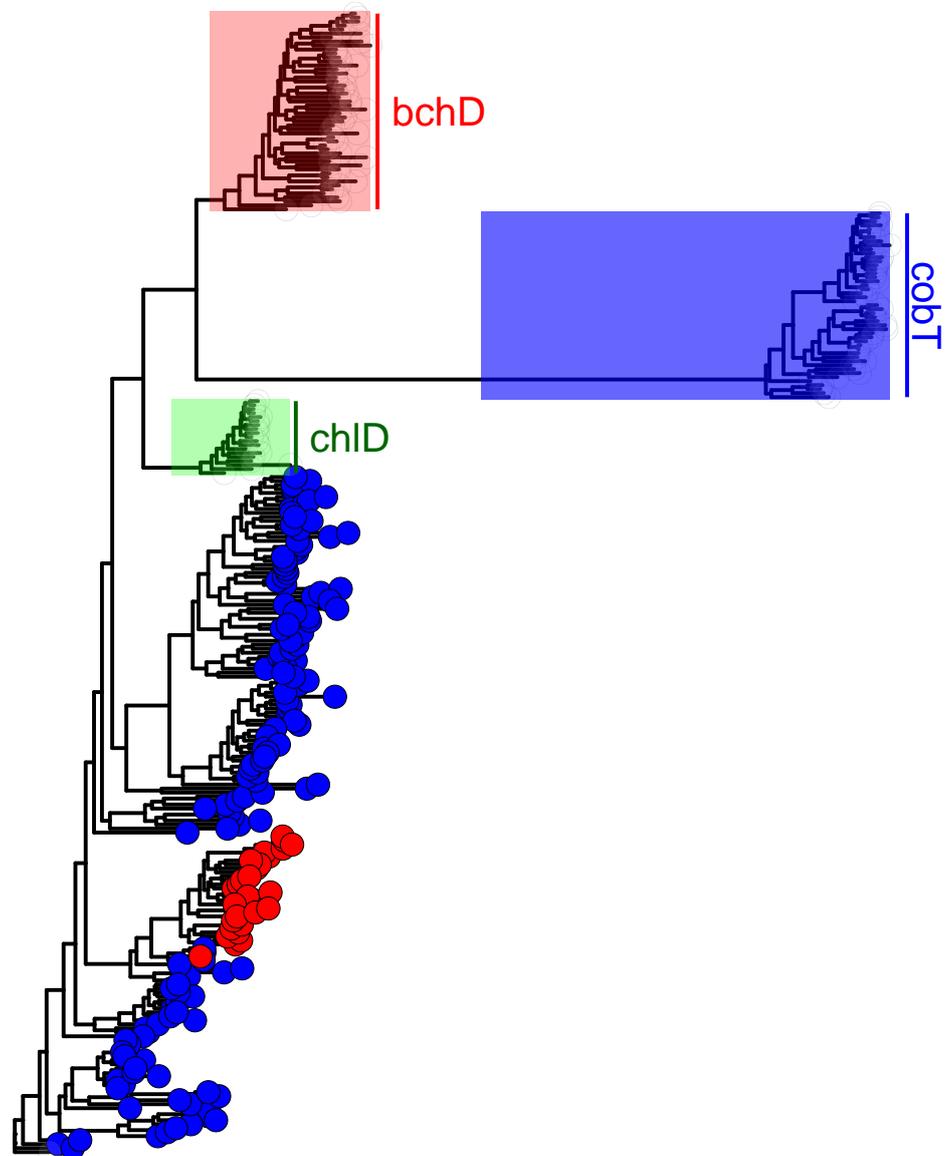
Supplementary Figure S1. The main strategies utilized by different prokaryotic species to encode Mg- and Co-chelatases. For visualization purposes, a chelatase is depicted as a sausage (small subunit) and a hot dog (medium subunit) on a plate (large subunit). This representation highlights the fact that each of the chelatases consists of three different subunits and that the small subunit is similar to the N-terminal part of the medium subunit (in other words "a sausage is a part of a hot dog").



Supplementary Figure S2. The number of the identified chID genes in the genomes from different prokaryotic phyla. In each group the fraction of the chID genes with predicted -1 and +1 frameshift is indicated by the red and green colors, respectively.

Total number of genes in the tree = 286

Number of fs-chID genes = 169

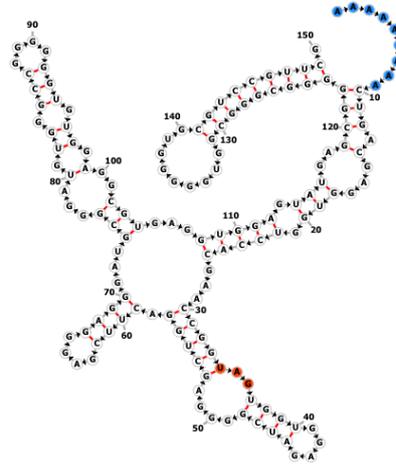


Frameshift ● -1 ● +1 ○ None

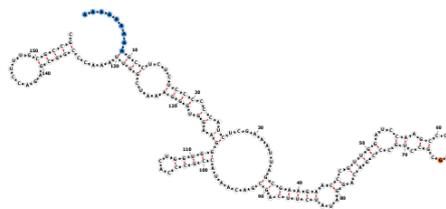
Supplementary Figure S3. The reference phylogenetic tree of the annotated medium chelataze subunit proteins and as well as the long products of the fs-chID genes. The only annotated archaeal bchD gene (METFOR_RS04410 from *Methanoregula formicica* SMSP) was used as the outgroup. The -1 and +1 frameshift types in the fs-chID genes are indicated by the blue and red colors, respectively.

A

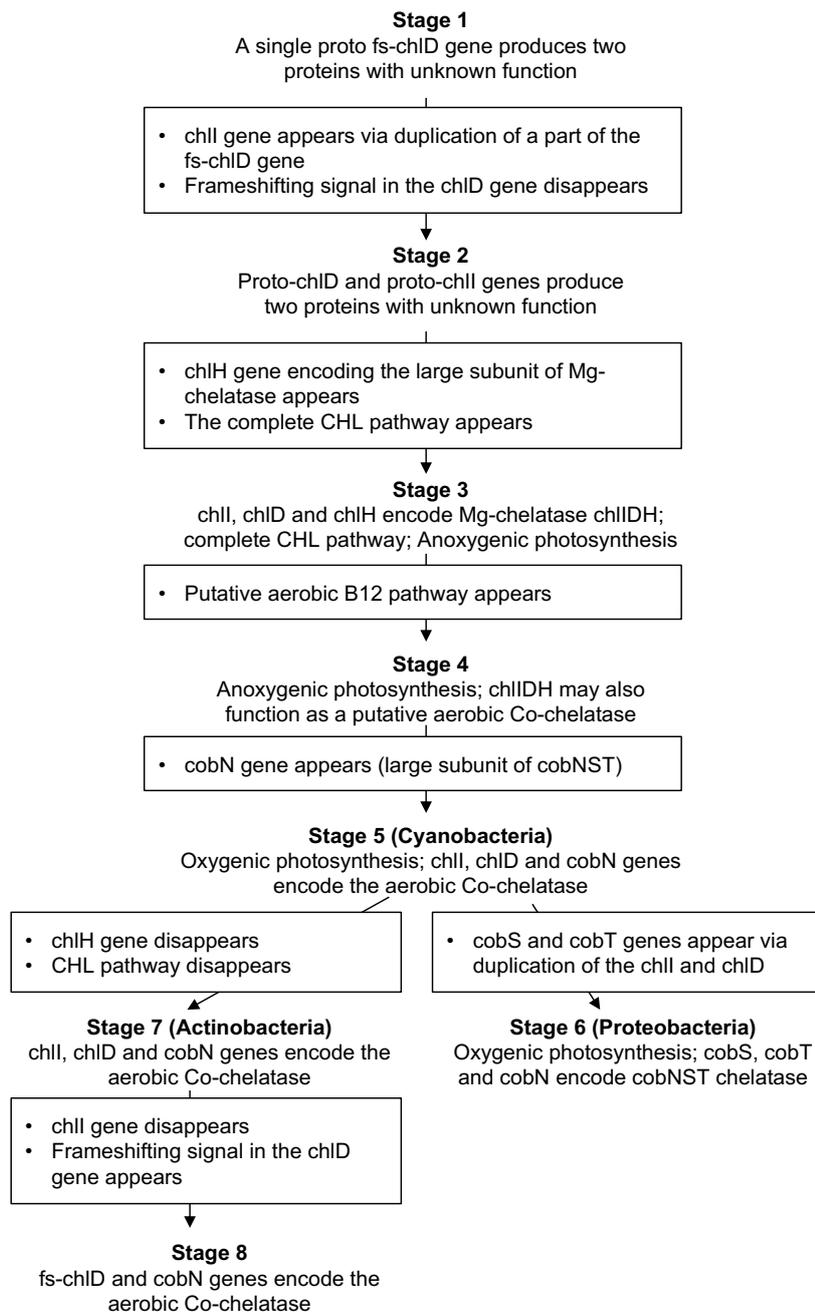
NC_017583.1:1673962:-1

**B**

NZ_LT859958.1:191466:-1



Supplementary Figure S4. Frameshifting signals identified in the *fs-chlD* genes from *Spirochaeta thermophila* DSM 6578 (A) and *Brevefilum fermentans* (B). The slippery site of the signal and the in-frame stop codon are marked by the blue and red colors, respectively. The coordinates of the predicted -1 frameshifts are NC_017583.1:1673962 and NZ_LT859958.1:191466.



Supplementary Figure S5. A putative chain of events that could happen between different stages during the evolution of the CHL and aerobic B12 pathways.