Supplementary Material

Navigating the Archaeal Frontier: Insights and Projections from Bioinformatic Pipelines

Val Karavaeva[1,2], Filipa L. Sousa[1*]

[1]Genome Evolution and Ecology Group, Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria

[2] Vienna Doctoral School of Ecology and Evolution, University of Vienna, Vienna, Austria

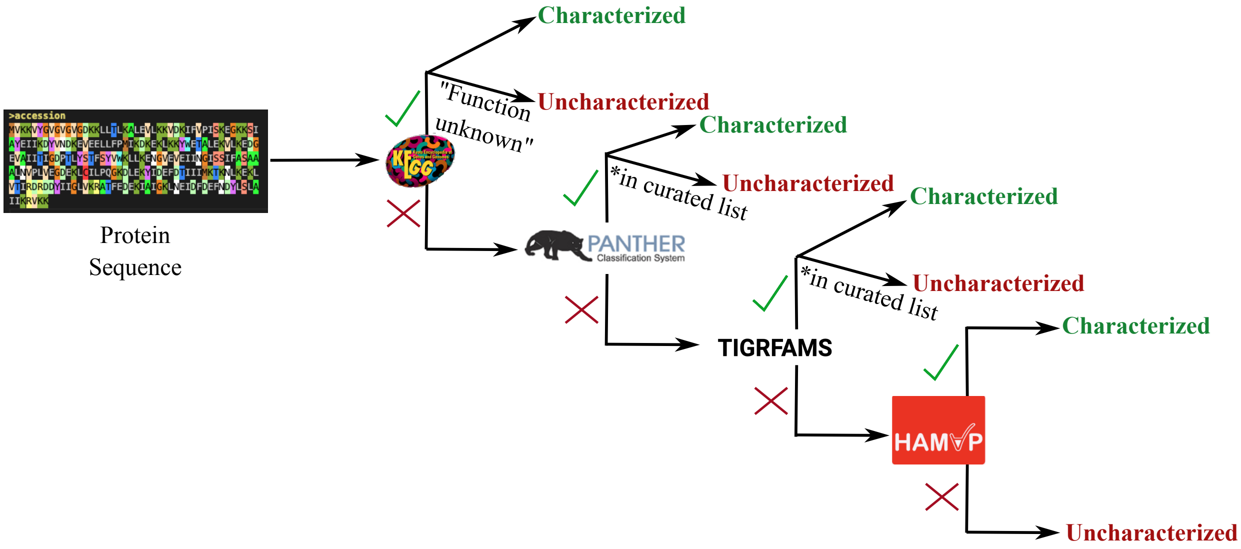
*** Correspondence:**
Filipa L. Sousa filipa.sousa@univie.ac.at

1 Supplementary Figures and Tables

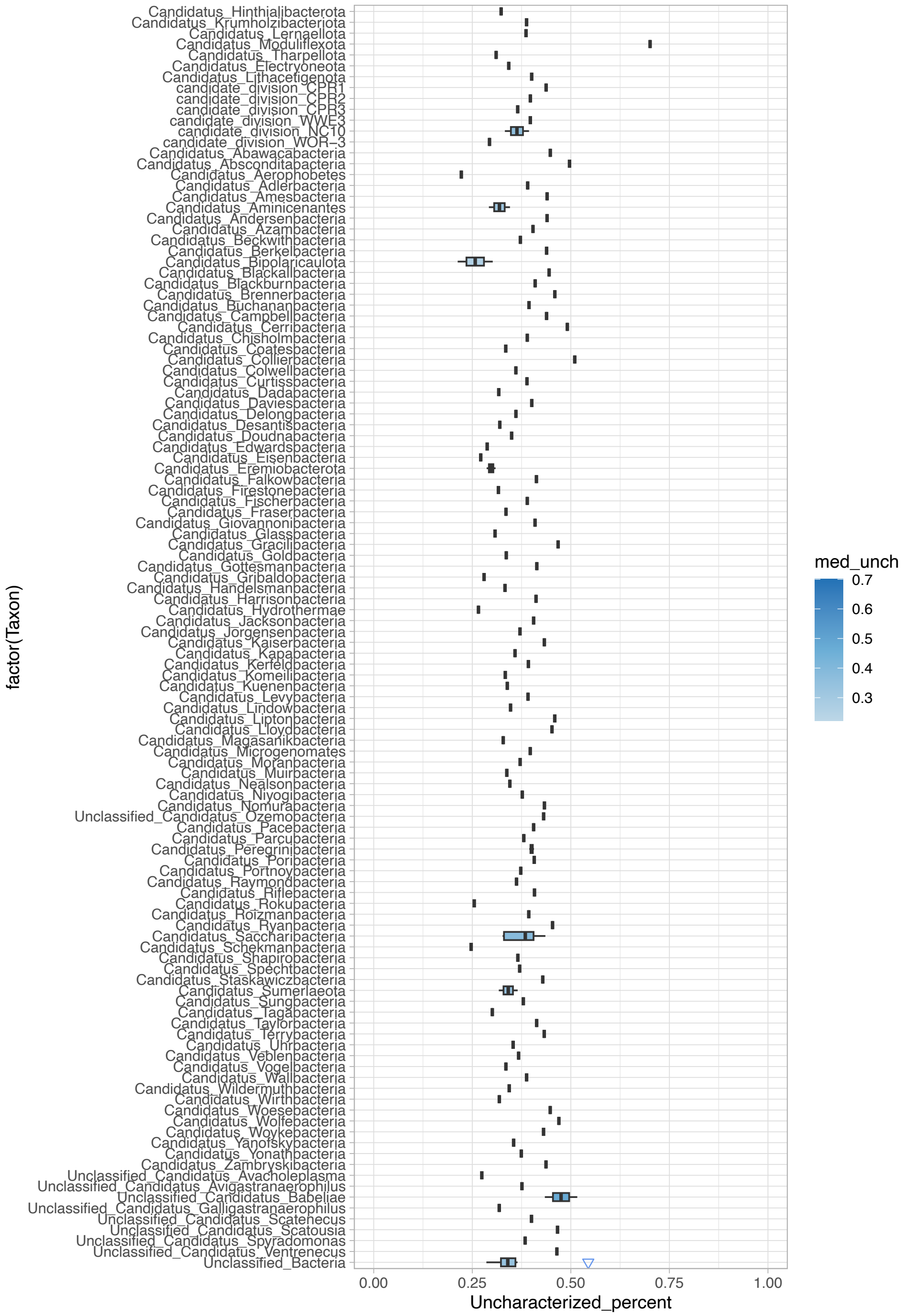
Supplementary Tables 1-7 are uploaded as separate files.

This supplementary material includes:

Supplementary Figures 1-4



Supplementary Figure 1. *The pipeline scheme to classify proteins as "characterized" or "uncharacterized" (as described in Materials and Methods).*



Supplementary Figure 2. *Percentage of unclassified bacterial proteins for Candidate phyla according to pipeline classification (see Materials and Methods) per phylum. For complete taxonomic information see Suppl. Table 1. For exact percentages, see Suppl. Table 4.*


Siroheme biosynthesis

Siroheme-dependent heme biosynthesis

Coproporphyrin-dependent heme biosynthesis

Protoporphyrin-dependent heme biosynthesis

Supplementary Figure 3. *Presence of cofactor biosynthesis in archaea (per phylum for most, per class for Euryarchaeota, with Methanobacteria, Methanococci, and Methanopyri grouped into the Methanomada supergroup) based on modified KEGG modules. From top to bottom: presence of siroheme biosynthesis, presence of siroheme-dependent heme biosynthesis, presence of coproporphyrin-dependent heme biosynthesis, presence of protoporphyrin-dependent heme biosynthesis. Dark purple indicates that the full module is present, light purple marks the presence of the incomplete module, white shows the absence of the module in a lineage.*

Ubiquinone biosynthesis

Menaquinone biosynthesis

Riboflavin biosynthesis

Supplementary Figure 4. *Presence of cofactor biosynthesis in archaea (per phylum for most, per class for Euryarchaeota, with Methanobacteria, Methanococci, and Methanopyri grouped into the Methanomada supergroup) based on modified KEGG modules. From top to bottom: presence of ubiquinone biosynthesis, presence of menaquinone biosynthesis, presence of riboflavin biosynthesis.*

Dark purple indicates that the full module is present, light purple marks the presence of the incomplete module, white shows the absence of the module in a lineage.