A044 · Repeated repurposing of nitrogenase-like proteins revealed by proteome-scale interaction prediction

Presenter: Subhadeep Chowdhury — Bioinformatics program, Faculty of Computing and Data Sciences, Boston University, Boston, MA, USA

Authors: Subhadeep Chowdhury

Session: Day 1 · Thu Oct 1, 2026 · 2:15–4:15 PM

Nitrogen is abundant in the atmosphere but inaccessible to nearly all life, and a single microbial enzyme (i.e., nitrogenase) is what makes it available to the rest of the biosphere. Nitrogenase has two parts: an ATP-powered electron donor (NifH) and the catalytic core that binds N2 (NifD and NifK). However, many genomes encode nifH-like genes without nifD/nifK, making nifH-only annotations potentially misleading. Here we separate true nitrogen-fixers (complete nifHDK) from pseudo-nitrogen-fixers (nifH-like genes only) across 6879 bacterial and archaeal genomes and ask what the retained proteins do. Using a genome-resolved framework, we compared their global distribution and carbon, nitrogen, and energy metabolism. Pseudo-nitrogen fixers are widely distributed, and we identified two new archaeal phyla (Halobacteriota and Methanobacteriota) that carry pseudo-nifH. Compared with true nitrogen fixers, pseudo nitrogen fixers are metabolically streamlined and lack carbon fixation and multiple nitrogen transformation pathways. In contrast, pseudo nitrogen fixers are enriched in strict anaerobic and host-associated lifestyles supported by acetate utilization, sulfur metabolism, hydrogen metabolism, and methane production. To infer functions of retained nifH-like proteins, we integrated genomic language models with protein-protein interaction prediction. Our analyses suggest that nitrogenase-like proteins in pseudo nitrogen fixers have undergone neofunctionalization, with predicted roles in tetrapyrrole biosynthesis (e.g., chlorophyll), sulfur scavenging and methanogenesis, and additional paralogs linked to nutrient import, siderophore-mediated iron uptake, and lipid flipping. By distinguishing true and pseudo nitrogen fixers, our study refines genome-based inference of nitrogen fixation and reveals hidden metabolic innovation in globally distributed nitrogenase-like proteins.

Repeated repurposing of nitrogenase-like proteins revealed by proteome-scale interaction prediction

Subhadeep Chowdhury[1], Juan José Pierella Karlusich[2], Daniel Segre[1,3], Robinson W Fulweiler[3,4]

[1]Bioinformatics program, Faculty of Computing and Data Sciences, Boston University, Boston, MA, USA
[2]Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC and University of Seville, Seville, 41092 Spain
[3]Department of Biology, Boston University, Boston, MA, USA
[4]Department of Earth and Environment, Boston University, Boston, MA, USA

Nitrogen is abundant in the atmosphere but inaccessible to nearly all life, and a single microbial enzyme (i.e., nitrogenase) is what makes it available to the rest of the biosphere. Nitrogenase has two parts: an ATP-powered electron donor (NifH) and the catalytic core that binds N_2 (NifD and NifK). However, many genomes encode *nif*H-like genes without *nif*D/*nif*K, making *nif*H-only annotations potentially misleading. Here we separate true nitrogen-fixers (complete *nif*HDK) from pseudo-nitrogen-fixers (*nif*H-like genes only) across 6879 bacterial and archaeal genomes and ask what the retained proteins do. Using a genome-resolved framework, we compared their global distribution and carbon, nitrogen, and energy metabolism. Pseudo-nitrogen fixers are widely distributed, and we identified two new archaeal phyla (Halobacteriota and Methanobacteriota) that carry pseudo-*nif*H. Compared with true nitrogen fixers, pseudo nitrogen fixers are metabolically streamlined and lack carbon fixation and multiple nitrogen transformation pathways. In contrast, pseudo nitrogen fixers are enriched in strict anaerobic and host-associated lifestyles supported by acetate utilization, sulfur metabolism, hydrogen metabolism, and methane production. To infer functions of retained *nif*H-like proteins, we integrated genomic language models with protein-protein interaction prediction. Our analyses suggest that nitrogenase-like proteins in pseudo nitrogen fixers have undergone neofunctionalization, with predicted roles in tetrapyrrole biosynthesis (e.g., chlorophyll), sulfur scavenging and methanogenesis, and additional paralogs linked to nutrient import, siderophore-mediated iron uptake, and lipid flipping. By distinguishing true and pseudo nitrogen fixers, our study refines genome-based inference of nitrogen fixation and reveals hidden metabolic innovation in globally distributed nitrogenase-like proteins.