Mapping the unseen structural diversity of hydrogenases *in silico*

James P. Lingford[1], Gavin J. Knott[2], Chris Greening[1].

1. Department of Microbiology, Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia.
2. Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia.

Hydrogenases are metalloenzymes that catalyse the simplest of chemical reactions: the interconversion of protons and electrons into hydrogen gas ($2H^+ + 2e^- \rightleftharpoons H_2$). This simple reaction is important for microbial life, and numerous microbes found across the tree-of-life and across diverse environments use hydrogenases in a variety of different metabolic pathways, where they either consume H_2 as an energy source, or produce H_2 during fermentation. Underpinning the diversity of hydrogenase metabolic functions is a diversity of protein structures. The limited number of experimentally determined hydrogenase structures in the PDB show a variety of multimeric forms, each of which confer different metabolic functions, catalytic behaviour, and protections to the metalloactive site from oxygen poisoning. These structures, however, likely represent only the tip-of-the-iceberg of hydrogenase structural diversity present in nature. The emergence of new protein structure prediction and comparison tools, like AlphaFold2 and Foldseek, now allows us to explore this hidden structural diversity at a massive scale. To this end, we searched the GTDB to retrieve hydrogenase sequences and their genomic contexts, clustered the hits using MMseqs2, and predicted unique structures for over 78,000 putative [NiFe]-hydrogenases and over 39,000 putative [FeFe]-hydrogenases using Boltz-2. Clustering by structural similarity recapitulates previous sequence-based classification scheme for hydrogenase catalytic subunits, and also reveals huge diversity in the hydrogenase small subunits. Understanding how structure determines the different functions of hydrogenases may help us to design bio-inspired hydrogen catalysts.