

HydDB: A web tool for hydrogenase classification and analysis

Dan Søndergaard^a, Christian N. S. Pedersen^a, Chris Greening^{b, c*}

^a Aarhus University, Bioinformatics Research Centre, C.F. Møllers Allé 8, Aarhus DK-8000, Denmark

^b The Commonwealth Scientific and Industrial Research Organisation, Land and Water Flagship, Clunies Ross Street, Acton, ACT 2060, Australia

^c Monash University, School of Biological Sciences, Clayton, VIC 2800, Australia

Correspondence:

Dr Chris Greening (chris.greening@monash.edu), Monash University, School of Biological Sciences, Clayton, VIC 2800, Australia

Dan Søndergaard (das@birc.au.dk), Aarhus University, Bioinformatics Research Centre, C.F. Møllers Allé 8, Aarhus DK-8000, Denmark

Figure S1. Sequence similarity networks showing the relationships between closely related subgroups of [NiFe]-hydrogenases as narrow $\log E$ filters.

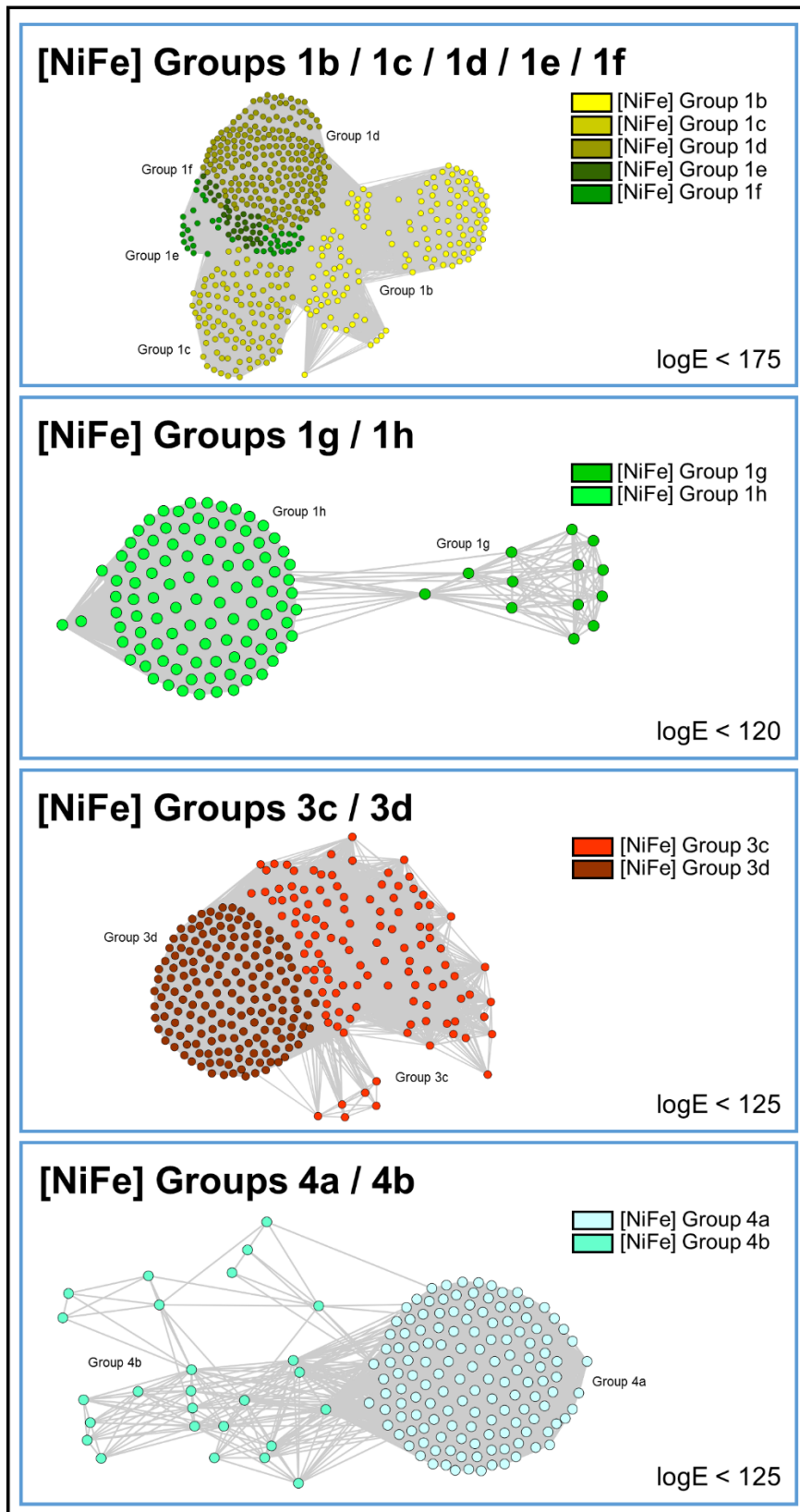


Figure S2. Screenshot showing interface of HydDB classification page.

HydDB Classify Browse Information Pages

Classify

HydDB provides access to an accurate classifier for hydrogenase sequences and a curated database of hydrogenases by known type. The service is provided by the School of Biological Sciences, Monash University and the Bioinformatics Research Centre, Aarhus University.

Classify

Sequences

Sequences File

No file chosen

Check sequences using CDD?
If enabled, HydDB will use CDD to check whether the submitted sequences encode catalytic subunits of putative before classification. Since this step is time-consuming, you may want to uncheck this option if you are certain your sequences encode hydrogenase catalytic subunits.

Mail

If an e-mail address is provided, a mail will be sent when the job succeeds or fails.

Instructions

To use the classifier to predict the type of one or more hydrogenases from sequence, either:

- paste your FASTA-formatted protein sequences into the text area, or
- upload a FASTA-formatted file with your protein sequences.

Press the "Submit" button to upload the sequences and begin the classification.

If you provided an e-mail address you will receive an e-mail when your job finishes or fails including a link to the results. You will also be able to download the results as a CSV file.

Only sequences encoding the catalytic subunits of hydrogenases will be classified, i.e. those binding the [NiFe]-centre (NiFe-hydrogenases), [FeFe]-centre (FeFe-hydrogenases), or [Fe]-centre (Fe-hydrogenases). Electron-transfer subunits, accessory proteins, and maturation factors cannot be classified by this service.

Limits

A job can at most run for 2 hours. This should be enough for about 2500 sequences to be classified. Results will be stored for 2 weeks. However, we recommend to download the results as they may be deleted due to the rare event of a power outage or server crash.

Statistics

| | |
|---|-----|
| Jobs completed in total | 40 |
| Sequences classified in total | 232 |
| Jobs completed in the last 24 hours | 0 |
| Sequences classified in the last 24 hours | 0 |

Figure S3. Screenshot showing the information provided in the data entry pages for 3248 individual hydrogenases in HydDB.

HydDB [Classify](#) [Browse](#) [Information Pages](#)

Entry WP_004030875.1

| | |
|--------------------------------|-----------------------------|
| Phylum | Euryarchaeota |
| Order | Methanobacteriales |
| Organism | Methanobacterium formicicum |
| Hydrogenase | [Fe] |
| Activity (Predicted) | Bidirectional |
| Oxygen Tolerance (Predicted) | Tolerant |
| Subunits (Predicted) | 1 |
| Metal Centres (Predicted) | Fe ion |
| Accessory Subunits (Predicted) | None |

```
MKLAILGAGCYRTHAASGITNFSRACEVAEQVQKPEIAMTHSTIAMGAEKELAGIDEIVVSDPVFDNDFTVIDDFEYEAVIEAHKDPESIMPQIREKVNVAKDLKPPKG
AIHFTHPEDLGFVTTDDNEAVQDADWMTWFPKGDMMQGIKKEFADNLKEGAILTHACTVPTTFQKIFEDLSSDEMNIAPKVVNVSYPHGAVPKMGQVYIAEGYASEDAI
CKLVDWGVAAARGDAFKLPAELLPVCDMCSALTAITYAGILSYRDSVMNIIIGAPAGFAQWIAKESLTQVTDLMNSVGDHMEKLDPGALLGTADSMNFGAAADVLPVLEVL
ENRKGKGP TCNI
```

Figure S4. Screenshot showing the capacity for browsing hydrogenase data entries in HydDB.

HydDB Classify Browse Information Pages

Browse

Filter (matches 3248 entries) Download

Phylum: Any

Order:

Organism:

Ncbi accession:

Hydrogenase class: Any

Subunits predicted: Any

Oxygen tolerance predicted: Any

Activity predicted: Any

Metal centres predicted: Any

Apply Clear

| NCBI Accession | Organism | Hydrogenase Class | Phylum | Order | Activity (Predicted) | Oxygen Tolerance (Predicted) | Subunits (Predicted) | Metal Centres (Predicted) | Accessory Subunits (Predicted) |
|----------------|--|-------------------|---------------|--------------------|----------------------|------------------------------|----------------------|---------------------------|--------------------------------|
| WP_004030875.1 | Methanobacterium formicum | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |
| WP_012955328.1 | Methanobrevibacter ruminantium | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |
| WP_019263574.1 | Methanobrevibacter smithii | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |
| WP_016357634.1 | Methanobrevibacter sp. AbM4 | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |
| WP_013296316.1 | Methanothermobacter marburgensis | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |
| WP_010876766.1 | Methanothermobacter thermautotrophicus | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |
| WP_013413799.1 | Methanothermus fervidus | [Fe] | Euryarchaeota | Methanobacteriales | Bidirectional | Tolerant | 1 | Fe ion | None |

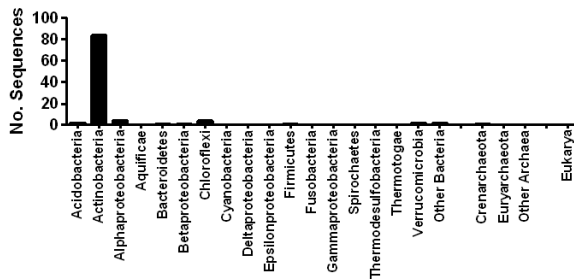
Figure S5. Screenshot showing the detailed content of the information pages about each hydrogenase class on HydDB. Equivalent information pages are available for all 38 hydrogenase classes defined in this work (**Table 1**).

[NiFe] Group 1h-hydrogenase

This entry was last updated at: June 13, 2016, 11:11 a.m.

| | |
|------------------|--|
| Properties | |
| Group | [NiFe] Group 1: Respiratory H ₂ -uptake [NiFe] hydrogenases |
| Subgroup | [NiFe] Group 1h: Actinobacteria-type |
| Function | Hydrogenotrophic respiration using O ₂ as terminal electron acceptor. Enzyme scavenges electrons from atmospheric H ₂ to fuel respiratory chain during carbon-starvation. Route of electron transfer unresolved. |
| Activity | H ₂ -uptake (unidirectional, high-affinity) |
| Oxygen tolerance | O ₂ -tolerant or O ₂ -insensitive |
| Localisation | Membrane-associated? |

| | |
|------------------------|--|
| Distribution | |
| Ecosystem distribution | Upland soils, plant tissues, possibly surface waters |
| Taxonomic distribution | Widespread among obligately aerobic soil bacteria, especially Actinobacteria, Acidobacteria, and Chloroflexi |



| | |
|---|--|
| Genetic Organisation | |
| e.g. <i>Mycobacterium smegmatis</i> (WP_003893634.1): | |
| | |

| | |
|---------------------|---|
| Architecture | |
| Structures | 5AA5 (<i>Ralstonia eutropha</i> , 2.5 Å resolution, active) |
| Subunits | 3? |
| Subunit description | HyhL (hydrogenase large subunit) HyhS (hydrogenase small subunit) HhyE (putative iron-sulfur protein and proposed physiological electron acceptor) |
| Catalytic site | [NiFe]-centre |
| FeS clusters | Proximal: 3Cys1Asp[4Fe4S] Medial: 4Cys[4Fe4S] Distal: 3Cys1His[4Fe4S] |

| | |
|---|--|
| Important Notes | |
| The <i>Robiginitalea biformata</i> and <i>Sulfolobus islandicus</i> enzymes are relatively to distantly related to the main group. No studies have yet tested whether these enzymes have a H ₂ -scavenging role like other Group 1h [NiFe]-hydrogenases. They may instead represent founding members of a functionally-distinct lineage. | |

| Sequences in this class | | | | | | | | | |
|-------------------------|---|-------------------|----------------|------------------|----------------------|------------------------------|----------------------|-------------------------------------|--------------------------------|
| NCBI Accession | Organism | Hydrogenase Class | Phylum | Order | Activity (Predicted) | Oxygen Tolerance (Predicted) | Subunits (Predicted) | Metal Centres (Predicted) | Accessory Subunits (Predicted) |
| WP_014267363.1 | <i>Granulicella mallensis</i> | [NiFe] Group 1h | Acidobacteria | Acidobacteriales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_011688202.1 | <i>Soilbacter usitatus</i> | [NiFe] Group 1h | Acidobacteria | Soilbacterales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_021597135.1 | <i>Actinomadura madurae</i> | [NiFe] Group 1h | Actinobacteria | Actinomycetales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_026402909.1 | <i>Actinomadura rifamycinii</i> | [NiFe] Group 1h | Actinobacteria | Actinomycetales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_018330638.1 | <i>Actinomycetospira chiangmaiensis</i> | [NiFe] Group 1h | Actinobacteria | Actinomycetales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_007735075.1 | <i>Rhodococcus qingshengii</i> | [NiFe] Group 1h | Actinobacteria | Actinomycetales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_003935326.1 | <i>Rhodococcus ruber</i> | [NiFe] Group 1h | Actinobacteria | Actinomycetales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |
| WP_005443931.1 | <i>Saccharomonospora azurea</i> | [NiFe] Group 1h | Actinobacteria | Actinomycetales | Aerobic Uptake | Tolerant | 3 | [NiFe]-centre, 3 x [4Fe4S] clusters | [FeS] protein |

Literature

Genetics:

- Berney, M., Greening, C., Hards, K., Collins, D., and Cook, G.M. (2014) Three different [NiFe] hydrogenases confer metabolic flexibility in the obligate aerobic *Mycobacterium smegmatis*. *Environ. Microbiol.* **16**: 318-330.
- Constant, P., Chowdhury, S.P., Hesse, L., and Conrad, R. (2011) Co-localization of atmospheric H₂ oxidation activity and high affinity H₂-oxidizing bacteria in non-axenic soil and sterile soil amended with *Streptomyces* sp. PCB7. *Soil Biol. Biochem.* **43**: 1888-1893.
- Constant, P., Chowdhury, S.P., Hesse, L., Pratscher, J., and Conrad, R. (2011) Genome data mining and soil survey for the novel group 5 [NiFe]-hydrogenase to explore the diversity and ecological importance of presumptive high-affinity H₂-oxidizing bacteria. *Appl. Environ. Microbiol.* **77**: 6027-6035.
- Greening, C., Biswas, A., Carere, C.R., Jackson, C.J., Taylor, M.C., Stott, M.B., Cook, G.M., and Morales, S.E. (2016) Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival. *ISME J.* **10**: 761-777.
- Khndiri, M., Hesse, L., Popa, M.E., Quiza, L., Lalonde, I., Meredith, L.K., Röckmann, T., and Constant, P. (2015) Soil carbon content and relative abundance of high affinity H₂-oxidizing bacteria predict atmospheric H₂ soil uptake activity better than soil microbial community composition. *Soil Biol. Biochem.* **85**: 1-9.

Physiology:

- Berney, M., Greening, C., Conrad, R., Jacobs, W.R., and Cook, G.M. (2014) An obligately aerobic soil bacterium activates fermentative hydrogen production to survive reductive stress during hypoxia. *Proc. Natl. Acad. Sci. U. S. A.* **111**: 11479-11484.
- Constant, P., Chowdhury, S.P., Pratscher, J., and Conrad, R. (2010) Streptomyces contributing to atmospheric molecular hydrogen soil uptake are widespread and encode a putative high-affinity [NiFe]-hydrogenase. *Environ. Microbiol.* **12**: 821-829.

Table S1. Validation that HydDB classifies only hydrogenase catalytic subunit sequences. HydDB excludes non-hydrogenase sequences through a combination of homology checks (sequences are only classified as hydrogenases if BLAST *E*-value of the closest hit in HydDB is less than 10^{-5}) and CDD checks (sequences are only classified as hydrogenases if signature conserved domains are found). In addition, the classifier has been specifically trained to exclude four protein families that are homologous to hydrogenase catalytic subunits (HmdII, Her, NuoD, NARF) but lack hydrogenase activity.

| NCBI Accession | Sequence type | Homology check | CDD check | Final result |
|--------------------|--|--|---|-----------------|
| WP_041979 300.1 | Validated hydrogenase catalytic subunit | Highest sequence homology with [NiFe] Group 1h ($E = 0$) | Ni,Fe-hydrogenase I large subunit (COG0374) | Hydrogenase |
| WP_011729 412.1 | P-type ATPase (unrelated to hydrogenases) | Low sequence homology with hydrogenases ($E = 5.6$) | Non-hydrogenase | Non-hydrogenase |
| WP_003895 387.1 | Chaperone (unrelated to hydrogenases) | Low sequence homology with hydrogenases ($E = 3.8$) | Non-hydrogenase | Non-hydrogenase |
| WP_013295 714.1 | HmdII (homologous with [Fe]-hydrogenases) | Highest sequence homology with HmdII ($E = 0$) | HMD (pfam03201) | Non-hydrogenase |
| WP_003901 794.1 | Ehr (homologous with [NiFe]-hydrogenases) | Highest sequence homology with Ehr ($E = 0$) | Complex1_49kDa superfamily (cl21493) | Non-hydrogenase |
| WP_003901 553.1 | NuoD (homologous with [NiFe]-hydrogenases) | Highest sequence homology with NuoD ($E = 0$) | NuoD (COG0649) | Non-hydrogenase |
| NP_114174. 1 | NARF (homologous with [FeFe]-hydrogenases) | Highest sequence homology with NARF ($E = 0$) | Fe_hyd_Ig_C (pfam02906) | Non-hydrogenase |

Table S2. Hydrogenase sequences where there is disagreement between classification by SSN and *k*-NN methods. These sequences represent six out of the total 3248 sequences analyzed, i.e. 0.0018%.

| NCBI Accession | Organism | <i>k</i> -NN Classification | SSN Classification |
|----------------|--------------------------------|-----------------------------|--------------------|
| WP_027414715.1 | Aneurinibacillus terranovensis | [NiFe] Group 1e | [NiFe] Group 1d |
| WP_027358538.1 | Desulforegula conservatrix | [NiFe] Group 3d | [NiFe] Group 3c |
| WP_012532312.1 | Geobacter bemidjiensis | [NiFe] Group 3d | [NiFe] Group 3c |
| WP_012469611.1 | Geobacter lovleyi | [NiFe] Group 3d | [NiFe] Group 3c |
| WP_004512544.1 | Geobacter metallireducens | [NiFe] Group 3d | [NiFe] Group 3c |
| WP_015839165.1 | Geobacter sp. M21 | [NiFe] Group 3d | [NiFe] Group 3c |

Dataset S1. Excel spreadsheet listing the sequence, taxonomy, and hydrogenase class of all 3248 hydrogenase catalytic subunit sequences listed in HydDB.

Dataset S2. Zip file containing the Cytoscape network for [NiFe]-hydrogenases.

Dataset S3. Zip file containing the Cytoscape network for [FeFe]-hydrogenases.

Dataset S4. Zip file containing the Cytoscape network for [Fe]-hydrogenases.