FEMS Microbiology Ecology, 2024, **100**, fiae147
DOI: 10.1093/femsec/fiae147
Advance access publication date: 30 October 2024
Research Article

Frankia [NiFe] uptake hydrogenases and genome reduction: different lineages of loss

Katharina Pawlowski[1,*], Daniel Wibberg[2], Sara Mehrabi[1,3], Nadia Binte Obaid[1], András Patyi[1], Fede Berckx[1,4], Han Nguyen[5], Michelle Hagen[1], Daniel Lundin[6], Andreas Brachmann[7], Jochen Blom[8], Aude Herrera-Belaroussi[9], Danis Abrouk[9], Petar Pujic[9], Ann-Sofi Hahlin[10], Jörn Kalinowski[2], Philippe Normand[9], Anita Sellstedt[5,*]

[1]Department of Ecology, Environment and Plant Sciences, Stockholm University, 10691 Stockholm, Sweden
[2]Center for Biotechnology (CeBiTec), Bielefeld University, 33594 Bielefeld, Germany
[3]Department of Ecology and Genetics, Evolutionary Biology, Uppsala University, 75236 Uppsala, Sweden
[4]Department of Crop Production Ecology, Swedish University of Agricultural Sciences, 75651 Uppsala, Sweden
[5]Department of Plant Physiology, UPSC, Umeå University, 90187 Umeå, Sweden
[6]Department of Biology and Environmental Science, Linnaeus University, 39182 Kalmar, Sweden
[7]Biocenter of the LMU Munich, Genetics Section, Grosshaderner Str. 2-4, D-82152 Planegg-Martinsried, Germany
[8]Bioinformatics and Systems Biology, Justus-Liebig Universität Giessen, 35392 Giessen, Germany
[9]Université de Lyon, Université Lyon 1, CNRS, UMR5557, Ecologie Microbienne, INRA, UMR 1418, 43 bd du 11 novembre 1918, 69622 Villeurbanne, France
[10]Department of Agricultural Research for Northern Sweden, Swedish University of Agricultural Sciences, 90183 Umeå, Sweden
*Corresponding authors. Katharina Pawlowski, Department of Ecology, Environment and Plant Sciences, Stockholm University, 10691 Stockholm, Sweden. E-mail: katharina.pawlowski@su.se; Anita Sellstedt, Department of Plant Physiology, UPSC, Umeå University, 90187 Umeå, Sweden. E-mail: anita.sellstedt@umu.se
Editor: [Paolina Garbeva]

Abstract

Uptake hydrogenase (Hup) recycles H_2 formed by nitrogenase during nitrogen fixation, thereby preserving energy. Among root nodule bacteria, most rhizobial strains examined are Hup^-, while only one Hup^- *Frankia* inoculum had been identified. Previous analyses had led to the identification of two different [NiFe] hydrogenase syntons. We analysed the distribution of different types of [NiFe] hydrogenase in the genomes of different *Frankia* species. Our results show that *Frankia* strains can contain four different [NiFe] hydrogenase syntons representing groups 1f, 1h, 2a, and 3b according to Søndergaard, Pedersen, and Greening (HydDB: a web tool for hydrogenase classification and analysis. *Sci Rep* 2016;6:34212. https://doi.org/10.1038/srep34212.); no more than three types were found in any individual genome. The phylogeny of the structural proteins of groups 1f, 1h, and 2a follows *Frankia* phylogeny; the phylogeny of the accessory proteins does not consistently. An analysis of different [NiFe] hydrogenase types in Actinomycetia shows that under the most parsimonious assumption, all four types were present in the ancestral *Frankia* strain. Based on Hup activities analysed and the losses of syntons in different lineages of genome reduction, we can conclude that groups 1f and 2a are involved in recycling H_2 formed by nitrogenase while group 1 h and group 3b are not.

Keywords: root nodules; actinorhiza; *Frankia*; nitrogen fixation; uptake hydrogenase

Introduction

The reduction of atmospheric dinitrogen to ammonia during biological nitrogen fixation (BNF) *via* the activity of the metalloenzyme complex nitrogenase represents one of the most important evolutionary innovations; its evolution preceded that of oxygenic photosynthesis. Phylogenetic analyses indicate that BNF most probably evolved in an anaerobic thermophilic ancestor of hydrogenotrophic methanogenic archaea and later diversified via lateral gene transfer into anaerobic, and eventually also into aerobic bacteria (Mus et al. 2019). Consistent with its origin in an anoxic environment, nitrogenase is extremely oxygen sensitive. The nitrogenase reaction involves the reductive elimination of two [Fe–H–Fe] bridging hydrides to form H_2; i.e. hydrogen evolution is part of the reduction of dinitrogen to ammonia (Khadka et al. 2017). Since hydrogen evolution represents a loss of energy, it is not surprising that many nitrogen-fixing organisms contain an uptake hydrogenase, which recycles the hydrogen produced in N_2 fixation (Schubert and Evans 1976). Hydrogen metabolism is of considerable ecological importance given the prevalence of hydrogen in the environment (Piché-Choquette and Constant 2019); the activities of H_2-producing and -consuming microbes shape the global H_2 cycle that is linked to the global cycling of other elements. Accordingly, different types of hydrogenase enzymes exist that fulfil different functions (Vignais et al. 2001, Vignais and Billaud 2007). Søndergaard et al. (2016) used the amino acid sequences to infer an expanded classification scheme predicting biological function.

In any case, the question arises why some nitrogen fixers have lost their uptake hydrogenase. This question is of particular importance for symbiotic nitrogen fixers. Two groups of nitrogen-fixing soil bacteria can enter intracellular symbioses with plants where they fix nitrogen while hosted within plant cells in special organs, the root nodules: rhizobia (alpha- or beta-*Pseudomonadota*) with legumes and one nonlegume genus, *Parasponia*, and *Actinomycetota* of the genus *Frankia* with a group of dicotyledonous plants from eight different families belonging to three orders,

Received 20 May 2024; **revised** 7 October 2024; **accepted** 29 October 2024

Fagales, Rosales, and Cucurbitales, collectively called actinorhizal plants (Pawlowski and Bisseling 1996). Many symbiotic nitrogen fixers examined can form an uptake hydrogenase (i.e. they are Hup^+), and it could be shown that the recycling of H_2 significantly increased the yield in agricultural legumes (Dixon 1972, Schubert and Evans 1976, Albrecht et al. 1979). Nevertheless, among surveyed legume symbioses Hup^+ symbioses are in the minority (Annan et al. 2012).

Actinorhizal symbioses are a more complex field. The microsymbionts, *Frankia* strains, are very diverse, with genome sizes ranging from 4.3 to 10.5 MB, and can be grouped into four phylogenetic clades called clusters, three of which are symbiotic and represent host specificity groups; the fourth cluster contains nonsymbiotic strains that cannot fix nitrogen (Normand et al. 1996, Nguyen and Pawlowski 2017). Cluster-1 strains nodulate actinorhizal Fagales (cluster-1c strains nodulate Casuarinaceae, while other cluster-1 strains nodulate actinorhizal Betulaceae and Myricaceae) except for members of two outlier genera. Cluster-3 strains nodulate most actinorhizal members within the Rosales families Elaeagnaceae and Rhamnaceae except for one outlier genus, and can nodulate the two outlier genera of the Fagales, *Gymnostoma* and *Morella*. Strains of cluster-2, the earliest divergent symbiotic cluster of *Frankia* (Sen et al. 2014, Gtari et al. 2015, Persson et al. 2015) nodulate all actinorhizal Cucurbitales, actinorhizal Rosaceae, and members of the outlier genus of the Rhamnaceae, *Ceanothus*. Cluster-2 strains can be separated in two divergent lineages, the continental lineage which encompasses strains from mainland Eurasia, Japan, and North America, and the island lineage encompassing strains from the Southern hemisphere, the Philippines, and Taiwan (Berckx et al. 2022). A phylogenetic tree including all *Frankia* species published up to the end of 2022 is shown in Fig. S1.

An early study examined different actinorhizal plants—*Alnus rubra* and *Myrica californica* (nodulated by cluster-a strains), *Purshia tridentata* (nodulated by cluster-2), and *Elaeagnus angustifolia* (nodulated by cluster-3)—and found a RE of nitrogenase between 0.97 and 0.99, indicating the activity of uptake hydrogenase (Schubert and Evans 1976). All actinorhizal nodules induced by cluster-1 strains analysed since then displayed Hup activity (Tjepkema et al. 1986, Winship et al. 1987, Pesce et al. 2017), with one exception: nodules induced on roots of *Alnus incana* (Betulaceae and Fagales) by the so-called 'local source' from Umeå in northern Sweden, an inoculum that could not be cultured (Sellstedt et al. 1986, Huss-Danell 1991). The 'local source' has another unusual feature in that here, *Frankia* can form sporangia within host plant cells, not just during saprotrophic growth. Meanwhile, more *Frankia* strains that sporulate in nodules have been found, particularly among strains that nodulate alder (Cotin-Galvan et al. 2016, Schwob et al. 2018) but also among some strains that nodulate *Myrica* sp. or *Comptonia* sp. (Kashanski and Schwintzer 1987), i.e. other host plants of cluster-1 *Frankia*. So far, no *Frankia* strains that sporulate in nodules could be cultured; sequencing showed strong genome reduction, which would be consistent with the loss of saprotrophic potential (4.3–5 MB genome size compared to 7.5–8 MB genome size in other cluster-1a *Frankia*; Herrera-Belaroussi et al. 2020, Pozzi et al. 2020).

Frankia strains contain [NiFe] hydrogenases (Mattsson et al. 2001, Richau et al. 2013), heterodimeric metalloenzymes consisting of a large subunit of ca. 60 kDa (HupL) and a small subunit of ca. 30 kDa (HupS). Six maturation proteins (HypA–HypF) are required for the biosynthesis of the $NiFe(CN)_2CO$ catalytic cluster and its insertion into HupL (Peters et al. 2015). HypE and HypF synthesize the thiocyanate, while the delivery of cyanide to iron occurs at the HypC–HypD complex. HypA and HypB are involved in acquisition and insertion of Ni^{2+} ions into HupL after the $Fe(CN)_2CO$ moiety has been added. In the final step, a peptide on the C-terminus of HupL is processed by a maturation endopeptidase (HupD), causing a conformational change which enables enzyme activity. After *Frankia* genome sequences began to become available in 2007 (Normand et al. 2007), it turned out that genomes of *Frankia* strains from clusters-1 and 3 contained two sets of uptake hydrogenase genes with associated genes of maturation proteins, organized in syntons-1 and -2 (Leul et al. 2007, 2009, Richau et al. 2013). Synton-1 was expressed at a higher level in culture, while synton-2 was expressed at higher levels in nodules (Leul et al. 2007, Richau et al. 2013). Proteins encoded by genes from synton-1 and synton-2, respectively, showed different phylogenies, which was interpreted to suggest HGT from other soil actinobacteria to *Frankia* (Leul et al. 2009).

Given that many more genomes from *Frankia* and other actinobacterial strains have been sequenced since 2009, our views on the uptake hydrogenase gene situation in *Frankia* strains should be reassessed. Moreover, the question has to be answered how hydrogenase activity was lost in one or more cluster-1a strain(s) that sporulate(s) in nodules; which genes/syntons were lost here? It also has to be considered that so far, only one cluster-3 genome and no cluster-2 genomes were analysed with regard to uptake hydrogenase syntons (Leul et al. 2007, 2009), and only nodules from one host plant of each cluster have been examined for Hup activity (Schubert and Evans 1976). What is the uptake hydrogenase gene/synton situation in *Frankia* strains from cluster-2 which, like the cluster-1a strains sporulating in nodules, show genome reduction (5–6 MB; Nguyen and Pawlowski 2017)?

Therefore, we analysed the Hup synton situation and phylogeny of the large and small subunits and several accessory proteins based on representatives of all *Frankia* species described before 2023. We also examined hydrogenase activity and transcription in nodules and, if available, cultures for a strain from *Frankia* cluster-1 (*Frankia alni* $ACN14a^T$; Normand et al. 2007), and for strains from cluster-1 that sporulate in nodules: the type strains *Candidatus* Frankia alpina AiOr and *Candidatus* Frankia nodulisporulans $AgTrS^T$ and two representative of the 'local source', the original inoculum and Ca. *F. nodulisporulans* UmASH1 (Sellstedt et al. 1986, Pozzi et al. 2020, Herrera-Belaroussi et al. 2020). Additionally, we examined a strain from the continental lineage of cluster-2 (*Candidatus* Frankia californiensis Dg2; Normand et al. 2017) and a strain from cluster-3, Hr75.2.1 that was isolated by André Moiroud in 1991 from a nodule from *Hippophae rhamnoides* growing in the Isère department in the village of Evrieu, France (Bautista-Guerrero et al. 2011), the genome of which was sequenced for this study. Sequence analysis [average nucleotide identity (ANI), core genome phylogeny] showed that Hr75.2.1 is a representative of *Frankia soli* (Gtari et al. 2020; Fig. S2). A summary of all hydrogenase syntons/types and the corresponding genes is presented in Table 1.

Experimental procedures

Plant and bacterial material and growth conditions

Seeds of *Alnus glutinosa* (L.) Gaertn were obtained from Svenska Skogsplantor (Bålsta, Sweden), while seeds of *A. incana* (L.) Moench were obtained from Levinsen (Gørløse, Denmark). Seeds were germinated on coarse-grained sand (1.2–2 mm diameter; Rådasand, Lidköping, Sweden) in a greenhouse at a light/dark rhythm of 16 h

Table 1. All hydrogenase types/syntons analysed in this study. When one gene is missing, it is mentioned by name; when more than one gene is missing, the synton is called 'incomplete'. Not shown, no sequence from this synton is discussed in this manuscript.

		NiFe hydrogenase synton/type			
Cluster	**Species, strain**	**hupL1 (group 1 h)** ***hypABCDEFhupDSL***[a]	**hupL2 (group 2a)** ***hypABCDEFGhupDSL***[b]	**hupL3 (group1f)** ***hypABCDEFhupCDSL***[a,c]	**hupL4 (group 3b)** ***hyhGBhupSLD***
Cluster-1	*Frankia alni* ACN14a	Yes	Yes	No	No
	Frankia canadensis ARgP5	Yes	Yes, *hypF2* missing	Only *hypF3*	No
	Frankia torreyi CpI1S	Yes	Yes	No	Yes
	Frankia torreyi ACN1ag	Yes (not shown)	Yes (not shown)	No	Yes
	Frankia nodulisporulans AgTrS	No	*hypF2* incomplete	No	No
	Frankia nodulisporulans UmASH1	No	*hypF2* incomplete	No	No
	Frankia nodulisporulans LS	No	*hypF2* incomplete	No	No
	Frankia alpina AiOr	Incomplete	Yes	No	No
	Frankia casuarinae CcI3	Yes	Yes	No	Yes
	Frankia casuarinae KB5	Yes (not shown)	Yes (not shown)	No	Yes
Cluster-2	*Frankia datiscae* Dg1	No	Yes	No	No
	Frankia coriariae BMG5.1	No	Yes	No	No
	Frankia californiensis Dg2	No	Yes	No	No
	Frankia meridionalis Cppng1	No	No	Yes	No
	Frankia sp. CiP3	No	No	Yes	No
Cluster-3	*Frankia elaeagni* BMG5.12	Incomplete[d]	*hypF2-hypD2-hypE2*	Yes	Yes
	Frankia discariae BCU110501	Yes, *hypF1* missing	Only *hypF2*	Yes	No
	Frankia soli NRRL B 16 219	Yes, *hypF1* missing	Only *hypF2*	Yes	Yes
	Frankia soli Hr75.2.1	Yes, *hypF1* missing	Only *hypF2*	Yes (not shown)	Yes
	Frankia soli EAN1pec	Yes, *hypF1* missing	Only *hypF2*	Yes	No
	Frankia irregularis DSM45899	yes, *hypF1* missing	Only *hypF2*	Yes	No
Cluster-4	*Frankia saprophytica* CN3	Yes	No	No	Yes
	Frankia asymbiotica NRRL B 16 386	Yes	No	No	No
	Frankia inefficax EuI1c	Yes, *hypF1* missing	No	Only *hypF3*	No
	Frankia sp. EUN1h	Yes (not shown)	No	No	Yes

[a]Two *hypC1* and *hypC3* copies in cluster-3.
[b]Two *hypC2* copies in cluster-2.
[c]*hypB3* and *hypF3* missing in cluster-3.
[d]*hypD1E1F1*, *hypA1B1* missing.

light/8 h dark, with 23°C during the light and 20°C during the dark phase. After germination, seedlings were transferred to larger pots with sand, and after 4 weeks, plantlets were transferred to experimental pots (13 cm diameter) containing a 1:1 mixture of autoclaved sand (1.2–2 mm quartz, Rådasand) and autoclaved germination soil (Såjord, Hasselfors, Sweden), and inoculum was added. Inocula were for prepared for eight plants each (four *A. glutinosa* and four *A. incana*): a culture of *Frankia alni* ACN14a grown in BAP medium without N source (Benoist et al. 1992), washed twice with sterile milliQ H_2O, and resuspended in sterile milliQ H_2O; nodules containing *Candidatus* Frankia nodulisporulans UmASH1 (Herrera-Belaroussi et al. 2020), nodules containing *Ca.* F. nodulisporulans AgTrS (Herrera-Belaroussi et al. 2020), nodules containing the original 'local source' (Sellstedt 1986) or nodules containing *Candidatus* F. alpina AiOr (Pozzi et al. 2020) crushed in 10 ml sterile milliQ H_2O per sample, respectively. The nodules representing the inocula for UmASH1, AgTrS, 'local source' and AiOr had been formed on greenhouse plants grown on autoclaved sand inoculated with surface-sterilized crushed nodules, and therefore were used without a new round of surface sterilization. Plants were watered with deionized H_2O on alternate days and with one-fourth strength Hoagland's medium (Hoagland and Arnon 1938) every 2 weeks. Before inoculation, one-fourth strength Hoagland's with 10 mM KNO_3 was used; after inoculation, one-fourth strength Hoagland's with 1 mM KNO_3 was used. Nodules were harvested for RNA isolation 6 months after inoculation; whole root systems were used for acetylene reduction assays (ARA) and H_2 evolution measurements ca. 23 weeks after infection, in case of 'local source' inoculum ca. 9 months after infection.

Datisca glomerata (C. Presl) Baill. seeds originating from plants growing at Gates Canyon in Vacaville, CA, USA, had been propagated in the greenhouse since 1997. Seeds were germinated on sand and transferred to a 1/1 (v/v) mixture of sand and germination soil (Såjord) after 3–6 weeks. Plants were nodulated with the Dg2 inoculum (Nguyen et al. 2016) 3 months after germination. Before inoculation, plants were watered with deionized H_2O on alternate days and with one-fourth Hoagland's medium with 10 mM KNO_3, after inoculation with the same medium but with 1 mM KNO_3 every 2 weeks. Temperature and light conditions were as described for alders. Plants were used for uptake hydrogenase assays 13 months after inoculation in August 2021. Nodules for RNA isolation were collected from *D. glomerata* plants inoculated with Dg2 6 months after inoculation; ARA and hydrogen evolution were measured ca. 9 months after inoculation.

Seeds of *H. rhamnoides* L. were obtained from Saflax (www.saflax.de), surface-sterilized with 3% sodium hypochlorite and 0.01% Tween 20 for 20 min, washed extensively with autoclaved milliQ H_2O, and kept in autoclaved milliQ H_2O at 4°C for 3 days. Then, they were sown onto moistened filter paper in Petri dishes covered with aluminium foil. Three weeks later, seedlings were transferred to small pots containing a mix of sterilized gravel and sand; 50% of the plants were inoculated with *Frankia* sp. Hr75.2.1 grown in PUM medium without nitrogen (Mattsson and Sellstedt

2000) and washed twice with sterile milliQ H_2O, while 50% were were left uninoculated. The plants were grown at 18 h light/6 h dark with 22°C during the light and 18°C during the dark phase, and 60% humidity. They were watered with Evan's medium with 5 mM NH_4NO_3 once per week for 4 weeks, and afterwards with Evan's medium without nitrogen once per week (Evans et al. 1972), otherwise with tap water. Nodules were harvested for RNA isolation ca. 6 months after inoculation; ARA and hydrogen evolution measurements were performed ca. 1 year after inoculation. In each case, temperatures were maintained through the entire growth period and during ARA and hydrogen evolution measurements as temperature affects the RE of ntrogenase (Walsh and Layzell 1986).

RNA isolation and reverse transcription-quantitative polymerase chain reaction

RNA was isolated from cultures and nodules as described by Nguyen et al. (2019) and Berckx et al. (2020). Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) was performed as described by Zdyb et al. (2018). *InfC* (encoding the translation initiation factor IF-3) was used as reference gene (Alloisio et al. 2010). Primers used in this study are listed in Table S1. Statistical evaluation was performed in RStudio (RStudio Team 2022) using one-way ANOVA with *post hoc* Tukey HSD test.

Analysis of *hypF2* in *Ca.* Frankia nodulisporulans AgTrS, UmASH1, and 'local source'

In order to find out whether the appearance of a broken *hypF2* gene (one operon ending with a 3′-truncated version, another operon beginning with a 5′-truncated version) was an assembly error, DNA was isolated from nodules induced by AgTrS, UmASH1, and the 'local source' on *A. glutinosa* as described by Berckx et al. (2022). Polymerase Chain Reaction (PCR) was performed for *hypF2* using DreamTaq DNA Polymerase (Thermo Fisher) and three different primer pairs from the 5′ and 3′ part (Table S1) with the program 95°C 5 min; 35 cycles of 95°C 30 s, 63°C or 64°C 30 s, 72°C 1 min; and 72°C 10 min. To further check the completeness of the *hypF2* gene, the raw assembly data in *.ace format were imported into Consed (Gordon et al. 1998). The region encoding *hypF2* was identified and manually inspected.

Hr75.2.1 and 'local source' genome sequencing and assembly

DNA was isolated from a Hr75.2.1 culture grown in basic alkaline propionate (BAP) liquid medium without nitrogen (Benoist et al. 1992) according to Nguyen et al. (2019). DNA was used for next generation sequencing (NGS) using Illumina technology. Raw sequencing data of 1.11 GB in size was obtained and used for genome assembly. The assembled genome was then annotated on the MaGe genomic Platform (https://mage.genoscope.cns.fr/microscope/home/index.php). Genome data were deposited at GenBank, BioProject PRJEB60167. The genome of Hr75.2.1 is 8.269 105 MB in size with 71.59% GC.

For sequencing the genome of the original 'local source' from Umeå (Sweden), a representative of *Candidatus* Frankia nodulisporulans (Herrera-Belaroussi et al. 2020), nodules of the original inoculum, (Sellstedt et al. 1986) were used which had been stored at −80°C. DNA was isolated and sequenced as described by Berckx et al. (2022). Genome data were deposited at NCBI, accession numbers CAWVPV010000001–CAWVPV010000515. The genome of the 'local source' (LS) is 4.238 073 MB in size with 71.68% GC.

Organism phylogeny and ANI

ANI comparisons were performed in order to determine which species the *H. rhamnoides*-nodulating strain Hr75.2.1 belonged to, using the usually applied ANI threshold range of 95% for species delineation (Goris et al. 2007), the EDGAR platform (Blom et al. 2009, 2016) and the type strains of all *Frankia* species available thus far. On the same platform, a core genome phylogenetic tree was inferred using the FastTree software (http://www.microbesonline.org/fasttree/; Price et al. 2009, 2010).

Protein phylogeny

Hyp and Hup trees were constructed as described previously with some modifications (Wibberg et al. 2020). Briefly, multiple alignments were generated for all selected protein sequences using MUSCLE (Edgar 2004). All amino acid alignments were subjected to automatic elimination of poorly aligned positions using the Castresana Lab Gblocks standalone application with default parameters (version 0.91b; Castresana 2000); the 'with-hal' gap treatment was applied and the minimum block length was set to 5 (Talavera and Castresana 2007). Phylogenetic trees based on curated alignments using the maximum likelihood method with 1000 bootstrap resampling were calculated using MEGA version 11 (Tamura et al. 2021).

Hydrogenase operon analysis

Frankia [NiFe] hydrogenase genes were identified via Blast searches on NCBI and JGI (https://img.jgi.doe.gov/cgi-bin/w/main.cgi). Plots of the hydrogenase regions were created by a modified in-house script as described before (Eikmeyer et al. 2012).

16S profiles

DNA was isolated from nodules induced by UmASH1 and the original 'local source' inoculum, respectively, on *A. glutinosa* as described by Berckx et al. (2022). DNA quality was assessed by gel electrophoresis and photometrically using a NanoDrop® ND-2000c UV–Vis spectrophotometer (NanoDrop Technologies). The V3–V4 region of the 16 s rRNA gene was amplified using barcoded versions of the universal primer set 341F (5′-CCT AYG GGR BGC ASC AG-3′) and 806R (5′-GGA CTA CNN GGG TAT CTA AT-3′). Amplicon quality was assessed by agarose gel electrophoresis prior to library preparation. Library quality was controlled using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific) and an Agilent Bioanalyzer 2100 system. Libraries were sequenced on an Illumina NovaSeq 6000 platform employing 250 bp paired-end reads at Novogene Europe (Cambridge, UK). Processing of raw data entailed: data split (paired-end reads was assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence), sequence assembly [paired-end reads were merged using FLASH (V1.2. 1 1; http://ccb.jhu.edu/software/FLASH/) (Magoč and Salzberg 2011), a highly efficient tool designed to merge paired-end reads by utilizing the overlap between reads generated from opposite ends of the same DNA fragment; this process results in the creation of sequences known as raw tags; data filtration [quality filtering on the raw tags were performed using the fastp (Version 0.23.1) software to obtain high-quality Clean Tags (Bokulich et al. 2013)]; chimera removal [the tags were compared with the reference database [Silva database (16S/18S), https://www.arb-silva.de/; Unite Database (ITS),

https://unite.ut.ee/] to detect chimera sequences, and the effective tags were obtained by removing the chimera sequences with the vsearch package (V2.16.0, https://github.com/torognes/vsearch] (Edgar et al. 2011).

Distribution of hydrogenase types in the class Actinomycetia

The hidden Markov model (HMM) of HupL was identified using the HupL1 sequence from *F. alni* ACN14a (CAJ61042.1) for a search on www.uniprot.org/blast. The raw HMM was downloaded from http://pfam.xfam.org/family/PF00374. In order to run an hmmsearch of PF00374 with HMMER (Eddy 2011), a database was created out of the amino acid fasta files of the PROKKA (version 1.14.6; Seeman 2014) annotations of all Actinobacteriota species representative genomes in the Genome Taxonomy Database (GTDB) (release R07-RS207; Parks et al. 2019). A ranking of the HMMER output per sequence was created using 'hmmrank' by Daniel Lundin (https://github.com/erikrikarddaniel/hmmrank). The best-ranked sequences were used to identify the hydrogenase type on hydDB (https://services.birc.au.dk/hyddb/) (Søndergaard et al. 2016). A set of counting tables for each taxonomic rank was created out of the hydDB output using RStudio (version 2022.07.1, R version 4.2.0; RStudio Team 2022) with the tidyverse package (https://www.tidyverse.org/) (Wickham et al. 2019).

In order to create a family tree for all Actinomycetia, the bacterial GTDB reference tree was used (https://data.gtdb.ecogenomic.org/releases/release207/207.0/bac120_r207.tree), subset to c_Actinomycetia and pruned to family level using Dendroscope (version 3.8.3) (https://uni-tuebingen.de/fakultaeten/mathematisch-naturwissenschaftliche-fakultaet/fachbereiche/informatik/lehrstuehle/algorithms-in-bioinformatics/software/dendroscope/) (Huson and Scornavacca 2012).

Determination of nitrogenase and uptake hydrogenase activities

Roots of intact whole-plant systems were incubated in gas-tight cuvettes, which were sealed with suba seals (Sigma-Aldrich, Sweden) and left in the greenhouse for 1 h. Measurement of hydrogen was performed by taking out 1 ml samples from each cuvette with a gas-tight syringe and injecting them into a GC-8AIT gas chromatograph with a TCD detector (Thermal Conductivity Detector; Shimadzu Scientific Instruments, Columbia, USA) (Mattsson and Sellstedt 2000). The TCD-GC was equipped with a column packed with Molecular Sieve 5A with argon as carrier gas, and run at 80°C column temperature and 120°C detector temperature. In the cases where hydrogen evolution could not be detected after 1 h, a second measurement was made after 24 h.

For measurements of nitrogenase activities, the ARA was used. Plants roots of intact whole-plant systems incubated in the above-mentioned gas-tight cuvettes were vented for 2 h; then, 10% (v/v) of the gas phase in the cuvette were exchanged with 99.5% (v/v) acetylene; then the cuvettes were kept in the greenhouse for 1 h. Measurement of ethylene was performed by taking 1 ml samples from each cuvette with a gas-tight syringe and injecting into a GC-8AIT gas chromatograph with a FID detector (Flame Ionization Detector; Shimadzu Scientific Instruments) (Mattsson and Sellstedt 2000, Richau et al. 2017). The FID-GC was equipped with a column packed with Porapak T, and run at 90°C column temperature and 220°C detector temperature, with N_2 as carrier gas.

The relative efficiency (RE) of nitrogenase is used to describe the extent of hydrogen loss during N_2 fixation (Schubert and Evans 1976). RE is a coefficient describing the electron allocation to nitrogenase. Divergence of the coefficient indicates the presence of other interacting enzyme systems, such as for example uptake hydrogenase.

$$RE = 1 - (\text{rate of } H_2 \text{ evolution in air}) / (\text{rate of } C_2H_2 \text{ reduction}).$$

According to Burns and Hardy (1975) and Schubert and Evans (1976), the rate of hydrogen evolution in the absence of other reducible substrates should depict the total electron flux to the nitrogenase system. Therefore, in the presence of acetylene at a saturating amount of 10% (v/v), no hydrogen is formed and the total electron flow to nitrogenase is most likely used to reduce acetylene (Dart and Day 1971). Consequently, the RE of electron transfer to nitrogen via nitrogenase can be expressed by the ratio defined the RE, and an RE above 0.75 indicates uptake hydrogenase activity (Schubert and Evans 1976).

Results and discussion

Four different types of [NiFe] hydrogenase can be found in the *Frankia* strain genomes

[NiFe] hydrogenases can be grouped into 29 classes based on the amino acid sequences of their large subunits (Søndergaard et al. 2016). Using BlastP searches on NCBI with different large subunits of hydrogenases from Streptomycetes, two groups of HupL proteins were found to be encoded by *Frankia* genomes, a large group of sequences of more than 500 amino acids, and a smaller group of proteins of ca. 430 amino acids. The proteins representing the 500+ amino acid large subunit of [NiFe] uptake hydrogenase, HupL, were used to build a phylogenetic tree. The results (Fig. 1A) clearly show three distinct forms of HupL, corresponding to HupL1 and HupL2 from the two syntons previously identified in representatives of *Frankia* cluster 1, and one form that was previously identified in a strain from cluster-3 but not distinguished from HupL2 (Leul et al. 2009). The latter, now termed HupL3, was found in all genomes of cluster-3 strains examined, in one strain of the nonsymbiotic cluster-4 and in all strains from the island lineage of cluster-2 (Nguyen et al. 2019, Berckx et al. 2022, Fig. 1A). The phylogeny of the proteins representing the small submit of [NiFe] uptake hydrogenases, HupS, was similar to that of HupL (Fig. S3).

HupL1 represents a ca. 595 amino acid large subunit of [NiFe] hydrogenase group 1 h (Søndergaard et al. 2016). HupL2 represents a ca. 535 amino acid large subunit of [NiFe] hydrogenase of group 2a, and HupL3 a ca. 573 amino acid large subunit of group 1f. All three types—1 h, 2a, and 1f—are classified as 'plasma membrane associated?' in HydDB (https://services.birc.au.dk/hyddb/). Also the small subunits (HupS) encoded by the three different syntons had distinct sizes: 350–360 amino acids for HupS from synton-1, 320–323 amino acids for the version from synton-2, and 410–420 amino acids for the version from synton-3 (Table S2).

The smaller version of HupL (409–451 amino acids), termed HupL4 in this manuscript and present in several strains, represents a type 3b cytosolic [NiFe] hydrogenase (Søndergaard et al. 2016; Table S2).

The phylogeny of the hydrogenase structural proteins of syntons 1, 2, and 3 follows that of *Frankia*

Within the *Frankia* genus, cluster-1 representatives contain synton-1 and synton-2, cluster-3 representatives contain synton-1 and synton-3, and the representatives of the earliest branching clade, cluster-2, either contain synton-1 (continental lineage) or

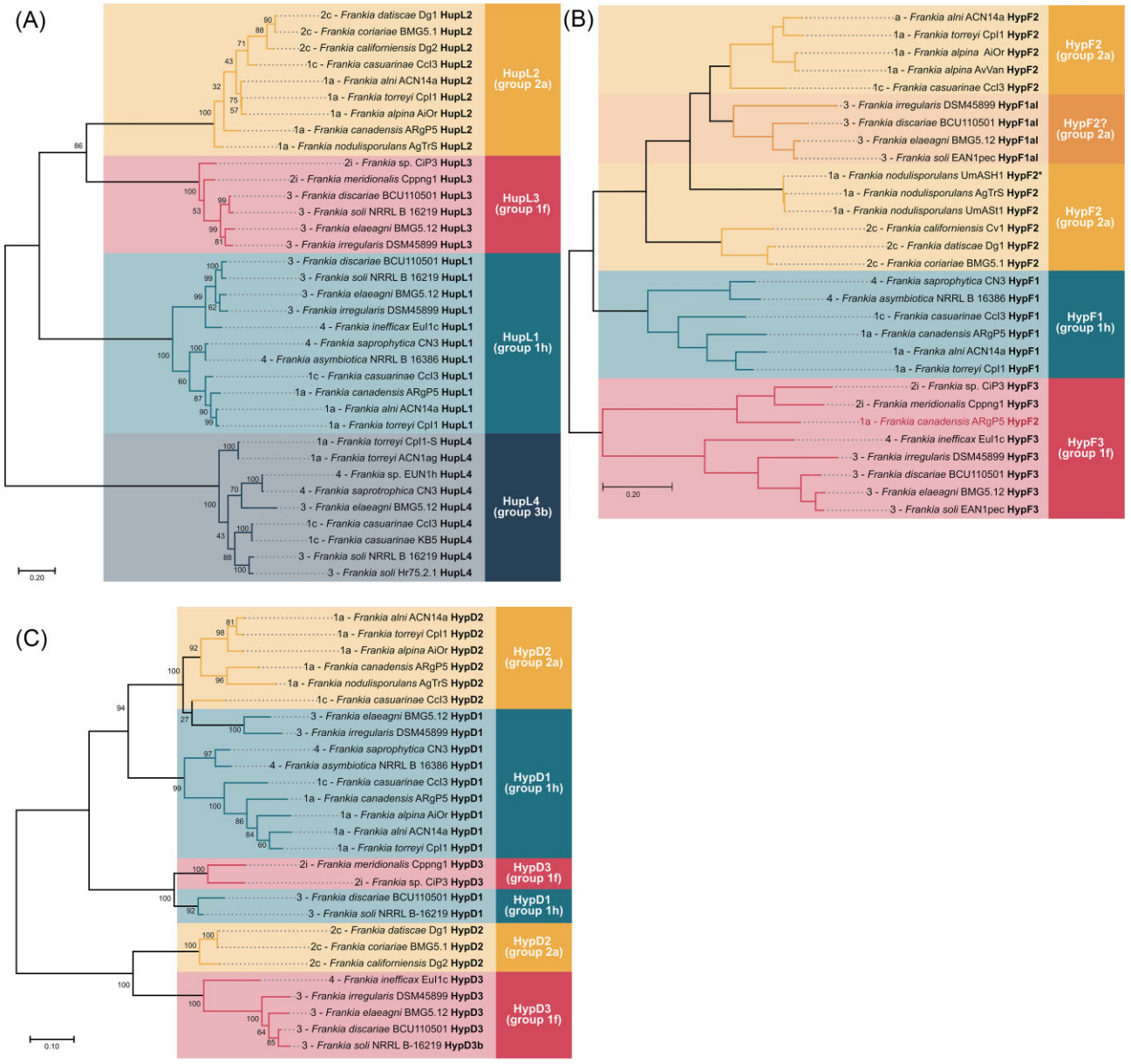

Figure 1. Phylogenies of [NiFe] hydrogenase large subunit and accessory proteins. Colouring is according to synton (synton-1, group 1 h; synton-2, group 2a; synton-3, group 1f; and synton-4, group 3b). (A) *Frankia* hydrogenase large subunit phylogeny. HupL proteins of *Frankia* species fall in four phylogenetic groups, HupL1 (synton-1), HupL2 (synton-2), HupL3 (synton-3), and HupL4 (synton-4). Numbering of synton-1 and synton-2 is based on the previously identified hydrogenase syntons published by Leul et al. (2009). (B) Phylogeny of HypF proteins of *Frankia* species. Protein numbering is based the HupL protein of the same synton; the genes encoding the HypF1al proteins from cluster-3 species are not linked to any synton in the genome (al, alone); the numbers were assigned because these genomes contain synton-1 without a *hypF* gene (and synton-3 with a *hypF* gene), nevertheless, they group with synton-2 HypF proteins. The numbering of *Frankia inefficax* Eul1c HypF3 is doubtful in so far that the corresponding gene is not linked to the only synton present in the genome, synton-1), but based on sequence, it represents a member of synton-3. *Frankia canadensis* ARg5P HypF2 was annotated as such since ARg5P contains synton-1 (which contains *hypF1*) and synton-2 (which does not contain *hypF2*); '*hypF2*' is not linked to either synton and the phylogenetic tree clearly shows it encodes a HypF3 protein. In this figure, *F. soli* EAN1pec was used instead of the type strain NRRL B-19216 because the latter contains two additional copies of *hypF3* (unlinked to any *hup* operon), which is atypical. (C) Phylogeny of HypD proteins of *Frankia* species. Protein numbering is based on the HupL protein of the same synton. The HypD3 (synton-3, group 1f) proteins of the island lineage of cluster-2 clearly represent synton-1 (group 1 h) proteins based on their sequence. The size bar shows the number of changes. Bootstrap values are included. Please note that *F. alpina*, *F. nodulisporulans*, *F. datiscae*, *F. californiensis*, and *Frankia* sp. (novel species described by Berckx et al. 2022) are *Candidatus* species.

synton-3 (island lineage) [Fig. 1A; see Berckx et al. (2022) for the two lineages of *Frankia* cluster-2]. The corresponding genes could have evolved by duplication and modification within *Frankia* or could have been acquired by horizontal gene transfer (HGT). Based on the comparison of the phylogeny of HupL1-4 (Fig. 1A) and HupS1-4 (Fig. S3) and the core genome phylogeny of the *Frankia* genus (Fig. S1), no HGT needs to be invoked; HupL1-3 and HupS1-3 clearly coevolved and their phylogeny more or less follows the core genome phylogeny.

Synton-4 (group 3b) was not found in any sequenced genome of *Frankia* cluster-2. It was found in two strains of the nonsymbiotic *Frankia* clade, cluster-4, *Frankia saprophytica* CN3 and *Frankia*

sp. EUN1h; however, two genes in the EUN1h operon are pseudogenized, so this strain cannot form a group 3b [NiFe] hydrogenase (Table S2). Synton-4 was found in some representatives of cluster-3, namely in *Frankia elaeagni* BMG5.12 and in some strains of *F. soli*, the type strain NRRL B-16219 (but two genes in the operon are pseudogenes, so a type 3b hydrogenase cannot be formed) and in Hr75.2.1 (operon intact), but not in EAN1pec (for the representatives of *F. soli*, see Fig. S2). In cluster-1, it was found among the alder-infective strains in both representatives of *Frankia torreyi*, but not at all in *F. alni*, *Ca.* Frankia alpina or *Ca.* Frankia nodulisporulans. However, it is present—and intact—in all 13 genomes available from *Casuarina*-infective strains (shown for the type strain CcI3 and for strain KB5 in Fig. 1A; see also Table S2). The phylogeny of the two subunits of group 3b [NiFe] hydrogenase, HupL4 and HupS4, does not follow *Frankia* species phylogeny (Fig. 1A; Fig. S3).

The phylogeny of the maturation proteins HypD and HypE, and of HypF, differs from the phylogeny of the structural hydrogenase proteins

The fact that both lineages of sporulating strains of cluster-1a have retained genes from different syntons, combined with the fact that synton-3 lacks *hypB3* in most cases examined, raises the question whether the phylogeny of the maturation protein genes follows the phylogeny of the structural genes, or whether we can find evidence for the exchange of maturation proteins between syntons. Therefore, the phylogeny of the largest maturation proteins, HypD, HypE, and HypF, was examined. For HypF, phylogeny clearly separated the HypF proteins of the different syntons, and within these group mostly followed *Frankia* phylogeny (Fig. 1B). However, the position of cluster-3 HypF proteins was one interesting exception. While syntons-3 of cluster-3 contain a *hypF* gene, syntons-1 of this cluster do not. However, each genome contains a *hypF* gene that is unlinked to any synton, termed '*hypF1alone*'. The corresponding proteins (HypF1al) group with the HypF proteins from synton-2 (Fig. 1B), suggesting that they represent leftovers of a lost synton-2. Furthermore, the HypF2 protein of the cluster-1 strain *Frankia canadensis* ARgP5—encoded by a gene that is not linked to synton-2, but synton-1 contains a hypF gene and synton-2 does not—clearly represents a HypF3 protein (Fig. 1B); thus, it could be a leftover of a lost synton-3. Similarly, the cluster-4 strain *Frankia inefficax* EuI1c contains a synton-1 (group 1 h) hydrogenase and a synton-3 (group 1f) HypF protein, possibly a remnant of a lost synton-3.

HypD and HypE clearly coevolved (Fig. 1C and Fig. S4). They can be grouped into five different clades, which neither represent genes from the same synton, nor follow *Frankia* phylogeny (A–E; Fig. 1C for HypD; for HypE see Fig. S4). The synton-2 HypD/E proteins of cluster-1 and of the continental lineage of cluster-2 are separated, the latter group with synton-3 HypD/E proteins of cluster-3 and of the cluster-4 species *F. inefficax* EuI1c. The synton-3 HypD/E proteins of cluster-3 and of the island lineage of cluster-2 are separated in that the latter group with synton-1 HypD/E proteins of two of the cluster-3 species examined. Thus, the phylogenetic analysis suggests that HypDE of the continental lineage of cluster-2 are remnants of a lost synton-3, and that HypDE of the island lineage of cluster-2 are remnants of a lost synton-2.

Altogether, these data show that the hydrogenase maturation protein genes evolved independently of the corresponding structural protein genes, and can be exchanged between syntons 1–3/[NiFe] hydrogenase groups 1 h, 1f and 2a. Given that HypE and HypF interact during hydrogenase maturation, while HypD interacts with HypC (Peters et al. 2015), it is surprising that HypE and HypD coevoled while HypF seems to have evolved independently (since HypC proteins of groups 1 h, 1f, and 2a are less than 100 amino acids in length, their phylogeny was not analysed).

Organization of the four different [NiFe] uptake hydrogenase syntons in the genus *Frankia*

The configuration of the three large syntons differs (Fig. 2A). Two versions of synton-1 (group 1 h) were found. In synton-1 of cluster-1 strains, all genes have the same orientation. It starts with *hypA1–hypB1*, which may represent a separate operon, followed by an operon encompassing a gene encoding a hypothetical protein of unknown function, the maturation protease (maturase) gene, *hupS1–hupL1*, a gene encoding the FeS cluster biogenesis protein NfuA, followed by three genes encoding hypothetical proteins and ending with *hypF1–hypC1–hypD1–hypE1*. The *nfuA* gene is missing in the genome of *Frankia casuarinae* CcI3. In contrast, synton-1 of cluster-3 strains begins with *hypC1a*, followed by *hupS1*, *hupL1*, four ORFs encoding hypothetical protein, the maturation protease gene, *hypC1b*, *hypD1*, *hypF1*, *hypA1*, *hypB1*, and then *hypF1* but in the opposite orientation. The cluster-4 strain *F. inefficax* EuI1c has a similar configuration of synton-1, but with *hypB1* and *hypF1* missing (data not shown); the cluster-4 strains *Frankia saprotrophica* CN3 and *Frankia asymbiotica* NRRL B-16386 have the cluster-1 configuration of synton-1.

Synton-2 (group 2a) consists of two operons in divergent orientation, one representing *hypA2–hypB2* and the large one representing *hupS2–hupL2–hupD2*, followed by a gene encoding a hypothetical protein and a gene encoding an NHL repeat containing protein, followed by *hypF2–hypC2–hypD2–hypE2*. In cluster-2 strains of the continental lineage, the structural genes are not part of the larger operon, and the arrangement of genes downstream of *hupD2* is different: *hypF2–hypC2b*, followed by a hypothetical protein gene and *gmhA*, then *hypE2–hypC2a–hypD2*. The *gmhA* gene found between *hypC2b* and *hypE2* (FsymDg_2476 in *Candidatus* Frankia datiscae Dg1; see Fig. 2A) encodes a D-glycero-beta-D-manno-heptose 1-phosphate adenylyltransferase, an enzyme involved in the synthesis of the lipopolysaccharide precursor ADP-L-glycero-β-D-manno-heptose (ADP-L-β-D-heptose) in Gram-negative bacteria and GDP-D-α-D-heptose in Gram-positive bacteria (Valvano et al. 2002, Kneidinger et al. 2002, Karan et al. 2020). *GmhA* is often found in group 1 h and group 2a hydrogenase operons (Islam 2020); however, *gmhA* is not present in synton-2 of cluster-1 *Frankia* strains nor in synton-1 (group 1 h) of any *Frankia* cluster.

Also synton-3 was represented by two versions. The first version of synton-3 was found in cluster-3 genomes. It consists of one large operon, *hypC3b–hupS3–hypL3*, followed by four genes for hypothetical proteins, followed by the maturation protease gene and then by *hypC3a–hypE3–hypD3–hypA3* (and in *Frankia irregularis*, by *hypB3* at the end; in the other cluster-3 species, *hypB3* was missing). The second version was found in the island lineage of cluster-2. Here, synton-3 consists of two operons in divergent orientation, one encompassing the maturation protein genes *hypC3–hypD3–hypE3–hypF3* and the other containing *hupS3–hupL3*, a gene encoding a [NiFe] hydrogenase cytochrome subunit (*hupC*), a gene encoding the maturation endopeptidase (maturase) *hupD*, and *hypA3*.

In summary, the operon organization and the order of the maturation protein genes *hypC*, *hypD*, *hypE*, and *hypF* differ between the three syntons. No *hypB3* gene could be identified in synton-3 in most of the cluster-3 genomes available, while the genomes of

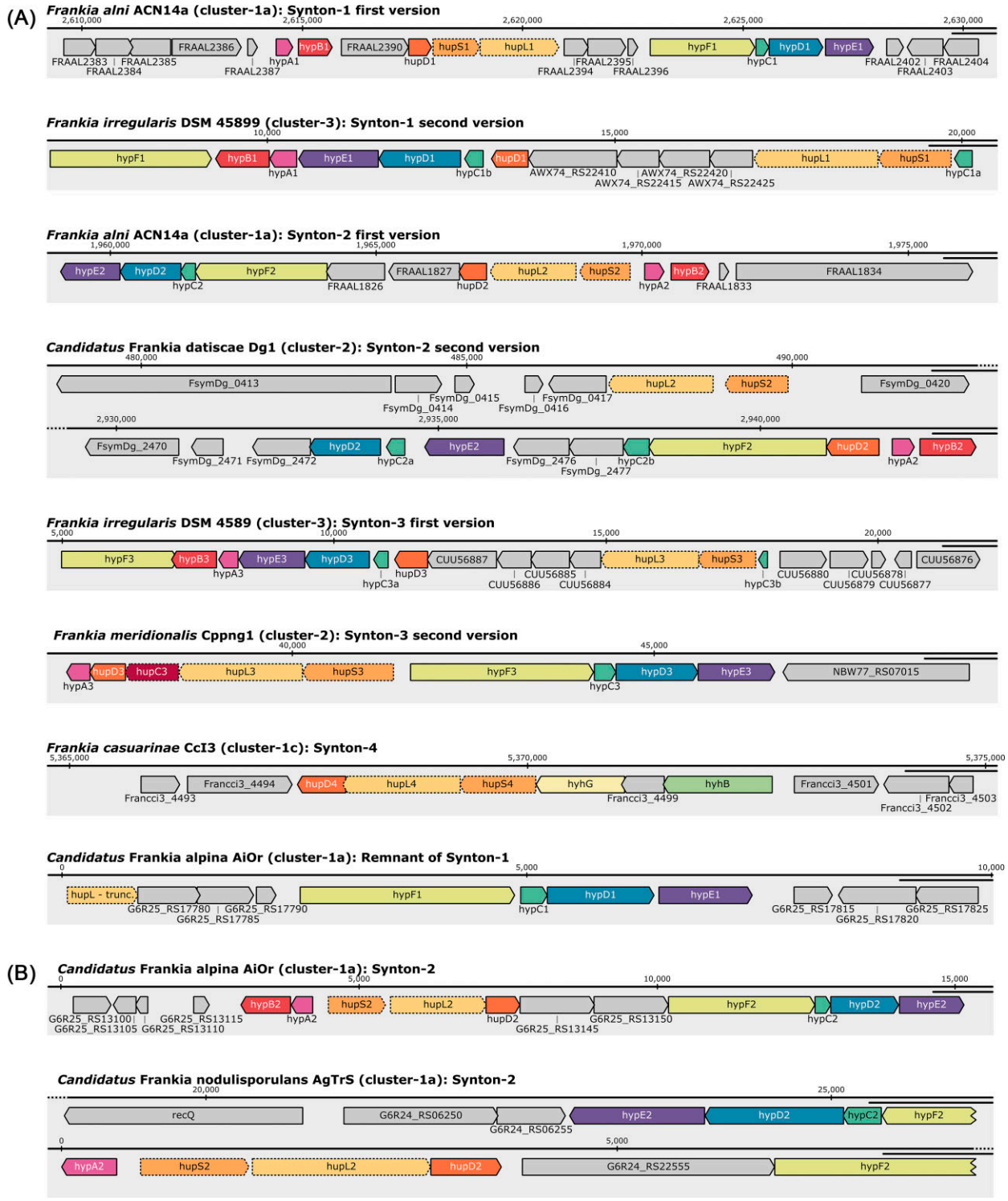

Figure 2. Complete operons of the four syntons and erosion of syntons-1 and -2. (A) Operon structure of the four different syntons in *Frankia* spp. Genes with known functions with regard to [NiFe] hydrogenase are depicted in the same colour for each synton. Genes encoding structural proteins are surrounded by dotted lines. Syntons-1, -2, and -3 have different configurations in different *Frankia* clusters. (B) Operon structure of eroding syntons in cluster-1a and cluster-2. The remnant of synton-1 in *Candidatus* Frankia alpina AiOr and its superficially intact synton-2 are shown. The two parts of synton-2 of *Candidatus* Frankia nodulisporulans AgTrS with the break in *hypF2* show that no Hup activity can be expected in this species. Erosion of synton-2 was most advanced in the 'local source' inoculum stored since 1986. The directly sequenced 'local source' inoculum (*Ca*. Frankia nodulisporulans LS) showed even further erosion of synton-2 in that the 3′-part of *hypF2* and the 5′- and 3′-end of *hypE* were missing (data not shown). Size bars denote 1 kb.

cluster-2 strains of the island lineage (Berckx et al. 2022), which only contain synton-3, lack both *hypA3* and *hypB3*.

Synton-4 did not contain any maturation protein genes besides the one encoding the maturation protease. It consists, as described in HydDB (Søndergaard et al. 2016), of the genes of the diphorase module [*hyhG*, encoding the iron–sulfur binding subunit (*F. casuarinae* CcI3: Francci3_4500), followed by *hyhB* (Francci3_4498), followed by the genes encoding the hydrogenase module with the [NiFe] center, *hyhS* (*hupS4*; Francci3_4497) encoding the small subunit and *hyhL* (*hupL4*; Francci3_4496) encoding the large subunit, followed by the maturation protease gene (Francci3_4495)]. However, in *Frankia*, between the two genes for the diaphorase module the gene for a small transcriptional regulator with a cyclic nucleotide binding domain (Francci3_4499) is present. In *F. casuarinae*, synton-4 is closely linked (ca. 7 kb distance) to the *nifHDK* operon.

Genome analysis shows loss of hydrogenase-related genes in cluster-1a spore formers and in all cluster-2 strains

Figure 2(B) shows the configuration of the four different syntons in lineages of genome reduction. In representatives of the continental lineage of cluster-2, *Frankia coriariae* and *Ca*. Frankia datiscae, synton-2 had been retained but is divided into two operons; the structural protein genes *hupS2–hupL2,* followed by the ferredoxin subunit gene, are not on the same operon as the maturation protein genes; instead, *HupS2–hupL2* had been transferred to another position in the chromosome. In *Ca*. F. datiscae Dg1, the only completely assembled genome of a cluster-2 strain, the maturation protein gene operon is located in position 2 941 034–2 934 475, while the structural protein gene operon in 488 967–488 785. The structural protein genes are followed by a gene encoding the ferredoxin subunit of nitrite reductase (F_symDg_0417 in Fig. 2B). In representatives of the island lineage of cluster-2 (Berckx et al. 2022), neither synton-1 nor synton-2 was found. Instead, synton-3 is present, but is lacking *hypB3*. While *hypB3* is also missing in the genomes of most cluster-3 strains examined, cluster-3 strains contain a complete synton-1 including *hypB1*.

In the two species of cluster-1 strains sporulating in nodules, *Ca*. F. alpina and *Ca*. F. nodulisporulans, (most of) synton-1 could not be found (Fig. 2B). *Ca*. F. alpina AiOr has retained the complete synton-2, i.e. the *hypA2–hypB2* operon and a larger operon representing *hupS2–hupL2*-maturation protein gene-two genes encoding hypothetical proteins-*hypF2–hypC2–hypD2–hypE2*. The AiOr genome also contains part of synton-1: it begins with a truncated version of *hupL1*, followed by four genes encoding hypothetical proteins, then *hypF1*, *hypC1*, a complete and a truncated version of *hypD1*, and *hypE1*. In the four representatives of *Ca*. F. nodulisporulans the genomes of which were sequenced, no remnants of synton-1 were found. The genome of the type strain AgTrS contains a truncated version of synton-2 consisting of *hypA2* and *hupS2–hupL2–hupD2* (maturation protease gene)-hypothetical protein gene, ending with a 3′-truncated version of *hypF2*. Another operon, *hypF2′–hypC2–hypD2–hypE2,* was also found, as was a *hypB2* gene. The genomes of the two representatives of the Swedish 'local source', strains AgUmASt1 and AgUmASH1, contained a complete *hypA2–hypB2* operon, while the larger operon was distributed over two contigs as in AgTrS. The genome of AgUmASH1 contains a hypothetical protein gene after the NHL repeat protein gene in the larger operon (Herrera-Belaroussi et al. 2020). Interestingly, the erosion of synton-2 is most advanced in the genome sequenced from nodules of the original 'local source' that were stored at −80°C since the 1980s. Here, only a truncated version of *hypE2* is left of the *hypF2′–hypC2–hypD2–hypE2* operon present in the three other genomes of *Ca*. F. nodulisporulans strains (Fig. 2B).

In summary, the cluster-1 strains sporulating in nodules have lost all or most of synton-1; all members of *Ca*. F. nodulisporulans have lost all of synton-1. Furthermore, the genomes of all four representatives of *Ca*. F. nodulisporulans contain no intact *hypF*; instead, AgTrS, UmASH1, and UmASt1 contained two truncated copies of *hypF2*, a 3′-truncated one in the operon with the structural genes, and a 5′-truncated one in the operon with the maturation protein genes. This should exclude any [NiFe] uptake hydrogenase activity. In order to check whether an assembly error had occurred and the strains actually contained a complete copy of *hypF2*, RT-PCR was performed with a 5′-primer derived from the 3′-truncated version and a 3′-primer derived from the 5′-truncated version of the gene. However, no product was obtained (data not shown). Lastly, the raw data of the AgTrS genome were analysed for the presence of overlapping reads between the two parts of *hypF2*, but none were found; on the contrary, the raw data showed that there were at least 300 bp of repetitive sequences between the two parts of *hypF2* (data not shown). Thus, *Ca*. F. nodulisporulans should not be able to form an active [NiFe] hydrogenase, while *Ca*. F. alpina might.

Exchangeability of maturation proteins between groups of [NiFe] hydrogenases

[NiFe] hydrogenase group 1f (synton-3) is consistently lacking *hypB* in all *Frankia* strains studied, including in cluster-3 strains of the island lineage where synton-3 encodes the only [NiFe] hydrogenase. It has been known for some time that accessory proteins for the incorporation of Ni can be exchanged between different systems, e.g. de Reuse et al. (2013) summarized that in *Heliobacter pylori* HypA/HypB contribute to the incorporation of Ni in urease, another enzyme that contains Ni as the active centre. Thus, the lack of a hypB gene in all but one representatives of synton-3 strongly suggests that HypB can be substituted for by the corresponding Ni incorporation protein from the other Hup synton (synton-1, group 1 h). However, that does not explain whether, and if yes, how the lack of *hypB3* in cluster-2 strains of the island lineage is compensated for.

Phylogenetic analysis of HypF proteins showed that in contrast with the structural proteins of [NiFe] hydrogenase, HypF phylogeny did not always follow *Frankia* phylogeny (Fig. 1B). HypF phylogeny did, however, show a separation between HypF3 and HypF1/2, with the exception that *Frankia canadensis* ARgP5 *hypF2*, which is not part of any Hup synton and was termed *hypF2* since synton-1 of ARgP5 does contain a *hypF* gene while synton-2 does not, grouped with HypF3. Furthermore, extant cluster-3 strains do not contain synton-2 (group 2a), but their syntons-1 lack *hypF1,* while their genomes contain unlinked *hypF* genes that encode HypF2 proteins (Fig. 1B). These data not only support a scenario where the ancestral *Frankia* strain had all three [NiFe] uptake hydrogenase syntons (groups 1f, 1 h, and 2a), but also suggest that accessory proteins can be exchanged between [NiFe] hydrogenases of different groups.

How did *Frankia* end up with four types of hydrogenases: evolution of hydrogenases in Actinomycetia

Phylogenetic data indicate that BNF evolved in an anaerobic, thermophilic ancestor of hydrogenotrophic methanogens and later

spread *via* lateral gene transfer (HGT) into anaerobic bacteria, and eventually into aerobic bacteria including Cyanobacteria (Mus et al. 2019). At one point, the ancestor of *Frankia* must have acquired nitrogenase genes and thus the capacity to fix nitrogen while producing hydrogen gas as a by-product. Thus, the ancestor of *Frankia* could profit from the acquisition of uptake hydrogenase: first, to reabsorb the H_2 produced by nitrogenase, second, to produce ATP via the oxyhydrogen (Knallgas) reaction; third, to remove oxygen from the oxygen-sensitive nitrogenase complex; and fourth, to provide reduction equivalents (Zhang et al. 2014). Or did the *Frankia* ancestor already contain hydrogenase genes, and if yes, for which type(s) of hydrogenase? In this context, it is interesting that in strains of the nonsymbiotic *Frankia* cluster-4, nitrogen fixation (*nif*) genes are missing while genes encoding [NiFe] uptake hydrogenase group 1 h (synton-1) are commonly found. Therefore, we set out to study the occurrence of different types of hydrogenase in the phylogenetic environment of *Frankia*.

According to the GTDB taxonomy (Chaumeil et al. 2019, Parks et al. 2022), Frankiaceae are part of the order Mycobacteriales in the class Actinomycetia in the phylum Actinobacteriota. We analysed the presence of different types of [NiFe] hydrogenases in the families that make up the Actinomycetia class (Fig. 3). It is obvious that several types of [NiFe] hydrogenases are present in different families of the Actinomycetia, specifically of the Mycobacteriales. Within the Mycobacteriales, two other families besides Frankiaceae contain the four types of hydrogenases present in *Frankia* spp.; these families (Mycobacteriaceae and Pseudonocardiaceae) contain also hydrogenases from two (Pseudonocardiaceae) or three (Mycobacteriaceae) other groups. Within the entire class Actinomycetia, altogether four families (Streptosporangiaceae, Streptomycetaceae, Mycobacteriaceae, and Pseudonocardiaceae) contain hydrogenases of (at least) the same four groups that *Frankia* spp. contain.

This means it is parsimonious to assume that the *Frankia* ancestor contained different types of [NiFe] hydrogenases. Furthermore, the phylogeny of the maturation protein HypF in *Frankia* shows evidence that the common ancestor of the four *Frankia* clades contained syntons 1–4, i.e. hydrogenases of the groups 1 h, 2a, 1f, and 3b, and that at least some maturation proteins can be exchanged between these different groups of [NiFe] hydrogenases. The exchangeability of maturation proteins between the groups 1 h, 2a, and 1f can be extended to group 3b hydrogenases (synton-4), given that synton-4 does not contain any genes encoding the proteins required for the insertion of the $Fe(CN)_2CO$ cofactor, or Ni, in the large subunit of [NiFe] hydrogenases. In this context, it is interesting that so far, synton-4 was never the only hydrogenase synton found in a *Frankia* genome.

Altogether, the assumption that the common ancestor of *Frankia* contained all four [NiFe] hydrogenase syntons (groups 1f, 1 h, 2a, and 3b) found in different extant strains, seems the most parsimonious.

Acetylene reduction and uptake hydrogenase activity in intact nodules

Uptake hydrogenase activity can be deduced when nitrogenase activity, as determined by the ARA, is compared with hydrogen evolution (Schubert and Evans 1976, Sellstedt et al. 1986). The results of measurements performed on all symbiotic systems covered in this study are shown in Table 2(A). Mature nodules of *D. glomerata* infected by *Ca.* F. californiensis Dg2, and of *H. rhamnoides* infected by *F. soli* Hr75.2.1, showed Hup activity in that no hydrogen evolution could be detected while nitrogen fixation took place. The data depicted in Table 2(B) show that *F. soli* Hr75.2.1 displayed uptake hydrogenase activity also during nitrogen fixation under aerobic conditions in culture.

Nodules of *A. glutinosa* or *A. incana* induced by *F. alni* ACN14a also showed no detectable H_2 production, i.e. due to Hup activity. In contrast, alder nodules induced by *Ca.* F. alpina AiOr or one of the three representatives of *Ca.* F. nodulisporulans used in this study, showed hydrogen evolution during nitrogen fixation. In nodules induced by *Ca.* F. alpina AiOr, which has retained the complete synton-2, the RE of nitrogenase was 0.60–0.94. Based on the stoichiometry of the nitrogenase reaction and the RE calculations, a RE below 0.75 means the strain lacks uptake hydrogenase activity (Schubert and Evans 1976, Sellstedt et al. 1986, Huss-Danell 1991). Thus, while no synton-2 gene was lost in AiOr, a promoter or one or more of the genes must have acquired mutations; the strain has lost synton-1 (group 1 h) and is in the process of losing synton-2 (group 2a).

Nodules induced by the type strain of *Ca.* F. nodulisporulans, AgTrS, displayed an RE of 0.70–0.76, i.e. barely any uptake hydrogenase activity. Nodules induced by the two representatives of the 'local source', the original inoculum and UmASH1, showed RE values of 0.56–0.77 and 0.26–0.47, respectively (Table 2A). In the original study, nodules induced by the 'local source' had shown an RE of 0.57, i.e. they did not display uptake hydrogenase activity (Sellstedt et al. 1986). In a later study by Huss-Danell (1991), several species of *Alnus* sp. were inoculated with the 'local source' of *Frankia,* and the RE ranged between 0.48 and 0.75. Thus, the RE values measured for the two representatives of the 'local source' analysed are consistent with previous publications for the 'local source' of *Frankia*, and all representatives of *Ca.* F. nodulisporulans lack Hup activity, consistent with the fact that based on the genome sequence, all representatives of *Ca. F.* nodulisporulans should be equally lacking in [NiFe] hydrogenase activity since they do not contain an intact *hypF* gene. However, the differences between the RE values of the different intrastrains are surprising.

The 'local source' (*Ca.* F. nodulisporulans) is represented by four inocula in this study which behave differently with regard to nitrogenase RE and nodule morphology

The 'local source' inoculum, which recently was shown to be a member of *Ca.* Frankia nodulisporulans (Herrera-Belaroussi et al. 2020) has been maintained since the 1980s in the form of nodules stored at −80°C. Occasionally, fractions were used to inoculate plants growing on a mix of autoclaved soil and autoclaved gravel, like the plant which was inoculated in 2014 and maintained in a growth chamber, from which the nodules were harvested in 2019 that were used to isolate and sequence the UmASH1 genome (Herrera-Belaroussi et al. 2020). Later, nodules were collected from the same plant and propagated on an *A. glutinosa* plant to be used as inoculum in this study. The differences between the genomes of the 'local source', UmASH1, UmASt1, and the type strain AgTrS are consistent with the observation made earlier for the equally uncultivable cluster-2 strains, namely that inocula represent mixtures of strains and different representatives will end up inducing nodules (Berckx et al. 2022). Thus, it cannot be excluded that the different inocula contained a minor contribution of strains with a complete synton-2, leading to the occasional occurrence of nodule lobes containing *Frankia* with uptake hydrogenase activity.

In independent experiments performed in Umeå and Stockholm, nodules induced by UmASH1 on roots of *A. glutinosa* and *A. incana* consistently did not resemble those induced by the 'local

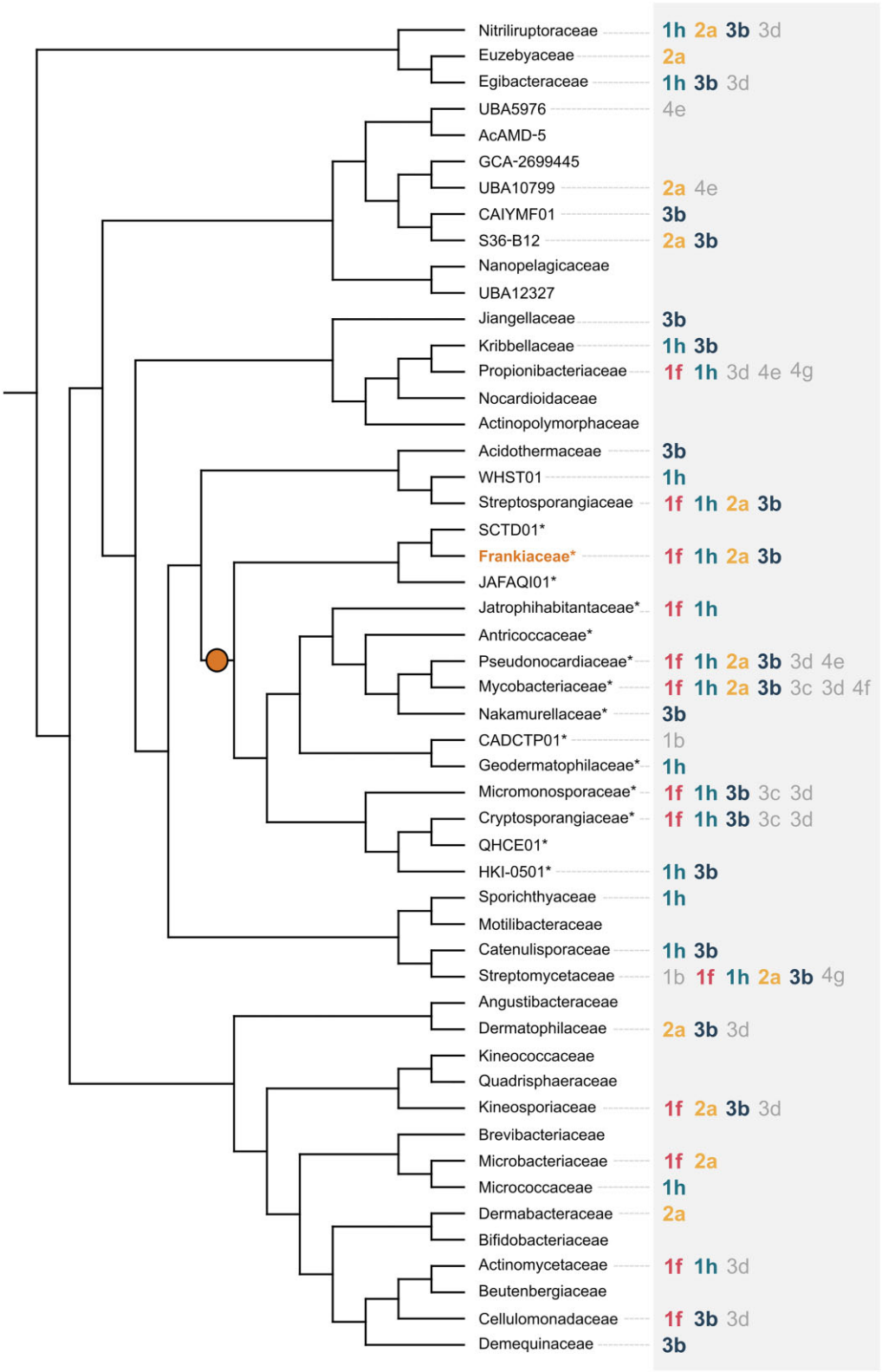

Figure 3. Distribution of [NiFe] hydrogenase types in Actinomycetia. To see whether the occurrence of four different types of hydrogenase in *Frankia* spp. is likely to have involved HGT, the occurrence of all hydrogenase types was examined using the HydDB tool (Søndergaard et al. 2016) in the entire class Actinomycetia. The phylogeny of Actinomycetia was taken from GTDB (Parks et al. 2019, 2022). The origin of the order Mycobacteriales, which contains the genus *Frankia* is indicated in the phylogenetic tree by an orange circle; families belonging to the Mycobacteriales are labelled by asterisks. The groups of hydrogenases that occur in *Frankia* are given in color and bold print for the entire class—1 h (synton-1), 2a (synton-2), 1f (synton-3), and 3b (synton-4)—, the other groups are given in grey.

Table 2A. Acetylene reduction and H_2 evolution in nodules.

Plant species	*Frankia* strain	Plant height[a] (cm)	H_2 (µmol H_2 h^{-1})	ARA (µmol C_2H_4 h^{-1})	RE	Average	StDEV	Nodule weight (g)	Average	StDEV	dpi	Season
A. incana	ACN14a	42.5	0	9.38	1.0	1.0	0	0.118	0.194	0.053	159	Spring 2021–05–15
A. incana	ACN14a	55.0	0	7.36	1.0			0.210				
A. glutinosa	ACN14a	46.5	0	8.48	1.0			0.209				
A. glutinosa	ACN14a	49.5	0	9.28	1.0			0.24				
A. incana	AiOr	54.0	0.92	14.37	0.94	0.81	0.15	0.410	0.303	0.081	162	Spring 2021–05–18
A. incana	AiOr	47.5	1.80	13.47	0.87			0.231				
A. glutinosa	AiOr	38.5	2.19	12.00	0.83			0.252				
A. glutinosa	AiOR	40.5	4.46	10.98	0.60			0.320				
A. incana	AgTrS	31.5	2.58	8.97	0.71	0.72	0.027	0.267	0.251	0.096	163	Spring 2021–05–19
A. incana	AgTrS	43.5	3.43	14.72	0.76			0.365				
A. glutinosa	AgTrS	53.0	3.98	13.73	0.71			0.131				
A. glutinosa	AgTrS	37.5	3.90	11.46	0.70			0.240				
A. incana	UmASH1	44.0	7.6	13.51	0.43	0.41	0.1	0.266	0.203	0.084	161	Spring 2021–05–17
A. incana	UmASH1	46.5	7.76	10.48	0.26			0.253				
A. glutinosa	UmASH1	38.0	7.48	14.04	0.47			0.212				
A. glutinosa	UmASH1	56.0	6.97	12.98	0.46			0.082				
A. incana	Local source	59.0	4.59	14.30	0.68	0.63	0.09	0.432	0.351	0.128	259	Summer 2021–08–23
A. Incana	Local source	77.0	5.22	12.95	0.60			0.421				
A. glutinosa	Local source	74.0	2.83	12.41	0.77			0.312				
A. glutinosa	Local source	38.5	7.66	17.36	0.56			0.157				
D. glomerata	Dg2	20.5	0	1.99	1.0	1.0	0	0.361	0.470	0.153	267	Summer 2021–08–31
D. glomerata	Dg2	18.5	0	2.40	1.0			0.578				
H. rhamnoides	Hr75.2	45.0	0	11.32	1.0	1.0	0	0.826	0.478	0.491	260	Summer 2021–08–24
H. rhamnoides	Hr75.2	30.0	0	2.13	1.0			0.132				
H. rhamnoides	Hr75.2	31.5	0	0.99	1.0	1.0	0	0.0926	0.091	0.002	161	Winter 2022–02–08
H. rhamnoides	Hr75.2	29.0	0	1.38	1.0			0.0895				

StDEV, standard deviation.
[a]The *H. rhamnoides* plants had three to four branches; the length of the longest one is given under 'plant height'.

Table 2B. Acetylene reduction and H_2 evolution in culture.

Frankia strain	N source	H_2 (nmol H_2 mg prot $^{-1}$)	ARA (µmol C_2H_4 mg prot^{-1})	StDEV
Hr75.2	Yes	0	0	0
Hr75.2	No	0	0.44	0.07
Hr75.2	Yes	0	0	0
Hr75.2	No	0	0.14	0.02

Averages of four replicates are shown. StDEV, standard deviation.

source' (*Ca*. F. nodulisporulans LS). While based on the biomass acquisition of the nodulated plants compared to control plants and on nitrogenase activity (Table 2), both inocula contained effective microsymbionts (data not shown), nodule phenotypes were strikingly different. Nodules induced by the 'local source' had much larger lobes than those induced by UmASH1 or by any of the other alder-compatible inocula used in this study, and this large-lobe phenotype has been stable since the 1980s (data not shown; A. Sellstedt, personal observations). In order to address the difference in nodule structure and nodule uptake hydrogenase activity, DNA was isolated from nodules induced by UmASH1 and the original 'local source' on *A. glutinosa*, and 16S profiles were generated. The results are depicted in Fig. S5. The profiles were rather similar but showed differences in some minor components in so far that the 'local source' nodules contained *Acidothermus*, *Asticcacaulus*, and *Variovorax* while the UmASH1 nodules contained *Pseudolabrys*, *Haliangium*, and *Reyranella*.

In the current study, all *Ca*. F. nodulisporulans strains behaved differently with regard to RE. The type strain AgTrS showed an RE of 0.7–0.76. The 'local source' had an RE of 0.56–0.77 while UmASH1, which is derived from nodules induced by the 'local source' showed an RE of 0.26–0.47, which is lower than should be permitted by the stoichiometry of the nitrogenase reaction in the absence of [NiFe] hydrogenase (Fig. 4). Thus, UmASH1 nodules produced more H_2 than the nitrogenase of the dominant *Frankia* strain. A comparison of the 16S profiles of nodules from nodules induced by UmASH1 and nodules induced directly by the 'local source' showed that when *Frankia* sequences were excluded, the rest of the microbiomes were dominated by pseudomonads with streptomycetes the next largest group; rhizobia were found as well. Streptomycetes include strains with [NiFe] hydrogenases (Constant et al. 2011). Furthermore, both microbiomes contain *Novosphingobium* spp. strains of which contain a group 2a [NiFe] hydrogenase and/or display nitrogenase activity (Islam et al. 2020, Bizjak et al. 2023). Furthermore, fermentative hydrogen production as found in the soil bacterium *Mycobacterium smegmatis* (Berney et al. 2014), might be performed by members of the nodule microbiome. Altogether, it is likely that the non-*Frankia* nodule microbiome contributes to H_2 evolution and H_2 uptake. Similarly, the nodule morphology phenotype might be due to gibberellin production by (a) member(s) of the non-*Frankia* nodule microbiome since in legume nodules, gibberellin increases nodule size (McAdam et al. 2018).

Expression of genes from the different syntons in nodules and in culture: is the expression of different syntons linked to different life styles?

Expression of one gene from each synton (for synton-2, two genes representing different operons were used) was analysed in nodules, and, for *F. alni* ACN14a and *F. soli* Hr75.2.1, also in nitrogen-fixing cultures. The results are shown in Fig. 4. Consistent with previous results (Leul et al. 2007, 2009), in ACN14a, both synton-1 and synton-2 were expressed in nodules and in nitrogen-fixing cultures. In cultures containing ammonium, synton-1 expression levels were significantly higher than those of synton-2, the expression of which was barely detectable. In nodules, expression of synton-2 was significantly higher than that of synton-1, suggesting that group 1 h hydrogenase does not play a great role during symbiosis. In the three strains representing *Ca*. F. nodulisporulans, AgTrS, UmASH1, and LS, synton-2 was expressed. Expression levels of the operon with the structural genes and the *hypA2B2* operon were not significantly different, but expression levels of the *hypF2′–hypC2–hypD2–hypE2* operon were higher in UmASH1 and LS than in the type strain AgTrS (Figs 2 and 4). However, the lack of an intact HypF protein should not allow uptake hydrogenase activity in any *Ca*. F. nodulisporulans strain (Fig. 2). In summary, we can say that the erosion of synton-2 in *Ca*. F. nodulisporulans does not seem to involve the promoters.

Both operons of synton-2 were expressed in both cluster-2 species examined, *Ca*. F. datiscae Dg1 and *Ca*. F. californiensis Dg2, consistent with the uptake hydrogenase activity shown by the corresponding nodules (Table 2A, Fig. 4). All three syntons, synton-1, synton-3, and synton-4, were expressed in nodules induced by the cluster-3 strain *F. soli* Hr75.2.1, as well as in nitrogen-fixing and non-nitrogen-fixing cultures. For synton-4, expression levels were significantly higher under nitrogen-fixing conditions (nodules and culture) than during growth on ammonium as nitrogen source, while for synton-1, expression levels were significantly higher in nitrogen-fixing cultures than in nodules.

Can the erosion of particular [NiFe] hydrogenase syntons in specific subgroups of *Frankia* strains provide information about their function?

Frankia strain genome sizes show an unusually large variation between 4.3 MB and more than 10 MB (Normand et al. 2007, Nguyen and Pawlowski 2017, Herrera-Belaroussi et al. 2020). There are three lineages of genome reduction within the genus: cluster-1 strain genomes are 7.5–8 MB in size while cluster-1c strains have genomes of 5–6 MB and cluster-1 strains that sporulate in nodules have genomes of 4.2–5 MB (Pozzi et al. 2020, Herrera-Belaroussi et al. 2020; this study). *Frankia* cluster-2 strains have genome sizes between 5 and 6 MB (Nguyen and Pawlowski 2017, Nguyen et al. 2019, Berckx et al. 2022). Strains of these lineages have limited saprotrophic capabilities. Cluster-1c strains apparently depend on the presence of their host plant to grow saprotrophically in soil (Samant et al. 2016), cluster-2 strains as well as cluster-1 strains that sporulate in nodules have lost some saprotrophic capabilities as exemplified by the fact that with two exceptions in cluster-2 (Gtari et al. 2015, Gueddou et al. 2019), they could never be grown in culture.

However, while cluster-2 strains as well as cluster-1 strains that sporulate in nodules have lost [NiFe] hydrogenase synton-1, members of *Ca*. F. nodulisporulans have also lost functionality of synton-2 in that the *hypF* gene has broken up in two parts. The loss of synton-2 did not take place in the type strain of the other lineage that sporulates in nodules, *Ca*. Frankia alpina AiOr

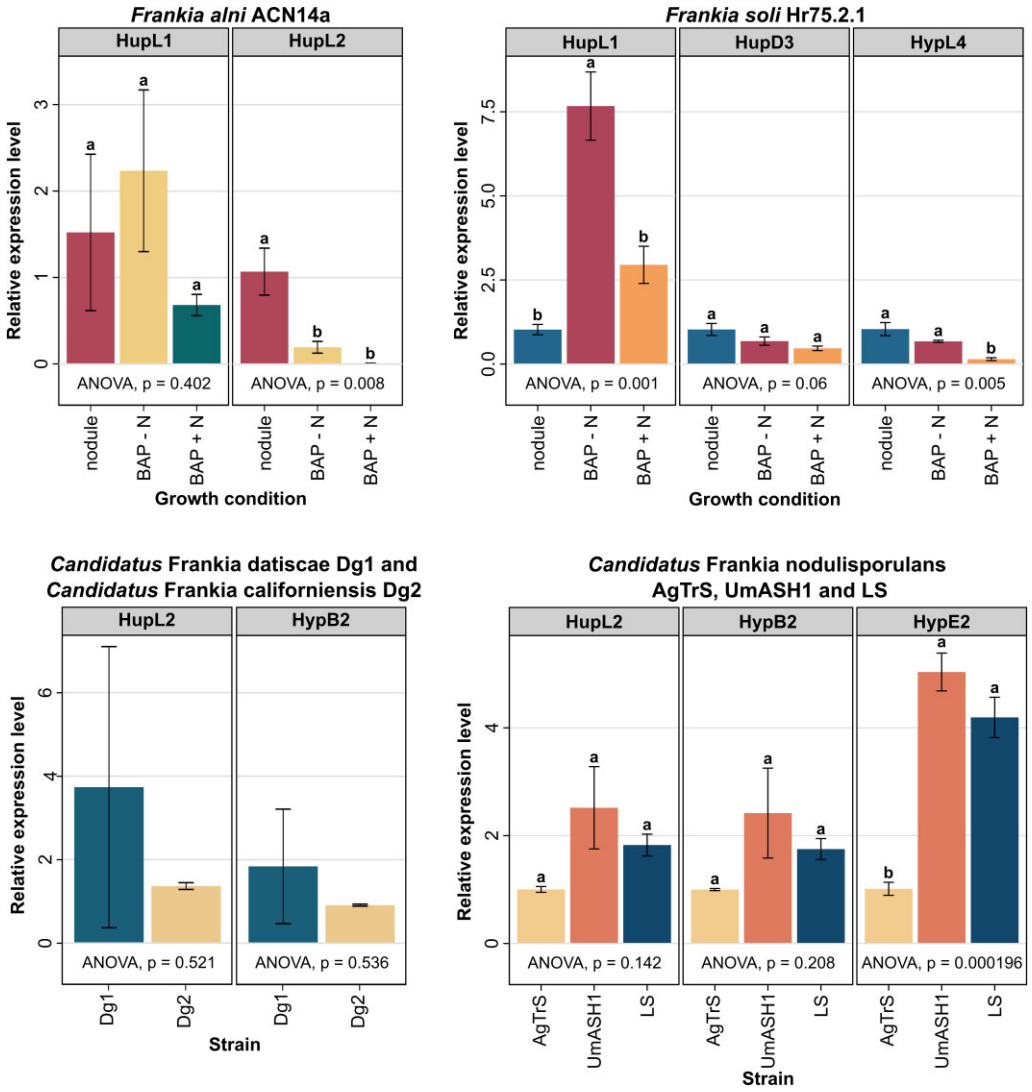

Figure 4. RT-qPCR to analyse the expression levels of different syntons. Expression levels of all available [NiFe] hydrogenase syntons were analysed: *F. alni* (synton-1 and synton-2), three strains representing *Candidatus* Frankia nodulisporulans (synton-2), the two cluster-2 species *Candidatus* Frankia datiscae Dg1 and *Candidatus* Frankia californiensisDg2 (synton-2), and the cluster-3 species *F. soli* Hr75.2.1 (synton-1, synton-3, and synton-4). Expression was analysed in nodules and, if the strains were cultivable, in cultures without nitrogen (nitrogen-fixing conditions) and with nitrogen, respectively. Each column represents the average of three biological replicates; error bars show standard error. Bars labelled with the same letter are not statistically different.

(Pozzi et al. 2020); however, the efficiency of uptake hydrogenase in AiOr was insufficient to reassimilate all hydrogen released during nitrogen fixation, suggesting that at least one of the genes of synton-2 had accumulated mutations. In summary, cluster-1c strains still have both synton-1 and synton-2, cluster-2 strains have lost synton-1, while cluster-1 strains that sporulate in nodules lost synton-1 (*Ca.* F. nodulisporulans) or are losing synton-1 (*Ca.* F. alpina), while the erosion of synton-2 is still in progress in both species.

The fact that synton-1 (group 1 h) was lost in species that show genome reduction and reduced saprotrophic potential indicates that while also expressed in nodules when present, synton-1 is not required for the symbiotic life style. On the contrary, the fact that cluster-2 *Frankia* strains, which show a reduced saprotrophic potential have lost synton-1 (group 1 h) and have retained either synton-2 (group 2a; continental lineage) or synton-3 (group 1f; island lineage), indicates that either hydrogenase group is well suited for the symbiotic lifestyle, and it seems to have been a matter of chance whether synton-2 (group 2a) or synton-3 (group 1f) was retained during the loss of saprotrophic potential. As stated above, Søndergaard et al. (2016) linked both hydrogenase group 1f and 2a to the recycling of H_2 produced during nitrogen fixation, which is consistent with the RE of 1 for nodules of *D. glomerata* (only group 2a present) and the RE of 1 for nodules of *H. rhamnoides* induced by Hr75.2.1 (group 2a and 1f present).

The role of syntons-2 and -3 (groups 2a and 1f) in symbiosis is consistent with the earlier hypothesis, based on the relative expression levels of synton-1 and synton-2/3 genes under different growth conditions, that syntons-2 and 3 (groups 2a and 1f) are more important for symbiosis and synton-1 (group 1 h) is more important for growth in soil (Leul et al. 2007, 2009). However, group 1 h (synton-1) has not been linked to the recycling of H_2 produced by nitrogenase by Søndergaard et al. (2016), but to hydrogenotrophic respiration using O_2 as terminal electron acceptor; the enzyme is supposed to scavenge electrons from atmospheric H_2 to fuel the respiratory chain during carbon-starvation.

The suggestion that the role of synton-1 (group 1 h) in *Frankia* is not linked to nitrogen fixation is consistent with the fact that all strains examined of the nonsymbiotic cluster-4 have lost the *nif* genes, but retained synton-1 (group 1 h). Thus, the data available thus far suggest that hydrogenases groups 1f and 2a are linked to the recycling of H_2 produced by nitrogenase, while hydrogenase group 1 h is not.

Synton-4 encodes a cytosolic bidirectional [NiFe] hydrogenase of group 3b that according to Søndergaard et al. (2016) can couple the oxidation of NADP to the fermentative evolution of H_2. Its pattern of occurrence is interesting: synton-4 was only found in two genomes of nonsymbiotic *Frankia* species (and is defective in one of them), was not found in any genome of cluster-2 *Frankia,* and its presence in genomes from the other symbiotic clusters, 1 and 3, is patchy (Table S2). However, there is one exception: all genomes of *Casuarina*-infective strains (*F. casuarinae,* cluster 1c) contain synton-4 (data not shown). One feature that distinguishes these strains from all other symbiotic *Frankia* strains is their environment in symbiosis: microaerobic conditions are established in the infected cells of nodules of *(Allo-)Casuarina* spp., abolishing the need for the differentiation of *Frankia* vesicles as places of nitrogen fixation; here, nitrogen fixation takes place in hyphae (Zhang and Torrey 1985, Berg and McDowell 1987, 1988, Schubert et al. 2013). It therefore seems plausible that group 3b [NiFe] hydrogenase is important for effective carbon metabolism of *Frankia* under microaerobic conditions. However, group 3b [NiFe] hydrogenases have been described as oxygen tolerant (Søndergaard et al. 2016), and another unusual feature of *F. casuarinae* is its comparatively fast growth on solid medium (Bassi and Benson 2007). In summary, further studies are required to elucidate the function of group 3b [NiFe] hydrogenase in *Frankia*.

Symbiotic nitrogen fixation and uptake hydrogenase

Most rhizobial strains examined cannot fix nitrogen *ex planta* and do not possess uptake hydrogenase (Annan et al. 2012, Brito et al. 2005, Schubert and Evans 1976). While the reabsorption of the H_2 produced by nitrogenase would save energy, rhizobia in nodules are provided with carbon sources by the host plant, i.e. they live in an energy-rich environment. The low occurrence of rhizobial uptake hydrogenase might indicate that the host plants can afford the energy loss, and/or that the release of hydrogen gas in the rhizosphere might even be advantageous for the host plant with regard to selection of the rhizosphere microbiome. The results of Greening et al. (2016) suggest that H_2 is a widely utilized energy source for microbial growth and survival; thus, H_2 evolution could be important for rhizosphere conditioning. Consistent with this hypothesis, Li et al. (2018) report that hydrogen treatment of rhizosphere soil samples of *Medicago sativa* plots influenced the composition of microbial communities.

However, symbiotic *Frankia* strains can fix nitrogen *ex planta;* therefore, here the selection for maintaining uptake hydrogenase might also depend on its being advantageous during saprotrophic growth. On the other hand, nitrogen fixation does not seem to be a benefit during saprotrophic growth of *Frankia* based on the fact that the nonsymbiotic strains, cluster-4, have lost the capacity for nitrogen fixation. Then why does the vast majority of symbiotic *Frankia* strains display uptake hydrogenase activity, while the vast majority of symbiotic rhizobia does not? It should be noted that the legumes analysed for uptake hydrogenase activity were herbaceous, while actinorhizal plants, with one exception, represent trees or woody shrubs with perennial nodules. This is likely to affect the need, or lack of it, for continuous rhizosphere conditioning. It is tempting to speculate that for annual plants, the benefits of quick rhizosphere conditioning by hydrogen evolution outweigh the loss of energy, while for perennial plants, rhizosphere conditioning is predominantly based on root exudation and energy preservation is of larger importance. Plant selection for Hup^+ or Hup^- microsymbionts would explain the differences of distribution of uptake hydrogenase in symbiotic rhizobia versus symbiotic *Frankia* strains.

The functionality of [NiFe] hydrogenases is well examined in bacterial species that, in contrast with *Frankia* spp., are user-friendly for biotechnological approaches, *Escherichia coli* and *Cupriavidus necator* (Lamont and Sargent 2017). While the heterologous expression of *Frankia* genes in Gram-negative bacteria is not an option due to the high GC content of *Frankia* genomes and the consequently different codon preference (Kucho et al. 2013), the use of synthetic genes is a promising future option for detailed analyses.

Conclusions

Frankia strains can contain four different [NiFe] hydrogenase syntons representing groups 1f, 1 h, 2a, and 3b according to Søndergaard et al. (2016); no more than three types were found in any individual genome. The phylogeny of the structural proteins of groups 1f, 1 h, and 2a follows *Frankia* phylogeny while the phylogeny of accessory proteins does not necessarily. Evidence suggests that accessory proteins can be exchanged between different groups of [NiFe] hydrogenases, which would change the selection pressure on them. An analysis of different [NiFe] hydrogenase types in Actinomycetia showed that under the most parsimonious assumption, all four types were present in the ancestral *Frankia* strain. Based on Hup activities analysed and the losses of syntons in different lineages of genome reduction, we can conclude that groups 1f and 2a are involved in recycling H_2 formed by nitrogenase while group 1 h is not. We postulate that the difference in selection for Hup^+ microsymbionts between legumes and actinorhizal plants is due to the fact that actinorhizal plants are trees or woody shrubs while the legumes analysed are herbaceous, a difference which is likely to affect the importance of short-term rhizosphere conditioning via release of H_2.

Author contributions

Katharina Pawlowski (Conceptualization, Formal analysis, Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing), Daniel Wibberg (Formal analysis, Visualization, Writing – review & editing), Sara Mehrabi (Investigation, Writing – review & editing), Nadia Binte Obaid (Investigation, Supervision, Visualization, Writing – review & editing), András Patyi (Investigation, Writing – review & editing), Fede Berckx (Supervision, Writing – review & editing), Han Nguyen (Investigation, Writing – review & editing), Michelle Hagen (Formal analysis, Visualization, Writing – review & editing), Daniel Lundin (Formal analysis, Supervision, Writing – review & editing), Andreas Brachmann (Investigation, Writing – review & editing), Jochen Blom (Software, Writing – review & editing), Aude Herrera-Belaroussi (Resources, Writing – review & editing), Danis Abrouk (Formal analysis, Writing – review & editing), Petar Pujic (Formal analysis, Investigation, Writing – review & editing), Ann-Sofi Hahlin (Investigation, Writing – review & editing), Jörn Kalinowski (Supervision, Writing – review & editing), Philippe Normand (Conceptualization, Resources, Supervision, Writing – review & editing), and

Anita Sellstedt (Conceptualization, Investigation, Resources, Supervision, Writing – original draft, Writing – review & editing)

Supplementary data

Supplementary data is available at FEMSEC Journal online.

Conflict of interest: None declared.

Funding

Open access funding provided by Stockholm University. This project was supported by two grants from the Swedish Research Council Vetenskapsradet (VR 2012–03061 and 2019–05540) to K.P. Funding to ASt from FORMAS (2007) and Department of Plant Physiology Emeriti Funds, Umeå University (Sweden) is greatly acknowledged. The bioinformatics support of the BMBF-funded project 'Bielefeld-Giesen Center for Microbial Bioinformatics' (BiGi) and the BMBF grant FKZ 031A533 within the German Network for Bioinformatics Infrastructure (de.NBI) are gratefully acknowledged. The analysis of hydrogenases in Actinomycetia was enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) partially funded by the Swedish Research Council through grant agreement number 2018–05973.

References

Albrecht SL, Maier RJ, Hanus FJ *et al.* Hydrogenase in *Rhizobium japonicum* increases nitrogen fixation by nodulated soybeans. *Science* 1979;**203**:1255–7.

Alloisio N, Queiroux C, Fournier P *et al.* The *Frankia alni* symbiotic transcriptome. *MPMI* 2010;**23**:593–607. https://doi.org/10.1094/MPMI-23-5-0593.

Annan H, Golding AL, Zhao Y *et al.* Choice of hydrogen uptake (Hup) status in legume-rhizobia symbioses. *Ecol Evol* 2012;**2**:2285–90.

Bassi CA, Benson DR. Growth characteristics of the slow-growing actinobacterium *Frankia* sp. strain CcI3 on solid media. *Physiol Plant* 2007;**130**:391–9.

Bautista-Guerrero HH, Cruz H, Nesme X *et al.* Genomospecies identification and phylogenomic relevance of AFLP analysis of isolated and non-isolated strains of *Frankia* spp. *Syst Appl Microbiol* 2011;**34**:200–6.

Benoist P, Müller A, Diem HG *et al.* High molecular-mass multicatalytic proteinase 577 complexes produced by the nitrogen-fixing actinomycete *Frankia* strain BR. *J Bacteriol* 1992;**174**:1495–504.

Berckx F, Bandong CM, Nguyen TV *et al.* A tale of two lineages: how the strains of the earliest divergent symbiotic *Frankia* clade spread over the world. *BMC Genomics* 2022;**23**:602.

Berckx F, Wibberg D, Kalinowski J *et al.* The peptidoglycan biosynthesis gene murC in *Frankia*: actinobacterial vs. plant type. *Genes* 2020;**11**:432. https://doi.org/10.3390/genes11040432.

Berg RH, McDowell L. Endophyte differentiation in *Casuarina* actinorhizae. *Protoplasma* 1987;**136**:104–17.

Berg RH, McDowell L. Cytochemistry of the wall of infected cells in *Casuarina* actinorhizae. *Can J Bot* 1988;**66**:2038–47.

Berney M, Greening C, Conrad R *et al.* An obligately aerobic soil bacterium activates fermentative hydrogen production to survive reductive stress during hypoxia. *Proc Natl Acad Sci USA* 2014;**111**:11479–84. https://doi.org/10.1073/pnas.1407034111.

Bizjak T, Sellstedt A, Gratz R *et al.* Presence and activity of nitrogen-fixing bacteria in Scots pine needles in a boreal forest: a nitrogen-addition experiment. *Tree Physiol* 2023;**43**:1354–64. https://doi.org/10.1093/treephys/tpad048.

Blom J, Albaum SP, Doppmeier D *et al.* EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinf* 2009;**10**:154.

Blom J, Kreis J, Spänig S *et al.* EDGAR2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Res* 2016;**44**:W22–8.

Bokulich NA, Subramanian S, Faith JJ *et al.* Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 2013;**10**:57–9. https://doi.org/10.1038/nmeth.2276.

Brito B, Baginsky C, Palacios JM *et al.* Biodiversity of uptake hydrogenase systems from legume endosymbiotic bacteria. *Biochem Soc Trans* 2005;**33**:33–5. https://doi.org/10.1042/BST0330033.

Burns RC, Hardy RWF. *Nitrogen Fixation in Bacteria and Higher Plants*. New York, NY: Springer, 1975.

Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000;**17**:540–52. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

Chaumeil PA, Mussig AJ, Hugenholtz P *et al.* GTDB-Tk: toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 2020;**36**:1925–7. https://doi.org/10.1093/bioinformatics/btz848.

Constant P, Chowdhury SP, Hesse L *et al.* Genome data mining and soil survey for the novel group 5 [NiFe]-hydrogenase to explore the diversity and ecological importance of presumptive high-affinity H_2-oxidizing bacteria. *Appl Environ Microbiol* 2011;**77**:6027–35.

Cotin-Galvan L, Pozzi AC, Schwob G *et al.* In-planta sporulation capacity enhances infectivity and rhizospheric competitiveness of *Frankia* strains. *Microbes Environ* 2016;**31**:11–8.

Dart PJ, Day JM. Effects of inoculation temperature and oxygen tension on nitrogenase activity of legume root nodules. *Plant Soil* 1971;**35**:167–84.

de Reuse H, Vinella D, Cavazza C. Common themes and unique proteins for the uptake and trafficking of nickel, a metal essential for the virulence of *Helicobacter pylori*. *Front Cell Infect Microbiol* 2013;**3**:94.

Dixon RO. Hydrogenase in legume root nodule bacteroids: occurrence and properties. *Archiv Mikrobiol* 1972;**85**:193–201.

Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol* 2011;**7**:e1002195. https://doi.org/10.1371/journal.pcbi.1002195.

Edgar RC, Haas BJ, Clemente JC *et al.* UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;**27**:2194–200. https://doi.org/10.1093/bioinformatics/btr381.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7. https://doi.org/10.1093/nar/gkh340

Eikmeyer F, Hadiati A, Szczepanowski R *et al.* The complete genome sequences of four new IncN plasmids from wastewater treatment plant effluent provide new insights into IncN plasmid diversity and evolution. *Plasmid* 2012;**68**:13–24. https://doi.org/10.1016/j.plasmid.2012.01.011.

Evans HJ, Koch B, Klucas R. Preparation of nitrogenase from nodules and separation into components. *Methods Enzymol* 1972;**24**:470–6.

Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. *Genome Res* 1998;**8**:195–202. https://doi.org/10.1101/gr.8.3.195

Goris J, Konstantinidis KT, Klappenbach JA *et al.* DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;**57**:81–91.

Greening C, Biswas A, Carere CR *et al.* Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely

utilised energy source for microbial growth and survival. *ISME J* 2016;**10**:761–77.
Gtari M, Ghodhbane-Gtari F, Nouioui I *et al.* Cultivating the uncultured: growing the recalcitrant cluster-2 *Frankia* strains. *Sci Rep* 2015;**5**:13112. https://doi.org/10.1038/srep13112.
Gtari M, Ghodhbane-Gtari F, Nouioui I. *Frankia soli* sp. nov., an actinobacterium isolated from soil beneath *Ceanothus jepsonii*. *Int J Syst Evol Microbiol* 2020;**70**:1203–9. https://doi.org/10.1099/ijsem.0.003899.
Gueddou A, Swanson E, Hezbri K *et al.* Draft genome sequence of the symbiotic *Frankia* sp. strain BMG5.30 isolated from root nodules of *Coriaria myrtifolia* in Tunisia. *Antonie Van Leeuwenhoek* 2019;**112**:67–74. https://doi.org/10.1007/s10482-018-1138-1.
Herrera-Belaroussi A, Normand P, Pawlowski K *et al. Candidatus* Frankia nodulisporulans sp. nov., an *Alnus glutinosa*-infective *Frankia* species unable to grow in pure culture and able to sporulate in-planta. *Syst Appl Microbiol* 2020;**43**:126134.
Hoagland DR, Arnon DT. The water-culture method for growing plants without soil. In: *California Agriculture Experiment Station Circular* 347. Berkeley, CA: University of California, Berkeley, 1938.
Huson DH, Scornavacca C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol* 2012;**61**:1061–7.
Huss-Danell K. Influence of host (*Alnus* and *Myrica*) on infectivity, N_2 fixation, spore formation and hydrogenase activity in *Frankia*. *New Phytol* 1991;**119**:121–7.
Islam ZF, Welsh C, Bayly K *et al.* A widely distributed hydrogenase oxidises atmospheric H_2 during bacterial growth. *ISME J* 2020;**14**:2649–58.
Islam ZF. *Determinants of atmospheric trace gas oxidation in bacteria*. Ph.D. Thesis, Monash University, 2020.
Karan S, Pratap B, Yadav SP *et al.* Structural and functional characterization of *M. tuberculosis* sedoheptulose-7-phosphate isomerase, a critical enzyme involved in lipopolysaccharide biosynthetic pathway. *Sci Rep* 2020;**10**:20813. https://doi.org/10.1038/s41598-020-77230-8.
Kashanski CR, Schwintzer CR. Distribution of spore-positive and spore-negative nodules of *Myrica gale* in Maine, USA. *Plant Soil* 1987;**104**:113–20.
Khadka N, Milton RD, Shaw S *et al.* Mechanism of nitrogenase H_2 formation by metal-hydride protonation probed by mediated electrocatalysis and H/D isotope effects. *J Am Chem Soc* 2017;**139**:13518–24.
Kneidinger B, Marolda C, Graninger M *et al.* Biosynthesis pathway of ADP-L-glycero-beta-D-manno-heptose in *Escherichia coli*. *J Bacteriol* 2002;**184**:363–9. https://doi.org/10.1128/JB.184.2.363-369.2002.
Kucho K, Kakoi K, Yamaura M *et al.* Codon-optimized antibiotic resistance gene improves efficiency of transient transformation in *Frankia*. *J Biosci* 2013;**38**:713–7. https://doi.org/10.1007/s12038-013-9361-4.
Lamont CM, Sargent F. Design and characterisation of synthetic operons for biohydrogen technology. *Arch Microbiol* 2017;**199**:495–503. https://doi.org/10.1007/s00203-016-1322-5.
Leul M, Normand P, Sellstedt A. The organization, regulation and phylogeny of uptake hydrogenase genes in *Frankia*. *Physiol Plant* 2007;**130**:464–70.
Leul M, Normand P, Sellstedt A. The phylogeny of uptake hydrogenases in *Frankia*. *Int Microbiol* 2009;**12**:23–8.
Li Z, Liu X, Liu R *et al.* Insight into bacterial community diversity and monthly fluctuations of *Medicago sativa* rhizosphere soil in response to hydrogen gas using Illumina high-throughput sequencing. *Curr Microbiol* 2018;**75**:1626–33.
Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;**27**:2957–63. https://doi.org/10.1093/bioinformatics/btr507.
Mattsson U, Johansson L, Sandström G *et al.* *Frankia* KB5 possesses a hydrogenase immunologically related to membrane-bound. *Curr Microbiol* 2001;**42**:438–41.
Mattsson U, Sellstedt A. Hydrogenase in *Frankia* KB5: expression of and relation to nitrogenase. *Can J Microbiol* 2000;**46**:1091–5.
McAdam EL, Reid JB, Foo E. Gibberellins promote nodule organogenesis but inhibit the infection stages of nodulation. *J Exp Bot* 2018;**69**:2117–30. https://doi.org/10.1093/jxb/ery046.
Mus F, Colman DR, Peters JW *et al.* Geobiological feedbacks, oxygen, and the evolution of nitrogenase. *Free Radical Biol Med* 2019;**140**:250–9.
Nguyen TV, Pawlowski K. *Frankia* and actinorhizal plants: symbiotic nitrogen fixation. In: Mehnaz S (ed.), *Rhizotrophs: Plant Growth Promotion to Bioremediation*. New York, NY: Springer, 2017, 237–61.
Nguyen TV, Wibberg D, Battenberg K *et al.* An assemblage of *Frankia* cluster II strains from California contains the canonical nod genes and also the sulfotransferase gene nodH. *BMC Genomics* 2016;**17**:796. https://doi.org/10.1186/s12864-016-3140-1.
Nguyen TV, Wibberg D, Vigil-Stenman T *et al.* *Frankia*-enriched metagenomes from the earliest divergent symbiotic *Frankia* cluster: they come in teams. *Genome Biol Evol* 2019;**11**:2273–91. https://doi.org/10.1093/gbe/evz153.
Normand P, Lapierre P, Tisa LS *et al.* Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. *Genome Res* 2007;**17**:7–15. https://doi.org/10.1101/gr.5798407.
Normand P, Nguyen TV, Battenberg K *et al.* Proposal of '*Candidatus* Frankia californiensis', the uncultured symbiont in nitrogen-fixing root nodules of a phylogenetically broad group of hosts endemic to western North America. *Int J Syst Evol Microbiol* 2017;**67**:3706–15. https://doi.org/10.1099/ijsem.0.002147.
Normand P, Orso S, Cournoyer B *et al.* Molecular phylogeny of the genus *Frankia* and related genera and emendation of family *Frankiaceae*. *Int J Syst Bacteriol* 1996;**46**:1–9.
Parks DH, Chuvochina M, Rinke C *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucl Acids Res* 2022;**50**:D785–94. https://doi.org/10.1093/nar/gkab776.
Parks DH, Chuvochina M, Waite DW *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 2018;**36**:996–1004. https://doi.org/10.1038/nbt.4229.
Pawlowski K, Bisseling T. Rhizobial and actinorhizal symbioses: what are the shared features?. *Plant Cell* 1996;**8**:1899–913.
Persson T, Battenberg K, Demina IV *et al. Candidatus* Frankia datiscae Dg1, the actinobacterial microsymbiont of *Datisca glomerata*, expresses the canonical nod genes nodABC in symbiosis with its host plant. *PLoS One* 2015;**10**:e0127630.
Pesce C, Swanson E, Simpson S *et al.* Draft genome sequence of the symbiotic *Frankia* sp. strain KB5 isolated from root nodules of *Casuarina equisetifolia*. *J Genomics* 2017;**5**:64–7. https://doi.org/10.7150/jgen.20887.
Peters JW, Schut GJ, Boyd ES *et al.* [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochim Biophys Acta Mol Cell Res* 2015;**1853**:1350–69.
Piché-Choquette S, Constant P. Molecular hydrogen, a neglected key driver of soil biogeochemical processes. *Appl Environ Microbiol* 2019;**85**:e02418–18. https://doi.org/10.1128/AEM.02418-18.

Pozzi ACM, Herrera-Belaroussi A, Schwob G *et al.* Proposal of '*Candidatus* Frankia alpina', the uncultured symbiont of *Alnus alnobetula* and *A. incana* that forms spore-containing nitrogen-fixing root nodules. *Int J Syst Evol Microbiol* 2020;**70**:5453–9.

Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum-evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 2009;**26**:1641–50.

Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 2010;**5**:e9490.

Richau KH, Kudahettige RL, Pujic P *et al.* Structural and gene expression analyses of uptake hydrogenases and other proteins involved in nitrogenase protection in *Frankia*. *J Biosci* 2013;**38**:703–12.

Richau KH, Pujic P, Normand P *et al.* Nitrogenase and hydrogenase of the actinobacteria *Frankia*: from gene expression to proteins function. *J Microbiol Res* 2017;**7**:79–92. https://doi.org/10.5923/j.microbiology.20170704.01.

RStudio Team. *RStudio: Integrated Development Environment for R. RStudio*. Boston, MA: PBC, 2022. http://www.rstudio.com/ (25 July 2024, date last accessed).

Samant S, Dawson JO, Hahn D. Growth responses of introduced *Frankia* strains to edaphic factors. *Plant Soil* 2016;**400**:123–32. https://doi.org/10.1007/s11104-015-2720-1.

Schubert KR, Evans HJ. Hydrogen evolution: a major factor affecting the efficiency of nitrogen fixation in nodulated symbionts. *Proc Natl Acad Sci USA* 1976;**73**:1207–11.

Schubert M, Koteyeva NK, Zdyb A *et al.* Lignification of cell walls of infected cells in *Casuarina glauca* nodules that depend on symplastic sugar supply is accompanied by reduction of plasmodesmata number and narrowing of plasmodesmata. *Physiol Plant* 2013;**147**:524–40.

Schwob G, Roy M, Pozzi AC *et al.* In planta sporulation of *Frankia* spp. as a determinant of alder-symbiont interactions. *Appl Environ Microbiol* 2018;**84**:e01737–18.

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–9. https://doi.org/10.1093/bioinformatics/btu153.

Sellstedt A, Huss-Danell K, Ahlqvist A-S. Nitrogen fixation and biomass production in symbioses between *Alnus incana* and *Frankia* strains with different hydrogen metabolism. *Physiol Plant* 1986;**66**:99–107.

Sellstedt A. Acetylene reduction, H_2 evolution and $^{15}N_2$ fixation in the *Alnus incana-Frankia* symbiosis. *Planta* 1986;**167**:382–6.

Sen A, Daubin V, Abrouk D *et al.* Phylogeny of the class *Actinobacteria* revisited in the light of complete genomes. The orders '*Frankiales*' and *Micrococcales* should be split into coherent entities: proposal of *Frankiales* ord. nov., *Geodermatophilales* ord. nov., *Acidothermales* ord. nov. and *Nakamurellales* ord. nov. *Int J Syst Evol Microbiol* 2014;**64**:3821–32. https://doi.org/10.1099/ijs.0.063966-0.

Søndergaard D, Pedersen CN, Greening C. HydDB: a web tool for hydrogenase classification and analysis. *Sci Rep* 2016;**6**:34212. https://doi.org/10.1038/srep34212.

Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 2007;**56**:564–77. https://doi.org/10.1080/10635150701472164.

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 2021;**38**:3022–7. https://doi.org/10.1093/molbev/msab120.

Tjepkema JD, Schwintzer CR, Benson DR. Physiology of actinorhizal nodules. *Annu Rev Plant Physiol* 1986;**37**:209–32.

Valvano MA, Messner P, Kosma P. Novel pathways for biosynthesis of nucleotide-activated glycero-manno-heptose precursors of bacterial glycoproteins and cell surface polysaccharides. *Microbiology* 2002;**148**:1979–89. https://doi.org/10.1099/00221287-148-7-1979.

Vignais PM, Billoud B, Meyer J. Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev* 2001;**25**:455–501. https://doi.org/10.1111/j.1574-6976.2001.tb00587.x.

Vignais PM, Billoud B. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem Rev* 2007;**107**:4206–72. https://doi.org/10.1021/cr050196r.

Walsh KB, Layzell DB. Carbon and nitrogen assimilation and partitioning in soybeans exposed to low root temperatures. *Plant Physiol* 1986;**80**:249–55. https://doi.org/10.1104/pp.80.1.249.

Wibberg D, Price-Carter M, Rückert C *et al.* Complete genome sequence of ovine *Mycobacterium avium* subsp. *paratuberculosis* strain JIII-386 (MAP-S/type III) and Its comparison to MAP-S/type I, MAP-C, and *M. avium* complex genomes. *Microorganisms* 2020;**9**:70. https://doi.org/10.3390/microorganisms9010070.

Wickham H, Averick M, Bryan J *et al.* Welcome to the tidyverse. *JOSS* 2019;**4**:1686. https://doi.org/10.21105/joss.01686.

Winship LJ, Martin KJ, Sellstedt A. The acetylene reduction assay inactivates root nodule uptake hydrogenase in some actinorhizal plants. *Physiol Plant* 1987;**70**:361–6.

Zdyb A, Salgado MG, Demchenko KN *et al.* Allene oxide synthase, allene oxide cyclase and jasmonic acid levels in *Lotus japonicus* nodules. *PLoS One* 2018;**13**:e0190884.

Zhang X, Sherman DM, Sherman LA. The uptake hydrogenase in the unicellular diazotrophic cyanobacterium *Cyanothece* sp. strain PCC 7822 protects nitrogenase from oxygen toxicity. *J Bacteriol* 2014;**196**:840–9.

Zhang Z, Torrey JG. Studies of an effective strain of *Frankia* from *Allocasuarina lehmanniana* of the *Casuarinaceae*. *Plant Soil* 1985;**87**:1–16.

Received 20 May 2024; **revised** 7 October 2024; **accepted** 29 October 2024